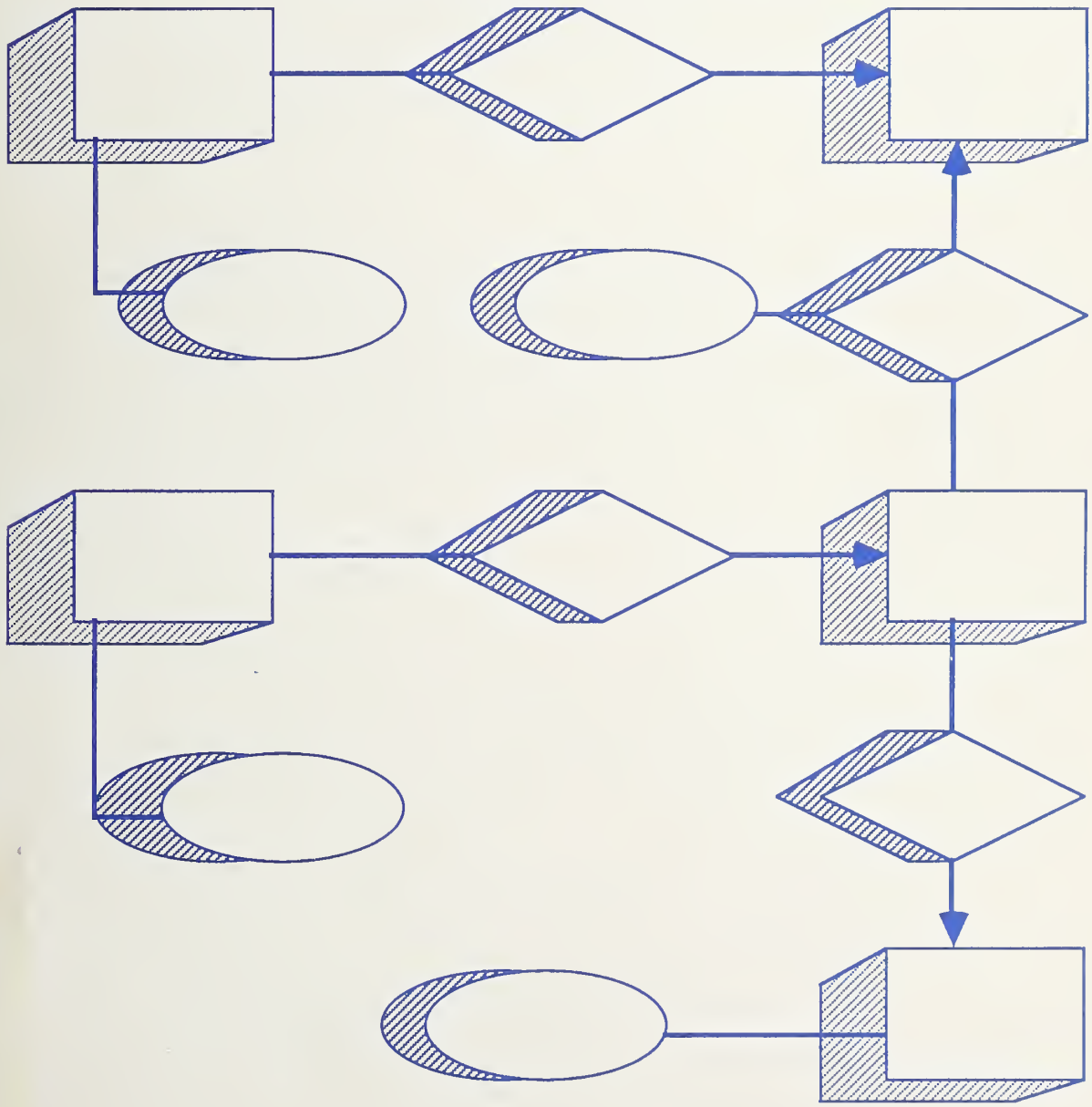
National Institute of
Standards and Technology
(formerly National Bureau of Standards)

nputer Science
Technology

NIST Special Publication 500-159

Data Administration: Management and Practice
Proceedings of the First DAMA Symposium

Judith J. Newton and Frankie E. Spielman, Editors



Computer Science and Technology

NIST Special Publication 500-159

Data Administration: Management and Practice

Proceedings of the First DAMA Symposium

Judith J. Newton and Frankie E. Spielman, Editors

Information Systems Engineering Division
National Computer and Telecommunications Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899

Sponsored by:

National Capital Region of the Data Administration Management Association (NCR DAMA)

Federal Data Management Users Group (FEDMUG)

Association for Federal Information Resources Management (AFFIRM)

October 1988

NOTE: As of 23 August 1988, the National Bureau of Standards (NBS) became the National Institute of Standards and Technology (NIST) when President Reagan signed into law the Omnibus Trade and Competitiveness Act.

U.S. DEPARTMENT OF COMMERCE
C. William Verity, Secretary

National Institute of Standards and Technology
(formerly National Bureau of Standards)
Ernest Ambler, Director

Library of Congress Catalog Card Number: 88-600586
National Institute of Standards and Technology
Special Publication 500-159, 122 pages (Oct. 1988)
CODEN: XNBSAV

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 1988

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402

FOREWORD

This document represents the proceedings of a one-day symposium at the National Bureau of Standards on May 17, 1988. It was the first in what we hope will become an annual series of symposia on the subject of data administration. As more and more organizations recognize the need to treat data as a corporate resource, data administration is gaining acceptance as an important area of specialization for information processing professionals.

The symposium was jointly sponsored by the National Capital Region of the Data Administration Management Association (NCR DAMA), the Federal Data Management Users Group (FEDMUG), and the Association for Federal Information Resources Management (AFFIRM). We wish to thank the following individuals for their commitment to and assistance with the symposium:

Jim Clancy, AFFIRM
Alice Cohen, DAMA
John Coyle, AFFIRM
Rene Fecteau, DAMA
Tina Knoeller, DAMA
Mary Lou Melley, DAMA
Tammar Paynter, DAMA
Ronald Shelby, DAMA
Rae Thompson, DAMA

We also wish to express our gratitude to the speakers, session moderators, and participants who made this symposium possible, and to James H. Burrows, the Director of ICST, for his fine welcoming talk.

With a few exceptions, the papers in this proceedings represent manuscripts submitted to the editors for publication. Those talks for which no manuscript was submitted have been summarized by the editors from audio recordings and are marked "Summary of Remarks."

Because the speakers in the symposium drew on their personal experience and knowledge, they may have expressed views which do not necessarily reflect those of the National Bureau of Standards, DAMA, or AFFIRM. Additionally, they sometimes cited specific vendors and commercial products. The inclusion or omission of a particular company or product does not imply either endorsement or criticism by NBS, DAMA, or AFFIRM.

Judith J. Newton
Frankie E. Spielman
Co-Chairs

CONTENTS

KEYNOTE ADDRESS

INFORMATION ASSET MANAGEMENT

Dan Appleton
The D. Appleton Company

INFORMATION ASSET MANAGEMENT

SUMMARY OF REMARKS

Dan Appleton
The D. Appleton Company

Information asset management is data-driven information resource management. An **asset** is reusable. One may invest in an asset, then use it to increase productivity. The quintessential information asset is data.

The information management function is a 'business within a business.' It must deliver products to customers. This represents a value shift in the system as it relates to data processing.

Processes are what people do; they drive technology. Processes must change before technology can have its promised effect. Change must occur on three levels: values, processes, and technology sets (fig. 2).

In 1968, it was possible to produce a system single-handed. Programmed in COBOL, a typical system took six months and cost about $100,000. That year, the data processing department developed seven new applications on the same order. Today, however, systems average $5 million, take over three years to develop and may involve 100 people. Though more code is produced, the concept of productivity has changed.

The concept of the chief information officer (CIO) has also changed ideas about productivity by introducing the **integrated environment.** This is perceived as a desirable improvement over the old system of segregated systems, even when it takes much longer to produce a product.

Another major value shift has occurred in regard to changes in systems as requested by users. Previously, a user would submit a change request to the data processing department and then wait until a change was produced. Now the users want to make the changes themselves.

New technologies such as relational database management systems have affected data management (fig. 3). Data modeling has brought a new concept to data processing: it has taken data out of the process context. Normalization even without relational databases is now seen as fundamental. Business rules have been the missing link in data management. Data dictionaries, renamed repositories or encyclopedias, are still seen as a hope for the near future.

All these new technologies are leading to data independence; the separation of data and process. Data is "leveragable" - reusable from process to process.

As more new technologies become available, they will lead to even more data independence. Object-oriented systems and artificial intelligence products which employ knowledge bases will strengthen the trend.

How can data administrators deal with new requirements? The traditional systems life cycle takes too long for big systems. There are three alternatives: use off-the-shelf, reusable modules; modify existing code; build new from scratch. The first alternative holds the most promise for the future, but it requires adoption of a new set of values. A new approach to defining requirements, such as designing the 'shelf' and how to put the data there, is needed. In the process, data administrators should recast themselves as **data asset engineers.**

The concept of 'an information' is important to the understanding of data administration. An information is the result of combining one or more data. Four hundred basic data can be combined to produce 10^{869} informations. A datum may be defined as fact plus stimulus (meaning). This notion of the ontological structure of data (that is, its meaning) allows the data administrator to leverage those meanings through the application of **business rules.** These rules are changing constantly, but managing at the data element level without the application of business rules is like trying to manage a beach one grain of sand at a time.

In figure 6, business rules may be applied to the conceptual view of the data. External schemas may be seen as requirements arriving from the user, and the internal schemas as the 'shelf' from which the reusable modules are pulled.

Figures 9 and 10 show how data engineering is a continuous process in the new system life cycle which will allow us to march, flags flying, into the bright new future of data asset engineering.

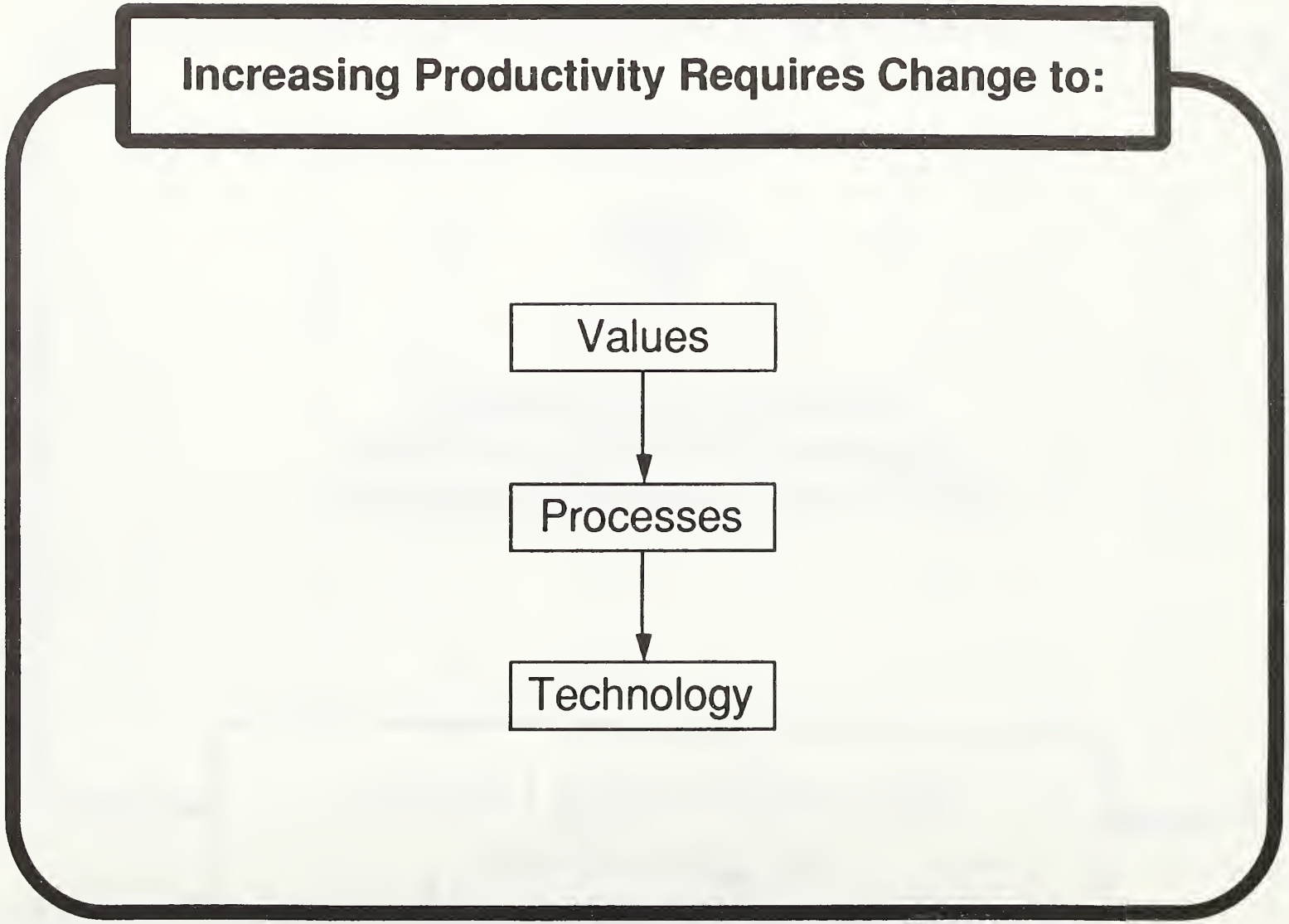
Mr. Appleton is President of D. Appleton Co. Inc. (DACOM), a firm that specializes in industrial modernization and data resource management methods and tools. He has many years of experience in Strategic Business Planning, Management Information Systems, and Systems Development. Mr. Appleton has published numerous technical papers and articles on manufacturing automation and database management.

Information Asset Management

Data Administration Symposium
National Bureau of Standards
Gaithersburg, Maryland

May 17, 1988

Figure 1



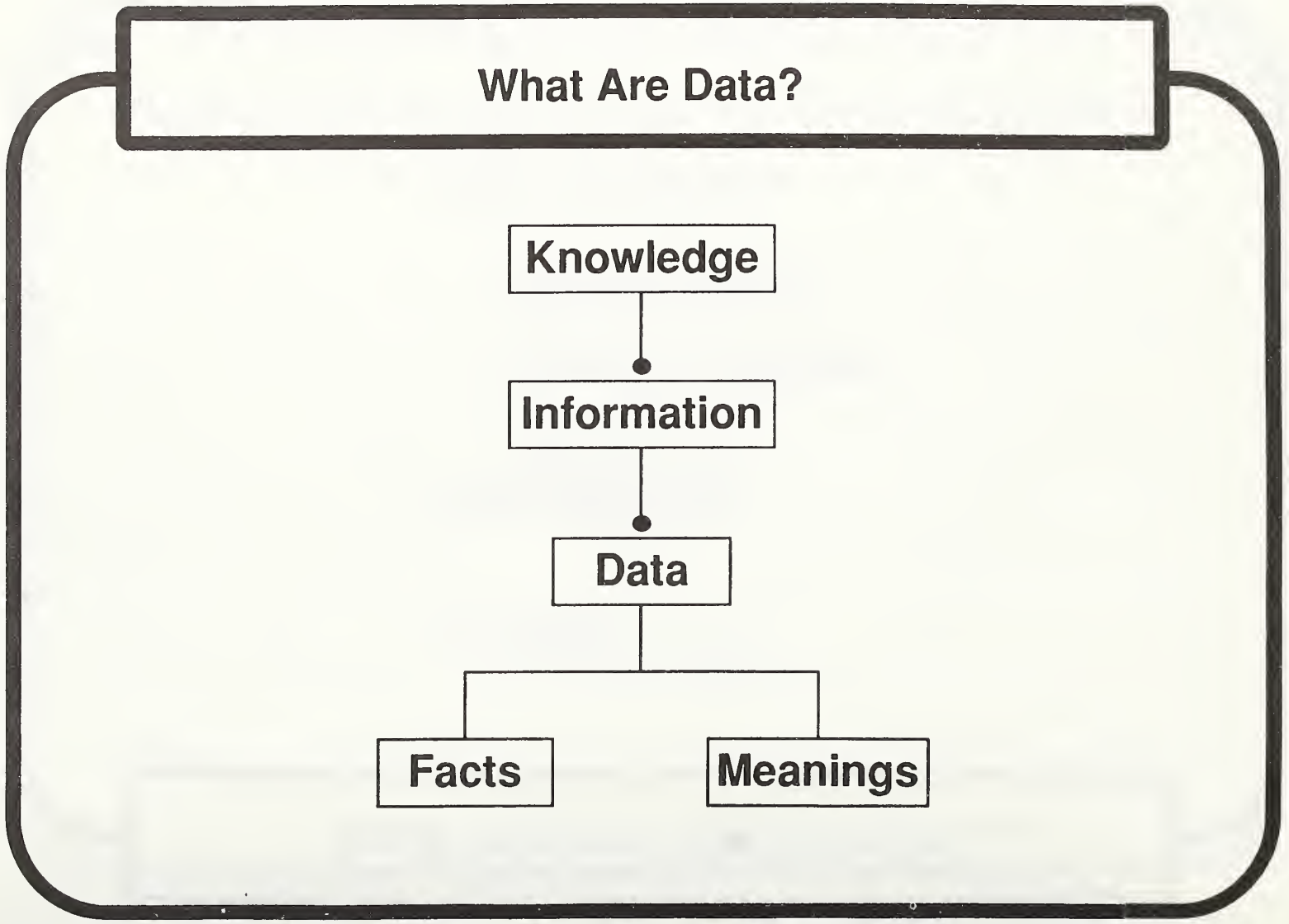
419-2

Figure 2

Basic Relational Technologies

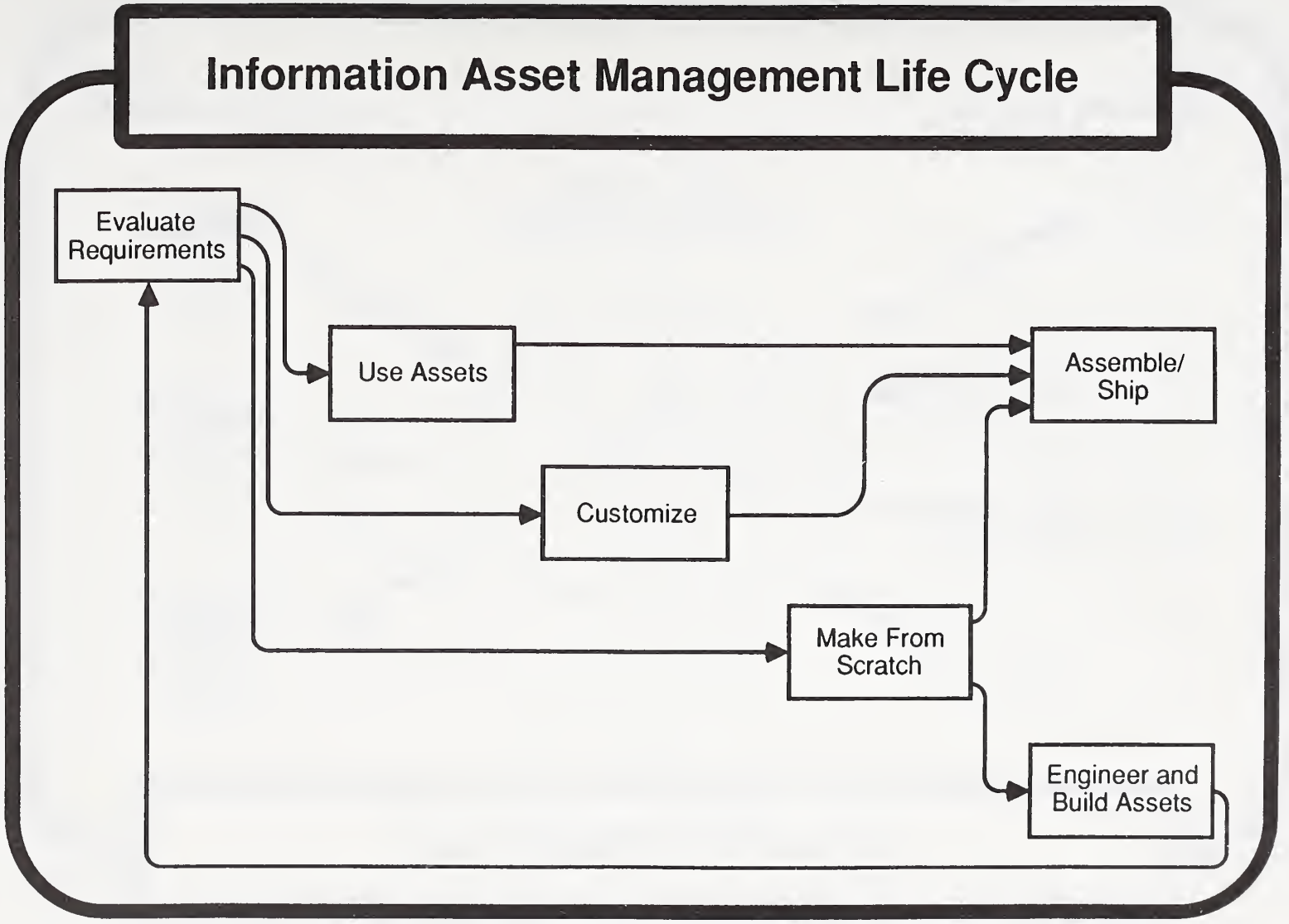
- RDBMS
- Data Modeling
  - ER Models
  - Normalized Models
  - Business Rules
- Data Dictionaries

Figure 3



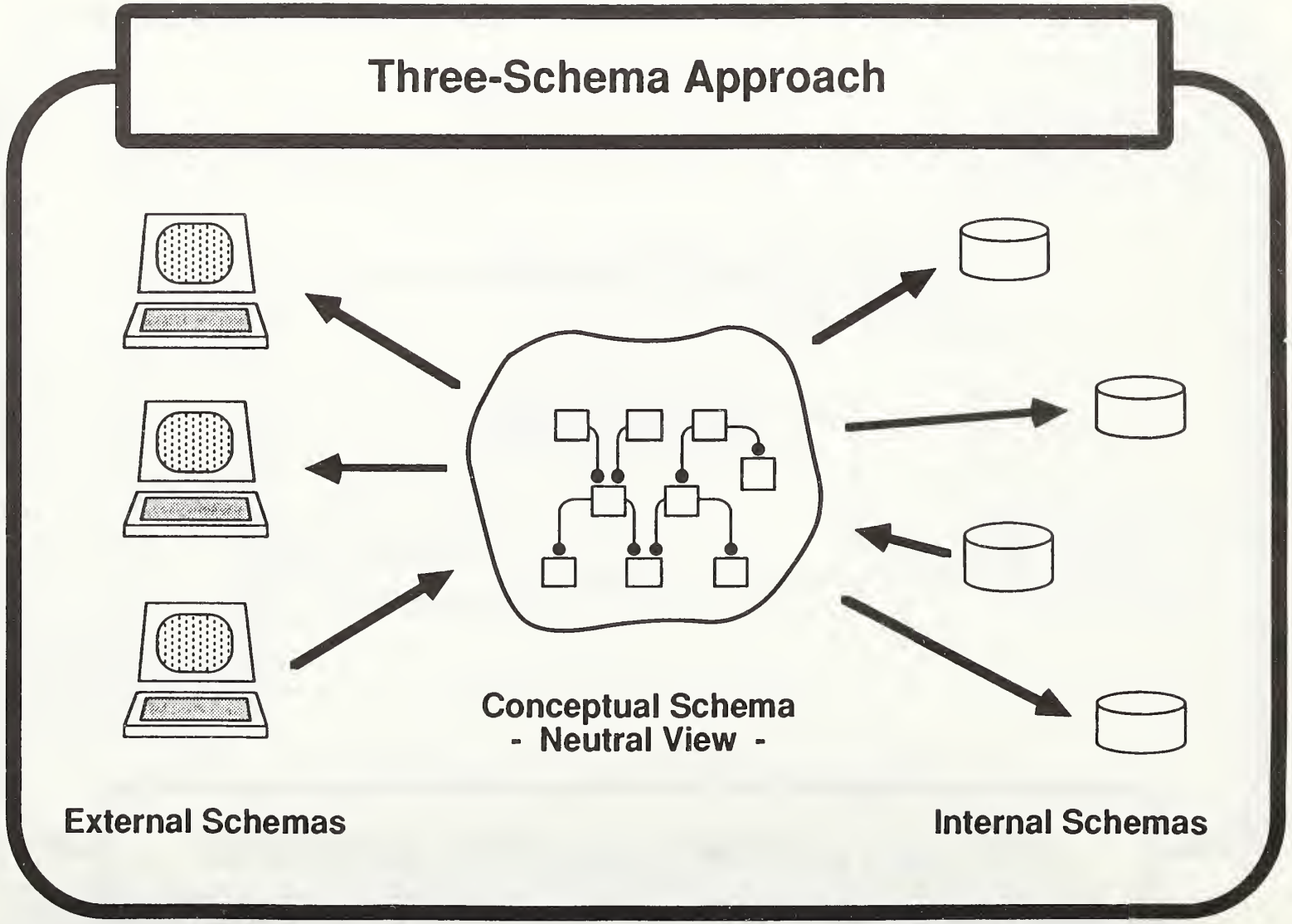
419-4

Figure 4



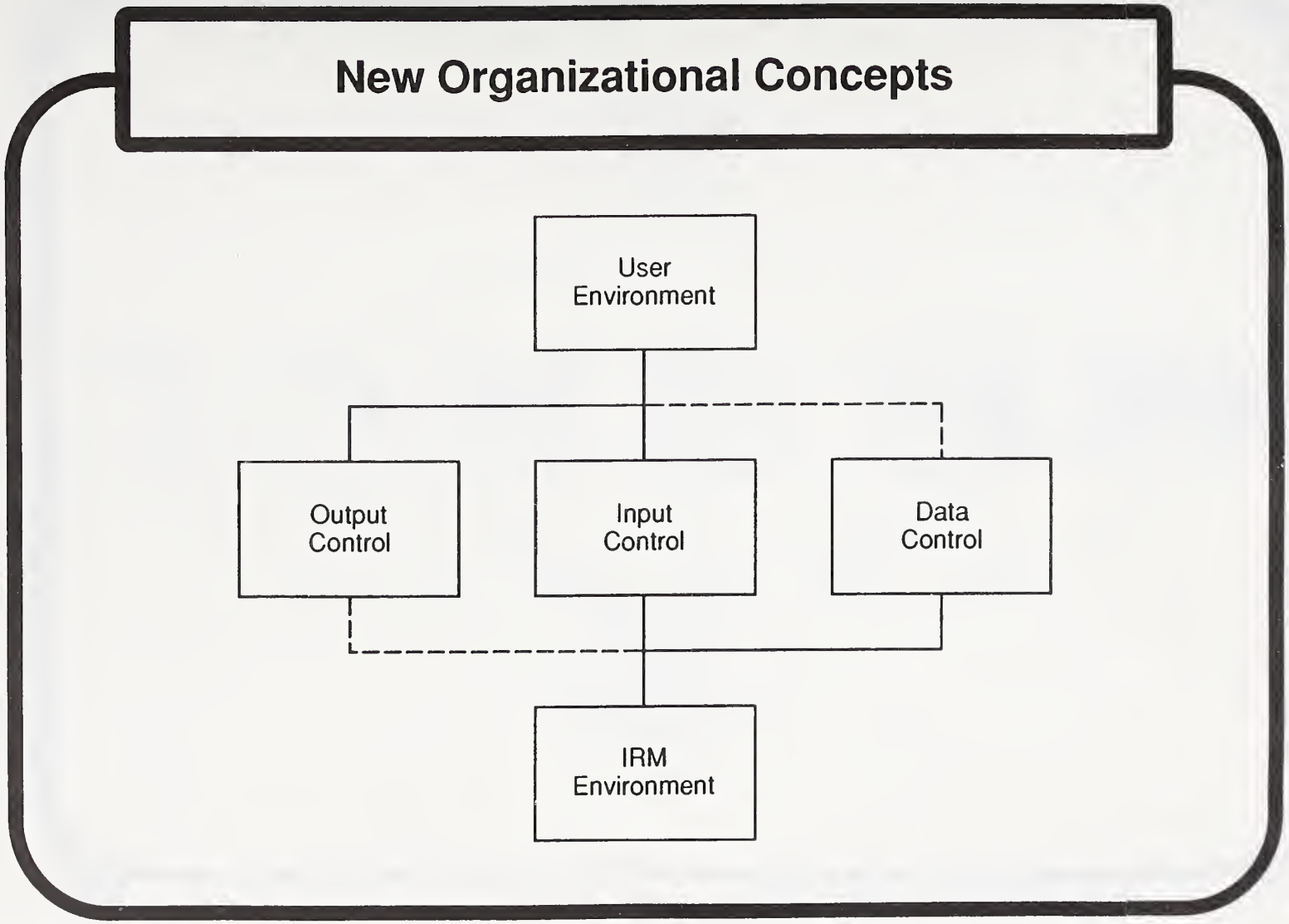
419-5

Figure 5



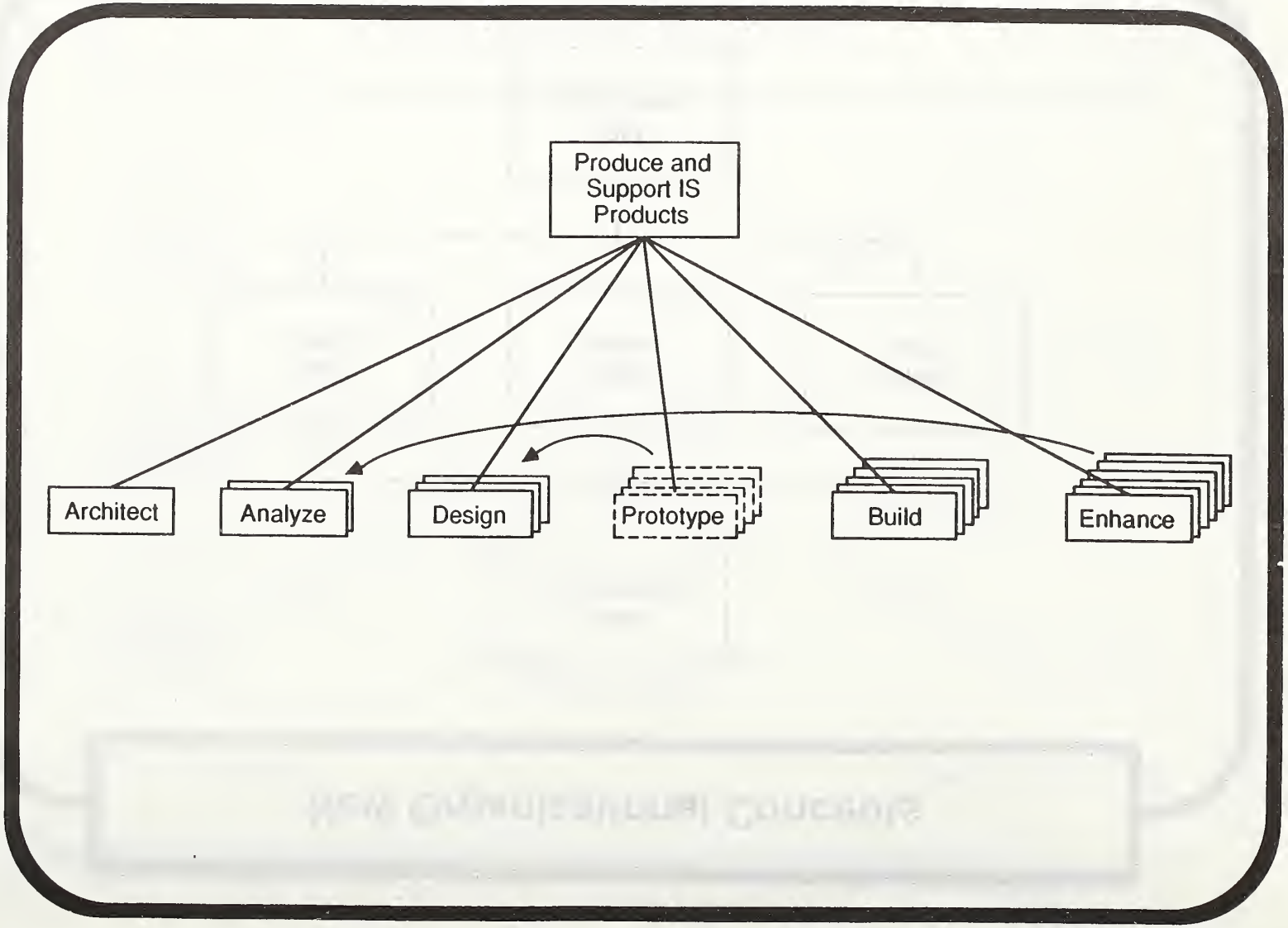
419-6

Figure 6



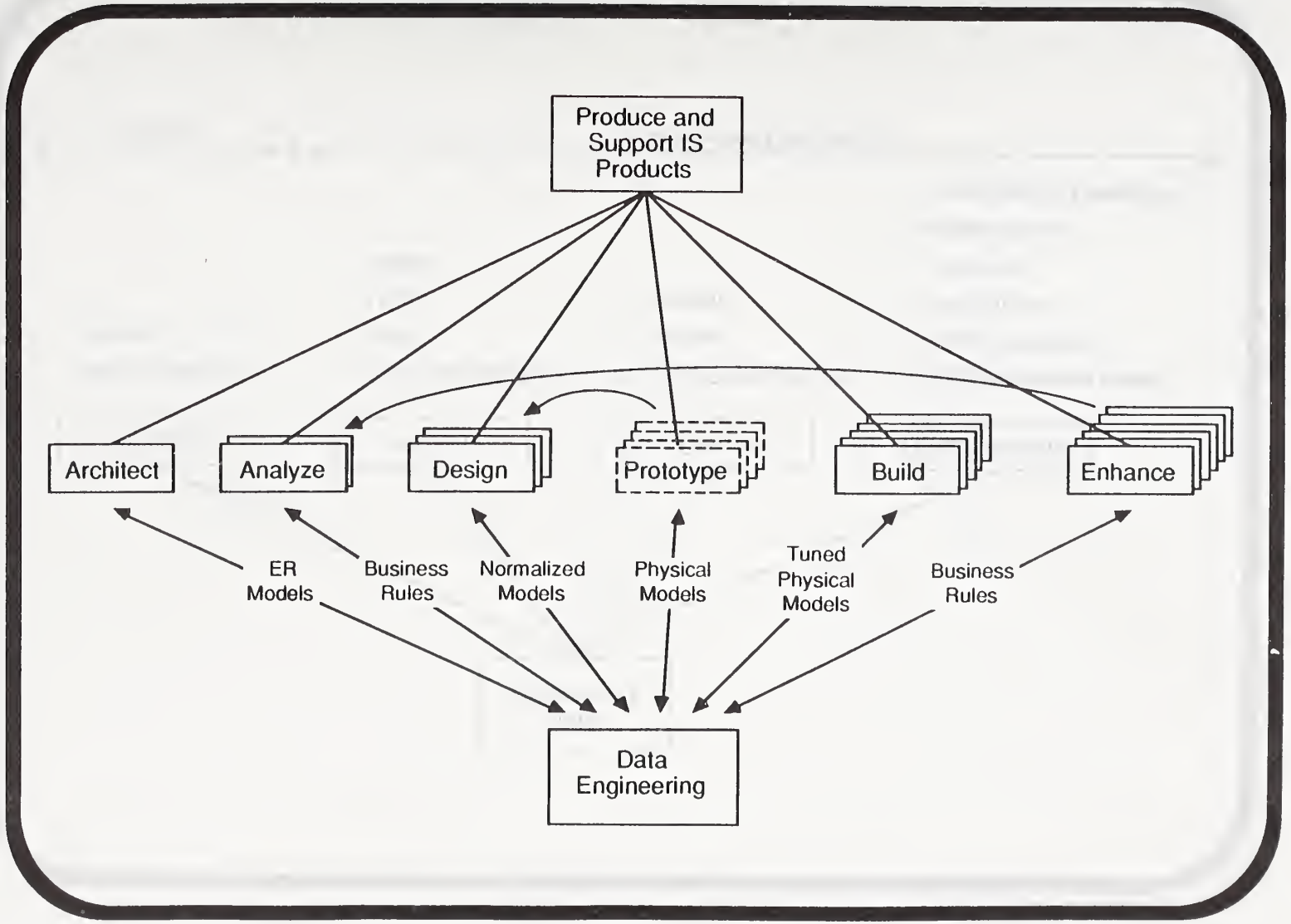
419-7

Figure 7



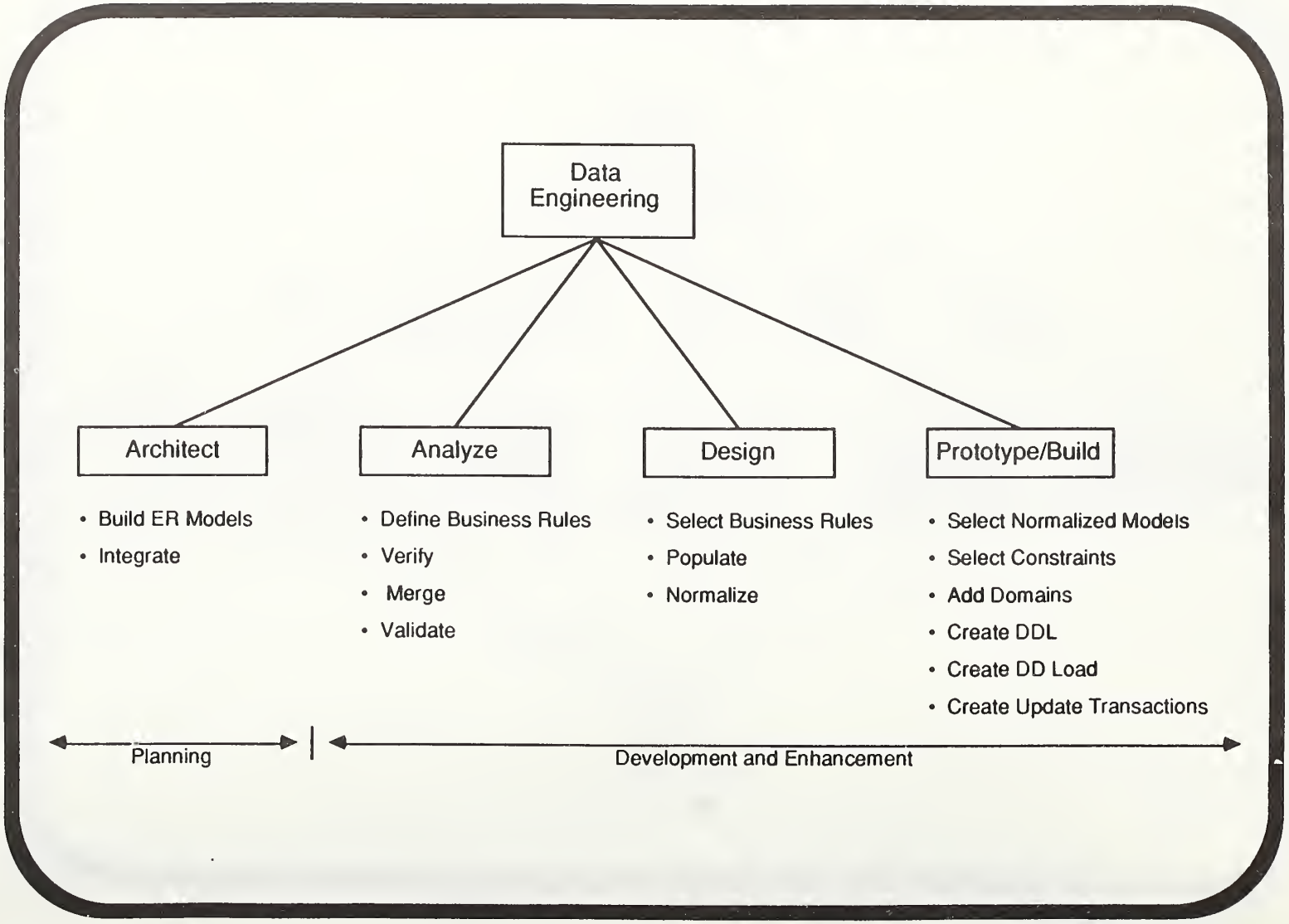
419-8

Figure 8



419-9

Figure 9



419-10

Figure 10

DATA ADMINISTRATION

AND THE

SYSTEMS DEVELOPMENT LIFE CYCLE

MODERATOR

Rene Fecteau
Veterans Administration

PANELISTS

Rick Barron
Thomas J. Bergin
Ellen J. Levin
Mary Ann Wallace

DATA ADMINISTRATION AND THE SYSTEMS DEVELOPMENT LIFE CYCLE

Rick Barron
McDonnell Douglas

Good Morning. I am Rick Barron with McDonnell Douglas and I am delighted to be here this morning to discuss Data Administration and the Systems Development Life Cycle at McDonnell Douglas. McDonnell Douglas is a large international, multi-divisional company. Within McDonnell Douglas, I am in a portion of the company called MDAIS or McDonnell Douglas Aerospace Information Services Company. This part of the company provides internal information services support to the aerospace components of McDonnell Douglas. It is a separate division and company from the information services that are offered to commercial customers.

Within McDonnell Douglas Aerospace Information Services Company I head an organization called Application Support which is part of the Professional Services Division of the Company. Heading this organization puts me in an excellent position to comment on today's topic since I have functional responsibility in both St. Louis and the West Coast for Application Data Management Services and an organization called Application Productivity Services. The Application Data Management Services is responsible for database management and data administration support across the corporation and the Application Productivity Services is responsible for programmer productivity, productivity tools, productivity measurement, and software quality assurance. Thus, I own the tools to support data administration and the tools to support application productivity in the Systems Development Life Cycle.

What do the following words have in common?

- Oxymoron
- Data Administration
- Data Driven Methodology
- Computer Aided Software Engineering (CASE)
- Relational Database or Distributed Relational Database
- Artificial Intelligence (AI)
- Three Schema Architecture

No, I am not trying to imply that each of those are oxymorons. Rather, each of these words seem to be in vogue today. I don't think I have heard a speech in the last six months or a year that didn't work these phrases somewhere into the speech. See, I have fulfilled my commitment by already mentioning them now.

Actually, these are all topics except for oxymoron, of course, that everybody seems to salute, have a lot of focus on and feel are the key to the future. Obviously, that is why we are all here today. I can't tell you the number of times I go out to a user organization and I am told that the user needs a relational database. I don't understand why the user cares how the data is physically stored internally. The same thing is true with data administration. I haven't decided whether there are a few data administrators or 106,000 data administrators within McDonnell Douglas. It depends how narrow or how broad you want to make the term. I think every new project we are starting is using data driven methodology, preparing to implement three schema architecture, using new integrated CASE tools, and working on Artificial Intelligence in their implementation of their distributed relational database. All of this is true if you listen to the project managers. If only the practice and commitment were as advanced as the vocabulary, we would be in great shape.

More specifically, what do the products PM/SS from Adpac Corporation and Arthur Anderson's Foundation have in common? Those are both products that have been requested by parts of my organization to evaluate. With two of these products, I have received requests from both St. Louis and the West Coast to evaluate them. That wouldn't be bad at all, in fact I had felt great about the fact that the two parts of the organization are becoming more common, if it weren't for the fact that in one center it is the data administration folks asking for the product evaluation while at the other center it was the systems development life cycle folks asking for the evaluation. It reminds me of the commercials "is it a candy mint or a breath mint, tastes great or less filling," etc.. It just highlights the fact that there is tremendous overlap between the systems development life cycle and data administration discipline.

More specifically, what do the following projects going on internally within McDonnell Douglas have in common: Corporate Data Administration Project as well as our component data administration projects, Product Definition Data, Integrated Composite Center, CALS (Computer Aided Logistic Support), AIM (Artificial Intelligence for Machining), Advance Business Management Systems (ABMS), etc., etc., in fact, the list goes on. Each one of those projects, as part of its activity, is looking at repositories, data administration, further integration in sharing data across disciplines, and the methodology to make it all happen. The methodology will dictate the CASE tools to be used to integrate with the repository selected for each project and, therefore, with the systems development life cycle automation processes. Each one of those projects thinks that the other projects will integrate with it. They all agree on the common goal. They all have data administration, they all think they are data driven instead of process driven, yet each one

considers a critical success factor its ability to define the methodology and the processes and believes it can't wait on global solutions if it is going to make its dates.

I was going to leave the labels off the picture I am showing here (fig. 3) and could probably talk to it for four days, constantly putting different things in the center circle, the outer circle and the sections in between. However, whether you call it the dictionary, of which we have 4,000 it seems, whether a repository or whatever name you wish to give it, it is in the center hub. The life cycle surrounds it, while encompassed in turn by the methodology. There, you see, I have solved all your problems. A nice simple solution.

Actually, it's a lot more complex than this picture shows. The data administrator is included as part of maintenance or a global support function. It could be shown as a ring on the inside. In fact, that is where I suspect it would be.

The bottom line is that the goals are the same. We need to be able to provide users with quicker, easier access to data. Both the systems development life cycle and data administration are enabling functions to make that happen. We all agree that we need to treat data as a corporate asset with more than just lip service. We need to have a breakthrough reduction in the systems development life cycle development by an order of magnitude. We need to get to reusable everything, code modules, analyst design elements, data and data definitions, etc. We need to support the concepts of the three schema architecture and separate data from the applications so that changing data stores and locations of where data is stored and database managers require no maintenance from the application perspective. We have to have high portability. We are getting a lot more complex, especially in our industry where you have secure projects that have to put up their own processors. You may have to port large Information Management System (IMS) applications onto hardware platforms that cannot afford to run on large IMS systems. The bottom line is to increase the availability and reliability of information for our users while reducing their cost of that information. Easily said, not so easily accomplished.

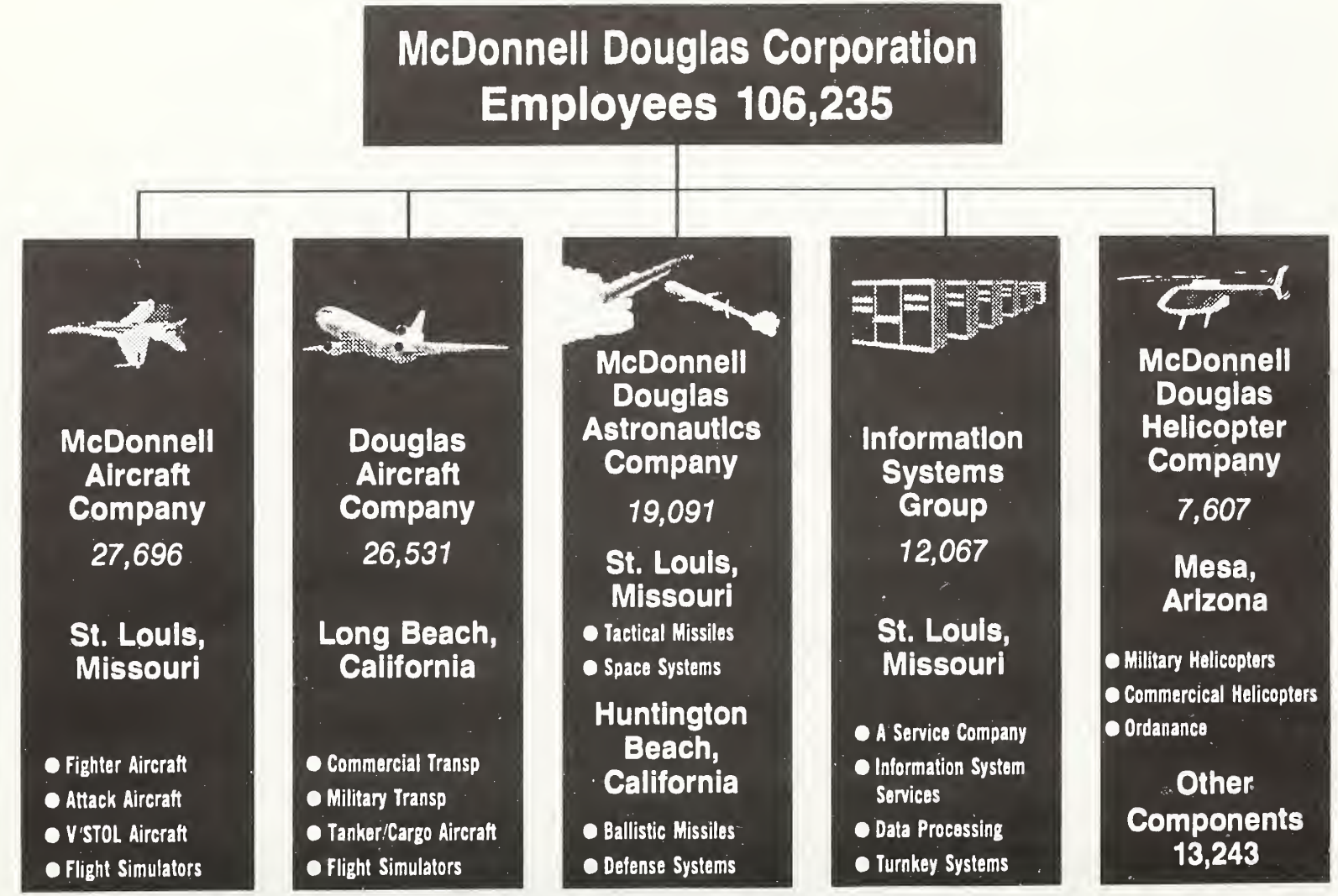
What I leave you with for consideration is that the political issues may be as great as the technological issues. When I look at both the data administration and the systems development life cycle discussions within McDonnell Douglas and around industry, we are arguing the "hows" of the implementation and getting frustrated when the tools that we want aren't there to meet our requirements. But, I am not sure we really understand all the "whats." Secondly, I see that middle management is a major inhibitor whether putting in programming tools or using data administration. They resist giving up their control or custodianship of data to a common source. They are deeply

entrenched and don't want to let go. We need to focus more time on solving that problem and not worrying about the fact that the technology is not there yet. The fact that we have many different data dictionaries popping up doesn't bother me if it facilitates resolving the political issues and changing mind set. What does bother me is if we keep evaluating the next one and next one because the 23 in house don't meet requirements. That is not productive. Facilitate the political solutions but don't proliferate the software.

That gets to our next point, we need implementation functions in step. Look at the Personal Computer (PC) and Video Cassette Recorder (VCR) market. Every week something is better, more functional and cheaper. You can't wait to make a decision because you get more functionality and lower cost the next day. You can't wait because every time you think that the function that was on your "must have" list was there you realize that three more "must have's" are just around the corner. You can't put in 412 different platforms because each week when a specific user wants to make an acquisition, they can succeed in buying something different because what is available then is better, faster, more functional and cheaper than what you have already got on board. Maintenance and support costs will become a nightmare. You won't get to common data in that process.

As more and more integrated tools and vendors enter the market the problem is going to get worse in the short term. I think about every six nanoseconds another new integrated dictionary and life cycle tool is announced. The key to success though is to keep the ball moving forward. Keep progress going. I would rather have the projects I listed earlier within McDonnell Douglas each talking data administration, common data methodology and changing that mind set solving some of the political problems, better helping us understand what works and doesn't work and what our needs are for repositories over time than the alternative of doing nothing. It is not ideal, but it is moving the ball forward. This is a very exciting and challenging area. It is a lot of fun. The next evolution of change in all these areas is getting our dreams to become reality, and not just talk. The glossies 15-20 years ago used the exact same words and had the same promises. Let's plan it carefully, not leave bad impressions in our minds or others that may inhibit reaching our common destiny, and we will have breakthroughs in productivity through both Data Administration and automating the System Development Life Cycle.

Mr. Barron is Director of Application Support at McDonnell Douglas Aerospace Services Company. He has given seminars, lectures and taught classes in the United States, Canada and Europe. He is currently an officer of GUIDE International (IBM user group).



GP71-1292-13-G

Figure 1

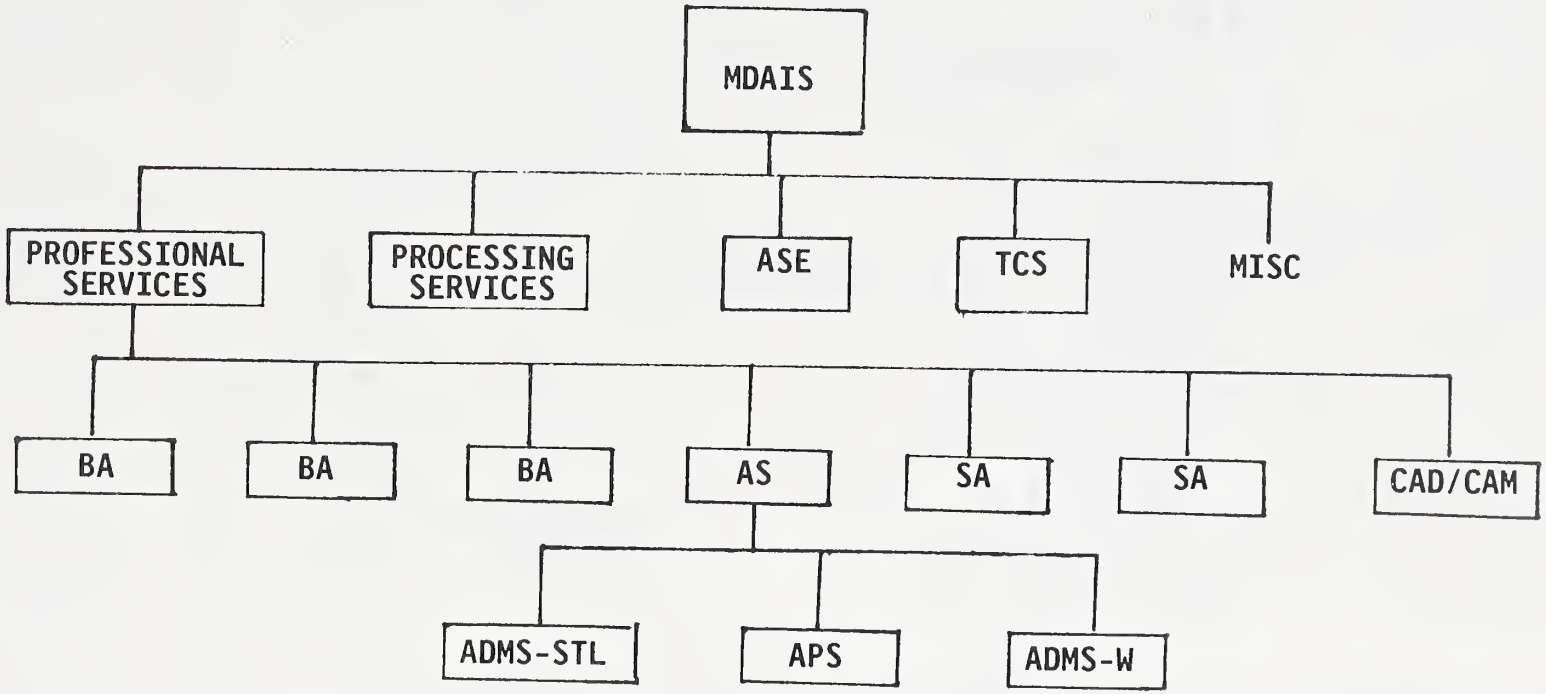

MDAIS
PROFESSIONAL SERVICES
PROCESSING SERVICES
ASE
TCS
MISC
BA
BA
BA
AS
SA
SA
CAD/CAM
ADMS-STL
APS
ADMS-W

Figure 2

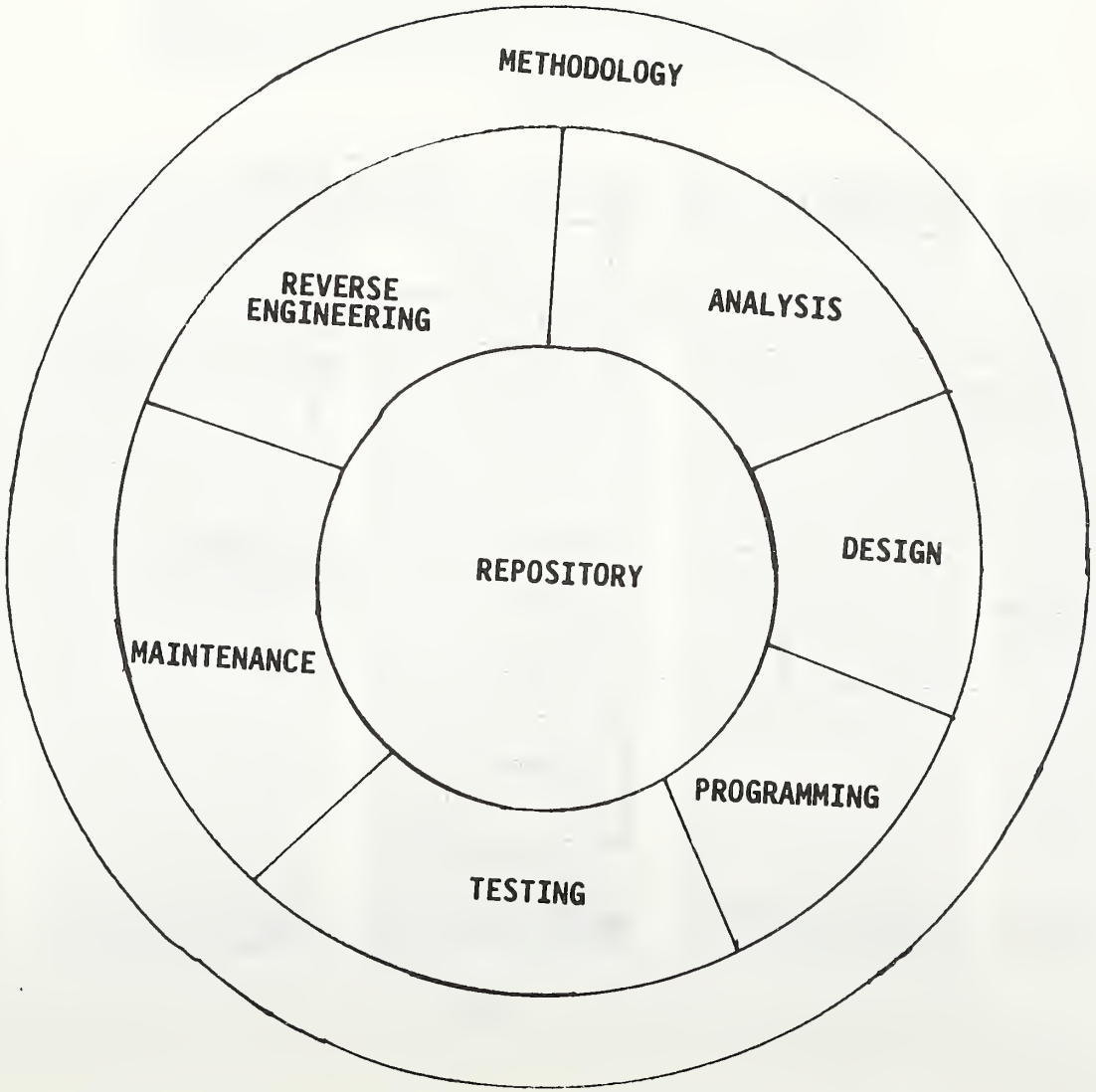

METHODOLOGY
REVERSE ENGINEERING
ANALYSIS
DESIGN
REPOSITORY
MAINTENANCE
PROGRAMMING
TESTING

Figure 3

DATA ADMINISTRATION AND THE SYSTEMS DEVELOPMENT LIFE CYCLE

Thomas J. Bergin

The American University

The Promise of Data Administration

The first time I heard the term "data administration" was at the Veterans Administration (VA) in the Fall of 1974. I was made project leader of the Data Administration Project! Our first job was to inventory the data elements (fields?) maintained in manual and automated systems. The Reports and Statistics Service of the Comptroller's Office had estimated that there were about 8,000 such elements. Another office, Management Engineering and Evaluation (which had responsibility for forms management) had estimated 12 to 14,000 data elements. We decided that we would look at automated systems of records first, and then worry about manual systems. Three months later, we had counted over 56,000 data elements in automated systems of records. We stopped! There was no need to increase the accuracy of bad news.

Like every other large user of computers, the Veterans Administration had created numerous automated systems. Each project team was separate, and worked with different users. The members of each project had different backgrounds: some had been functional users, others had a few years of systems work (1401, 7080 and assembly languages), and some were new to the business of computers and computing. The result was that each data processing system (either manual or automated) was created independently. Yes, we tried to coordinate efforts between projects in the same functional area, but....

The net result was that by 1974, the VA had a great number of manual and automated systems to capture, manipulate, and retrieve information. As I recall, we had 38 places where we stored the veteran's name and address. Depending on which system you were in, there were three different meanings for "SSN": Service Serial Number (Army identification number, Navy...), Station Serial Number (for hospitals), and Social Security Number. We weren't alone! The rush to use computers resulted in similar situations in all organizations.

I also remember the first time I set foot on the National Bureau of Standards campus. In the Summer of 1974, Harry White, the Director of the Federal Information Processing Standards (FIPS) Program, asked me to Chair TG-17, the NBS Task Group on Data Element Dictionaries. It was an exciting task, thinking about data as an entity. I recall sitting around a large table with 12

or 13 people from other Federal agencies talking about "data dictionaries" and "data directories", and what the differences were between them. I remember hearing a colleague from NSA using the term "metadata" for the first time. In March 1975, I nervously gave a lecture on "Data Administration in the Federal Sector" in this very auditorium.

The Promise of Life Cycle Management

The term Life Cycle Management (LCM) came into vogue a few years ago in an attempt to amalgamate the existing Systems Development Life Cycle (SDLC) with the planning and budgeting processes. In a sense, we can think of LCM as an attempt on the part of management to pull control of the SDLC back toward the functional user. At the same time, Yourdon, DeMarco, and Constantine provided technical management with new methodologies and techniques, in an attempt to make the Systems Development Life Cycle more rigorous.

The reasons for these efforts were obvious years ago. The problems which triggered them still exist today: application development projects typically cost more than estimated, take longer than anticipated, and do less than expected. When we couple these criticisms with the systems development backlog, we can understand the rationale behind efforts to improve the planning, analysis, design, and development of automated systems.

Reality in Large Complex Organizations (Has Anything Changed?)

Life Cycle Management today is little better than it was 10 or 15 years ago. It is focused on the creation and maintenance of individual (stand-alone) systems. Analysts are still using out-of-date methodologies and tools. The application plans are still tightly coupled to hardware and software. Management still focuses on the short-term costs and benefits of individual applications or application areas. The systems development process is still slow, ponderous, and resistant to change.

Indeed, a recent article in Infosystems stated:

> Despite what adherents of structured methodologies consider obvious benefits, less than 10 percent of the 750,000 programmers in this country are estimated to use these techniques.

The bottom line is that most organizations have out-of-date Life Cycle Management policies and procedures. I have seen life cycle manuals that still contain flowcharts. Worse still, some of the flowcharts show punched cards. Systems Development organizations are not using structured methods such as data dictionaries, data flow diagrams, and entity relationship diagrams. They are doing things much the way they did them twenty years ago. Data

Administration is even harder to characterize. Today, we have few data standards, poor data quality, and high data redundancy. Most large organizations do not know how much data they collect, what they do to it, and why they need it. Few organizations are serious about managing data, or have well-staffed data administration programs. Indeed, some organizations are still wrestling with the meaning of data administration. To paraphrase Justice Potter Stewart:

> data administration is a little like pornography; it is difficult to define, but everyone is sure that they will know it when they see it.

This past summer, I worked with Software Solutions, Inc., to re-write and update the Life Cycle Management policies and procedures for the Manpower, Personnel and Training arena of the Department of the Navy. In so doing, I called a number of Information Resources Management (IRM) organizations around the city. In each case, there was a small staff working at the policy level. There seemed to be limited understanding of the need for data resource management, data standards, or data dictionaries, as a tool for implementing a data administration program. This finding is consistent with the Caudle Report on Federal IRM.

<u>On a Positive Note</u>

At the Department of the Navy, on the other hand, there is a concerted effort to make data administration a reality. There is a sophisticated, and well staffed, organization in place. A data standards program exists as do data standards. A data naming standard was recently promulgated and is in use. An information resource encyclopedia (data dictionary/directory) has been populated and is being used in many phases of the systems development life cycle.

As I mentioned above, the Life Cycle Management policies and procedures have been re-written. When these policies and procedures are officially implemented, the use of these data administration tools will be part of the applications development process.

Although much of this is new, there are already some success stories. The Department of Defense Documentation Center needed permanent change of station information. Through a "data issue resolution process" the data needs were evaluated and it was determined that available data could be used to meet the new need. Had this problem been examined in a purely technical context, there is no doubt that the Navy would have an additional data processing application under development.

Observations and Conclusions

To help map the reality of the present to the promises of data administration and life cycle management, the following observations and conclusions are offered:

1. Neither data administration or life cycle management have changed much since the 1970's. Evolutionary changes take forever. A revolution is needed in the way we develop applications, and in the way we think about data. The focus of life cycle management must shift from the management of technology to the management of data.

2. Data must be recognized as an organizational resource. This has to become more than just a platitude. Data Administration must start to manage data from an organizational perspective, and not from the limited perspective of specific functional units or applications.

3. System Development organizations must recognize that quality data is crucial to any efforts to improve the systems development and life cycle management processes. Recent approaches to improving the systems development process, such as prototyping, fourth generation languages, and subject area databases all require quality data.

4. Data Administration must take the lead in modifying the Life Cycle Management policies and procedures, to incorporate data administration methods and techniques.

5. Systems analysts, systems designers, programmers and others in the systems development process must become familiar with, use, and appreciate data element dictionaries and other data administration processes, techniques, and tools.

6. Data Administration must recognize that Life Cycle Management is the vehicle for institutionalizing data administration and data resource management.

These are difficult prescriptions. They will require risk taking, time, and considerable resources. It is only though these processes, however, that data administration theory and promise will become reality.

Dr. Bergin served at NBS in 1974-76 as first chairman of Task Group 17, Data Element Dictionaries. In 1982, he became the Director of the Quantitative Teaching and Research Lab at American University. In 1983, he was Senior Technical Specialist to a joint American University-Institute of Public Administration advisory group to the Government of Bangladesh. He is co-author of Microcomputer Based Primer on Structural Behavior (1966).

LIFE CYCLE MANAGEMENT TODAY:
==========

PROCESS DRIVEN: APPLICATIONS

USERS

TECHNOLOGIES

SLOW AND PONDEROUS

TIGHTLY COUPLED TO OUT-OF-DATE:

HARDWARE

SOFTWARE

METHODOLOGIES

TOOLS

PROCEDURES & STANDARDS

PRESENT ENVIRONMENT:

RELIANCE ON EXISTING APPLICATIONS

COST/BENEFIT SENSITIVE

LACK OF ORGANIZATIONAL PERSPECTIVE

Figure 1

DATA ADMINISTRATION TODAY:

POOR DATA QUALITY

HIGH DATA REDUNDANCY

ACCOUNTING PARADIGM

TECHNOLOGY DEPENDENT

POORLY RESOURCED

SLOW RATE OF ADOPTION

EMERGING METHODOLOGIES:

- DATA STANDARDIZATION
- DATA ELEMENT NAMING
- DATA QUALITY ASSURANCE
- DATA ENGINEERING
- DATA MANAGEMENT TOOLS

GROWING AWARENESS ACROSS DISCIPLINE

Figure 2

WHY LCM NEEDS DA:
==========

QUALITY DATA

IS THE KEY

TO QUALITY SYSTEMS

MOVE AWAY FROM TECHNOLOGY AS FOCUS

PROMOTE IMPROVED FUNCTIONAL CONTROL

MAKE NEW TECHNOLOGICAL TOOLS WORK:

PROTOTYPING

C.A.S.E.

4GLs

DBMS & DDBMS

ENTERPRISE NETWORKING

OFFICE AUTOMATION

Figure 3

WHY DA NEEDS LCM:

TRANSFORM THEORY INTO PRACTICE

INTEGRATION VEHICLE FOR:

- DATA MANAGEMENT
- DATA STANDARDIZATION
- DATA ELEMENT NAMING
- DATA QUALITY ASSURANCE
- DATA ENGINEERING
- DATA MANAGEMENT

GAIN ADDITIONAL RESOURCES:

- PEOPLE
- FUNDING
- TECHNICAL ASSISTANCE (TOOLS)

LEGITIMIZE & INSTITUTIONALIZE DA

Figure 4

RECOMMENDATIONS:

DATA ADMINISTRATION & DATA MANAGEMENT

MUST BECOME CHANGE AGENTS!

RE-WRITE LIFE CYCLE MANAGEMENT:

POLICIES

STANDARDS & GUIDELINES

OPERATING PROCEDURES

SEPARATE DATA ADMINISTRATION

AND SYSTEMS DEVELOPMENT

BE PREPARED TO TAKE SOME RISKS,

INCREMENTAL CHANGE TAKES FOREVER!

Figure 5

DATA ADMINISTRATION AND THE SYSTEMS DEVELOPMENT LIFE CYCLE

Ellen J. Levin
Federal Home Loan Mortgage Corporation

This paper describes some of the efforts to introduce a system development life cycle management approach at the Federal Home Loan Mortgage Corporation, more commonly known as Freddie Mac. This approach involves a development methodology that incorporates the data driven and process-driven approaches to application system development. This paper offers an analysis of the factors which have contributed to its successful implementation and offers directions for future development toward an approach more heavily weighted toward a data-driven methodology.

The life cycle development manual was developed by a small team of individuals representing different sets of experiences toward system development -- entity and data modeling, functional decomposition/structured analysis, and testing/quality assurance. Established by high-level technical development management, the team's purpose was announced to all information systems personnel. An advisory committee, representing key systems development, database, and maintenance areas, provided regular feedback on the team's efforts, and helped to publicize the life cycle activity.

The team had two major objectives which it followed to assemble the methodology. The first was that the methodology had to be a balance between "data and process," and second, that the life cycle should be an "organizing principle" which would provide the framework for a common basis of understanding. Both of these goals were achieved. The necessity for addressing both data and process recognized that there were two vocal, equally determined schools of thought within the organization, each convinced that one approach or the other was the preferred method. In fact, the "process people" far outweighed the "data defenders" in numbers. This is a reflection of the prevalence of process-driven structured analysis techniques currently practiced by information systems professionals today. The balanced approach required that each side would have to understand the other and that the result would be a workable method that would address all concerns. The organizing principle involved the definition and description of products to be produced at each phase of the life cycle, categorized as pertaining either to the process model or the data model. Additionally, products were specified. The life cycle manual was summarized on one page as a matrix of products to be produced at each phase within each model (see fig. 1).

Definition of products irrespective of the functional area that might be responsible for producing them helped to focus attention on information content. This avoided potential conflicts between organizational groups and also a premature concern about resources. A further benefit of this approach was that the skills needed to produce each product could be identified. This has formed the basis for hiring and training decisions, which will ultimately change the organizational makeup and the approach to system development.

Successful implementation of the life cycle methodology is dependent on a number of factors. Clearly, management commitment is essential and can be seen in several key areas. One is the official endorsement which management can provide. Second, and perhaps more significant, is the encouragement and support that management can provide by rewarding quality products. This could be a departure from a cultural norm where the primary focus is getting things done on time. When implementing a new methodology, inevitably there will be impacts on schedules, since there are new techniques to understand and use effectively. While management should endorse and reward use of the new life cycle, it should not, initially, mandate its use. Preferably, pilot projects involving team members with a willingness to try new things should volunteer to exercise the life cycle out of an interest in providing more structure to their projects. This voluntary nature produces willing participants who are then able to identify the benefits, as well as some of the difficulties, to management and to other developers. A critical component is providing ongoing training in the use of the new life cycle methods.

One area of difficulty which must be addressed involves some apparent duplication of effort and products inherent in producing discrete products for both the data and process models, some of which do overlap. For example, structured analysis techniques identify data items such as data flows, data stores and data elements. The developers must reconcile these with the entities, entity-relationships, and entity attributes that form the data model. A second difficulty is the large volume of information that must be documented and managed for the project. Both the structured analysis techniques and entity-relationship approaches require the production of graphic materials, data flow diagrams and entity-relationship diagrams, as well as narrative and descriptive text. These difficulties can be effectively addressed through the use of software engineering tools that fully integrate and cross-validate the data and process model products of the life cycle in an automated way.

The use of automated CASE tools also helps to assure that life cycle products meet quality standards. The life cycle specifies that certain products will be produced at certain phases. A level of rigor is imposed on developers to produce the

deliverables at the specified time. This encourages complete analysis before beginning design, and completed design before implementation. There may be a tendency to postpone clear specification until the design and development phases, reflecting a tendency of developers to devise physical solutions to problems before completely understanding the logical system requirements. A quality assurance function can play a valuable role in this area by educating project teams, conducting product reviews at identified milestones and by informing management of risks when life cycle method standards are not met.

The life cycle methodology, with its emphasis on a balance between process and data, provides a certain familiar frame of reference for a majority of the developers, because it includes the functional decomposition, structured analysis approach. Because of that, it tends to represent an application view of the data. The structured analysis methodologies define scope based on the functionality of the systems, not on the data being used. This is a fundamental weakness of the balanced approach. This may result in continued redefinition of data, since each application project has a limited, partial view of the data.

One solution to this dilemma is to use the life cycle within the perspective of a global conceptual entity-relationship model, which documents organization-wide business rules. The global model becomes the starting point for all development projects. In this way, project scope boundaries can be established in a rational manner. The starting point of the life cycle process is the validation of the global entity relationship model and the definition of all entity attributes. The process portion of the life cycle then consists primarily in determining what business functions and processes need to be performed on the data. All possible questions and activities required by the functional areas must be satisfied by the logical database structure as depicted in the entity-relationship diagram.

Defining the data model prior to the process model requires an understanding of the importance of focusing on the data. This data-driven understanding, in many organizations, cannot be achieved quickly. Rather, a gradual progression of increasing understanding is more achievable. Using a life cycle methodology that combines the data and process approaches, while initially resulting in some inefficiencies, serves to highlight the importance of defining the data model while acquainting developers with the purpose and methods used in data analysis. It sets the stage for moving beyond process driven requirements analysis to an organizational environment that is prepared to adopt a data-driven approach.

Ms. Levin is Senior Information Systems Analyst at the Federal Home Loan Mortgage Corporation (Freddie Mac), designing the Life Cycle Methodology and Corporate Enterprise Model. Previously, she has worked at Intelsat and Comsat.

Lifecycle Products Chart

MODELS	PHASES							
	Project Definition	**General Requirements**	**Detailed Requirements**	**General Design**	**Detailed Design**	**Development**	**Testing**	**Implementation**
Process	Business Functions Statement Change Statement Functional Project Classification	Existing Business Functions Functional Requirements	Functional Specification Non-Functional Requirements System Conversion/ Initialization Requirements Requirements Priorities Transaction Definitions	Hardware Description Software Description Software Subsystem Description Technical Implications Implementation Strategy System Security Specification	External Design System Control Structure Module Specifications Checkpoint/ Restart & Locking	Built, Tested, and Secured Modules Draft User's Guide Draft System Documentation Module Test Data Training Plan	System/Approval Test Data Regression/ Integration Test Data Tested System Tested User's Guide Tested System Documentation Trained Users	Production System System Documentation User's Guide Conversion Signoff
Data	Major Subjects Data Base Project Classification	Entities Entity Relationships Entity Relationship (E-R) Diagram	Data Element Definitions Normalized Data Structure Data Conversion Requirements Data Integrity Requirements Normalized Data Structure Diagram Data Security Policy Requirements Record Retention, Archiving, and Backup Requirements	Data Implementation Recommendation Initial Physical Data Base Design Prototype Data Tables Data Integrity Specifications Data Security Specifications Transaction I/O Estimate Application Data View	Refined Physical Data Base Design Data Base Storage Specification Physical Data Base Structure Definition Data Conversion Specification	Compiled and Secured Data Base Structures Data Base Performance Reports Data Base Backups Archived Data Base Journal Data Base Log File	Data Base Performance Reports	Production Data Base Structures Data Base Backups Archived Data Base Journal Data Base Log File Data Base Performance Reports
Qualification	Product Review Summary Reports	Project Test Plan General Requirements Test Cases Function-Entity Matrix Product Review Summary Reports	Detailed Requirements Test Cases Transaction-Data Matrix Functional Access Map Application Controls Matrix Regression/ Integration Test Approach Product Review Summary Reports	System Test Plan User Approval Test Plan Requirements-Design Matrix Regression/ Integration Test Plan Product Review Summary Reports	System Test Cases User Approval Test Cases Module Test Cases Transaction-Module Matrix Module-DB Record Matrix Regression/ Integration Test Cases Product Review Summary Reports	Module Walkthru Reports Module Test Reports	System Test Report User Approval Test Report Migration Signoff Regression/ Integration Test Report	Production Migration Signoff
Lifecycle Management	Project Charter Project Impact Analysis Cost/Benefit/Risk Analysis Lifecycle Project Plan General Requirements Phase Work Plan	Detailed Requirements Phase Work Plan Management Summary and Analysis Milestone Review Report	General Design Phase Work Plan Management Summary and Analysis Milestone Review Report	Detailed Design Phase Work Plan Management Summary and Analysis Milestone Review Report	Development Phase Work Plan Management Summary and Analysis Milestone Review Report	Testing Phase Work Plan Management Summary and Analysis Milestone Review Report System Conversion Plan	Implementation Phase Work Plan Management Summary and Analysis Milestone Review Report	Project Performance vs. Estimate Project Lifecycle Assessment Post-Implementation Review Report

Figure 1

DATA ADMINISTRATION AND THE SYSTEMS DEVELOPMENT LIFE CYCLE

SUMMARY OF REMARKS

Mary Ann Wallace
National Archives and Records Administration

Records managers act as a 'drain' for information systems, assuring that archived information is stored safely and systematically. In addition, they identify information already available, ensure its existence for a certain length of time, then dispose of it. Two percent of all government records reside permanently in the National Archives.

A 'record' has been defined by Congress as any information related to the government's business, regardless of physical form. This is not too far from the data management definition of a set of related data treated as a unit.

Figures two through five illustrate the interaction between records management and the life cycle management of information systems.

Ms. Wallace is Director of the Agency Services Division of the National Archives and Records Administration. She is responsible for assisting agencies in the management of recorded information.

RECORDS MANAGEMENT
AND THE
LIFE CYCLE MANAGEMENT OF INFORMATION SYSTEMS

------ WHAT IS RECORDS MANAGEMENT?

------- PURPOSE

------- TOOLS

------- ORGANIZATIONAL RELATIONSHIPS

------- AUTHORITIES

Figure 1

RECORDS MANAGEMENT
AND THE
LIFE CYCLE MANAGEMENT OF INFORMATION SYSTEMS

WHAT DOES THIS HAVE TO DO WITH THE LIFE CYCLE MANAGEMENT OF INFORMATION SYSTEMS?

---- MANAGING INFORMATION AND THE TOOLS NEEDED TO SUPPORT IT

---- RECORDS ARE RECORDED INFORMATION

---- THEY MUST BE PROPERLY MANAGED JUST AS THE TOOLS THAT CREATE, COLLECT, PROCESS, TRANSMIT, USE, AND STORE THE INFORMATION

---- RECORDS MANAGEMENT CONSIDERATIONS APPLIED TO LIFE CYCLE MANAGEMENT CAN:

--- REDUCE INFORMATION SYSTEM COSTS

--- ENSURE AVAILABLITY OF INFORMATION TO USERS FOR LENGTH OF TIME REQUIRED

--- PROVIDE FOR THE PROTECTION AND INTEGRITY OF THE INFORMATION

--- SEE THAT AGENCY OPERATIONAL NEEDS, INCLUDING LEGAL REQUIREMENTS, ARE MET

--- OBTAIN THE NECESSARY AUTHORIZATION TO DISPOSE OF THE INFORMTION WHEN NO LONGER NEEDED

Figure 2

RECORDS MANAGEMENT
AND THE
LIFE CYCLE MANAGEMENT OF INFORMATION SYSTEMS

WHERE AND HOW DO THEY GET TOGETHER

INITIATION PHASE

IDENTIFY:

INFORMATION SUPPORTING MISSION AND ITS CURRENT LOCATION, ORGANIZATION, AND DISPOSITION

AGENCY/EXTERNAL SOURCES OF INFORMTION

OTHER INFORMATION USERS

INFORMATION DOCUMENTATION REQUIREMENTS, INCLUDING LEGAL REQUIREMENTS

ESTABLISH SYSTEM DOCUMENTATION AND RECORDKEEPING REQUIREMENTS

DOCUMENT ACCESS REQUIREMENTS, INCLUDING PUBLIC ACCESS AND RESTRICTED INFORMATION

Figure 3

RECORDS MANAGEMENT
AND THE
LIFE CYCLE MANAGEMENT OF INFORMATION SYSTEMS

DEVELOPMENT PHASE

IDENTIFY RECORDS SUPPORTING INFORMATION IN CURRENT SYSTEM

IDENTIFY INFORMATION REQUIREMENTS FROM THE SYSTEM AND THEIR RECORDKEEPING PRACTICES (ROUTINE, SITUATIONAL AND EXPECTIONAL REPORTS)

DETERMINE IF THE SYSTEM WILL RESULT IN NEW RECORDKEEPING REQUIREMENTS - WHAT ARE THEY, HOW WILL THEY BE COMPLIED WITH

ANALYZE THE LENGTH OF TIME INFORMATION IS NEEDED TO SUPPORT THE INFORMATION REQUIREMENTS OF ALL USERS

DETERMINE FORMAT THAT BEST MEETS LENGTH OF TIME REQUIREMENTS

ADVISE ON VITAL RECORDS PROCEDURES

OBTAIN A PRELIMINARY DETERMINATION FROM THE NATIONAL ARCHIVES ON PERMANENT VALUE OF INFORMATION CONTAINED IN THE SYSTEM

DETERMINE HOW SYSTEM WILL MEET NATIONAL ARCHIVES REQUIREMENTS FOR THE MAINTENANCE AND DISPOSITION OF ARCHIVAL INFORMATION

ASSIST WITH DOCUMENTATION OF SYSTEM DESIGN TO ENSURE INTEGRITY OF THE RECORDS

Figure 4

RECORDS MANAGEMENT
AND THE
LIFE CYCLE MANAGEMENT OF INFORMATION SYSTEMS

DEVELOPMENT PHASE (CONT.)

DETERMINE AND EFFECT DISPOSITION OF RECORDS ASSOCIATED WITH INFORMATION SYSTEMS BEING REPLACED

OBTAIN DISPOSITION AUTHORITY FOR INTERIM DOCUMENTATION

PROVIDE GUIDANCE ON RECORDS MAINTENANCE, ACCESS, AND DISPOSITION FOR INCLUSION IN USER AND OPERATIONS MANUALS

ENSURE RECORDKEEPING AND RECORDS DISPOSITION REQUIREMENTS ARE PART OF POST IMPLEMENTATION REVIEW PLAN

OPERATION PHASE

ENSURE ALL RECORDS MAINTENANCE, ACCESS, AND DISPOSITION PROCEDURES ARE INCLUDED AND WORKABLE

PERIODICALLY REVIEW SYSTEM OPERATION

OBTAIN DISPOSITION AUTHORITY FROM THE NATIONAL ARCHIVES AS NECESSARY

PARTICPATE IN THE POST IMPLEMENTATION REVIEW AND REFINE RECORDKEEPING AND DISPOSITION REQUIREMENTS AS NECESSARY

Figure 5

MANAGING ORGANIZATIONAL CHANGE

AS DATA ADMINISTRATION IS IMPLEMENTED

MODERATOR

Rae Thompson
Smithsonian Institution

PANELISTS

Michael Menard
Joan Shapiro
Richard Lytle
Margaret Skovira

MANAGING ORGANIZATIONAL CHANGE AS DATA ADMINISTRATION IS IMPLEMENTED

Michael P. Menard
President, Bush Menard, Inc.

Objectives of the Presentation

The team responsible for bringing new technology into an organization faces a formidable series of challenges. Technical success does not guarantee project success; equally important is that you gain the commitment of employees and managers at all levels to integrate the new technology into the fabric of the organization. The purpose of this presentation is for both myself and the other panel members to discuss with you some of the techniques which we have found to be successful in this task.

The Dynamics of Technical Change

I find it helpful to look at the introduction of new technology in relation to three other organizational factors: the overall goals of the organization, the operations or tasks and activities performed in the course of doing business, and, most important, the employees themselves. Figure 1 shows the state of these four factors before new technology is introduced. Ideally, the existing technology has been designed to facilitate the current operations and both factors are strongly tied to the goals. The people know how to do their jobs, how to use the technology and are committed to the goals.

Figure 2 shows the disruption caused by the introduction of new technology and suggests a possible agenda for its successful integration into the organization. In Phase 1 not only does new technology replace the old, but it breaks down technology's existing relationships with goals, people and operations. In Phase 2 new operations are put into place to take advantage of and support the new technology. At the same time, however, operation's previous relationships with goals and people are seriously altered. Phases 3 and 4 complete the restructuring process through training of users and demonstration of benefits.

The job of the team, then, is not just to install new technology, but also to construct appropriate new operations and then re-knit the fabric of the organization.

Critical Success Factors

The specific issues you choose to address and the actual techniques for implementation are highly dependent upon the culture of your own organization. No single approach will work in all cases. Here are a number of ideas to consider. The chance to alter existing operations presents a significant opportunity. Instead of merely using new technology to "automate the problem," use this time to reexamine the effectiveness of current procedures. Often employees will view new technology as an increased burden on their work load. If better operations are part of your overall package, then you will be able to sell the technology to employees on the basis of making their job easier. Use one of the existing business modeling technologies to be sure that you fully understand the existing operations and the impact that your technology will have.

Do everything possible to transfer ownership of the new technology to the employees and managers who are going to be affected. People are your most valuable resource and the most important actions you can take are those which will inspire the rest of the organization to work actively for the success of the project. Early involvement in project planning is critical. People who feel that they have some control over their job environment will work hard to make improvements. There is a dangerous tendency to try to maintain tight control over a project and not get too many people or organizations involved. But early on in a project you should identify all the people whose help is critical to your success and involve them in the decision making process. This includes not only your own employees, but managers and key employees from other work groups.

Have an honest but active marketing plan. Contrary to popular belief, good ideas do not sell themselves. You must anticipate potential adverse reactions and develop plans to channel these energies in useful directions. You must, however, resist over selling. Every beneficial change has costs of some sort or another. Be forthright in your explanations of these costs. People will appreciate the respect they are being given by having all aspects of the situation explained to them. Even more importantly, they will not feel resentful later when the hidden costs must be paid.

Make your training program activity-based. It is rarely enough merely to teach people how to use the new technology. You must teach them how to do their jobs with the new technology. During the course of the training, take key work activities and explicitly demonstrate how they will be carried out with the new tools. Ignoring this critical step in the transfer of technology will only result in delayed productivity for the new system.

Finally, it should be an ongoing daily task of every manager to improve the bonding between employees and the goals of the organization. If employees are committed to the overall mission of an organization AND THEY UNDERSTAND THAT AN IMPORTANT COMPONENT OF THAT MISSION IS TO PROVIDE A REWARDING AND PRODUCTIVE WORK ENVIRONMENT, then managing change is really quite simple. Building such trust cannot be done just for one project. It must exist long before the project is even conceived. But in those organizations where these activities occur, change is always an opportunity to be savored rather than a crisis to be solved.

Dr. Menard is Assistant Professor of Information and Communications Systems at Fordham University's Graduate Business School. He has spent eight years at Exxon Corporation in a variety of information systems positions. He holds a Doctorate in Adult Education. His book, The End User's Guide to Computer Systems Development... will be published soon.

TECHNOLOGY AND ORGANIZATIONAL DYNAMICS

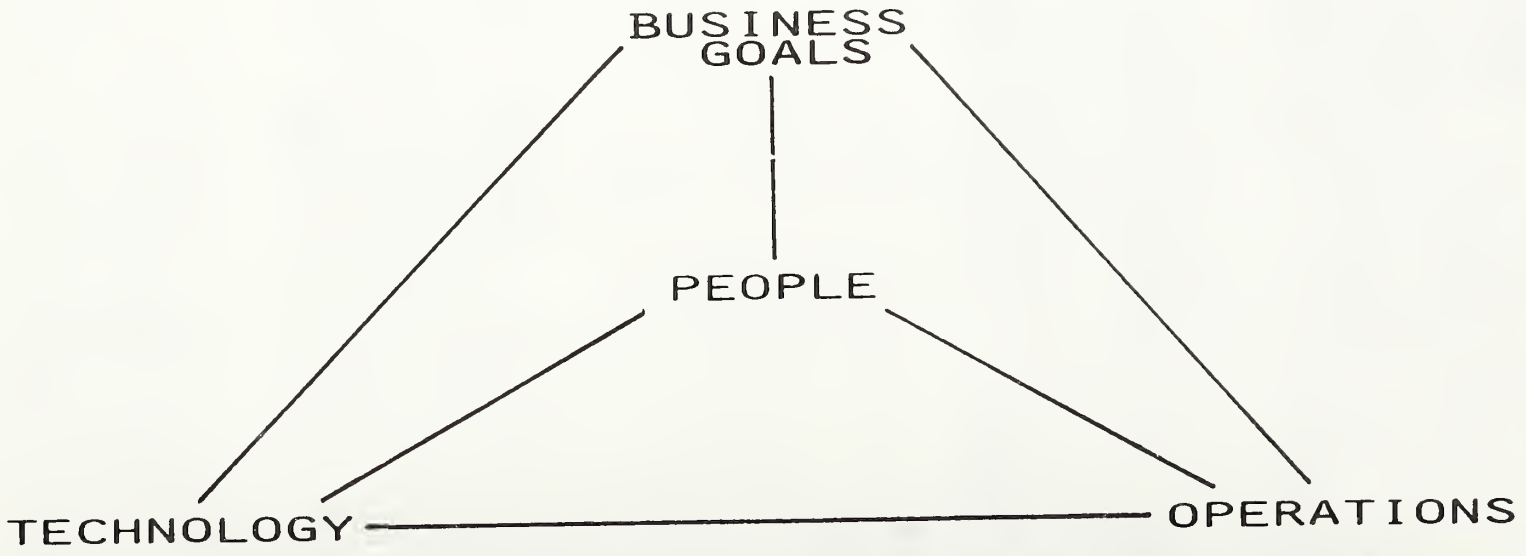
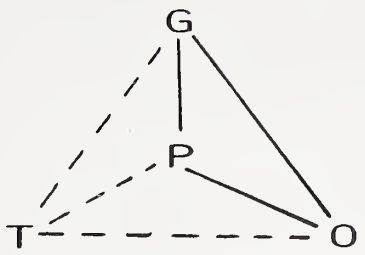
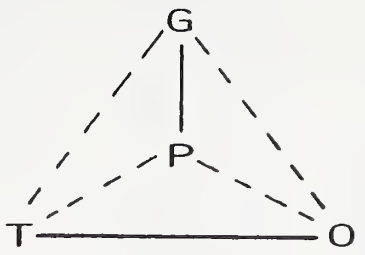


FIGURE 1

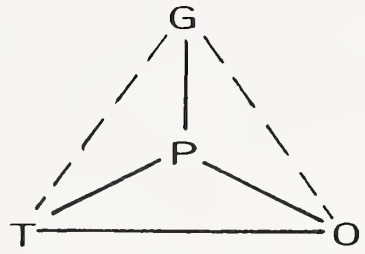
A CHANGE SCENARIO



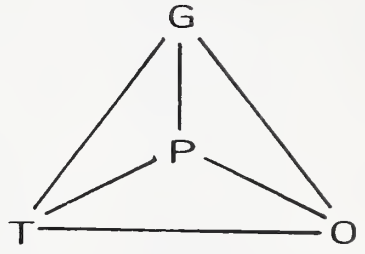
1. INSTALL TECHNOLOGY



2. MODIFY OPERATIONS



3. REDIRECT USERS



4. SHOW BENEFITS

FIGURE 2

MANAGING ORGANIZATIONAL CHANGE AS DATA ADMINISTRATION IS IMPLEMENTED

SUMMARY OF REMARKS

Joan Shapiro
Chemical Bank, New York

A real problem in data administration has been business entities within the corporation which consider themselves, and their data, independent of all other organizational units. They consider their data model their own. There is confusion between sharing **record structure** and **data values**, especially at the top of the company.

The data administration function has been renamed Corporate Information Architecture. The stress is on **interfacing** and standard definitions of such common data elements as 'customer' and product definitions, and the relationships between data elements. Information of cross-corporate use is stored in the corporate dictionary. This data is defined by a very high-level business information model.

Ms. Shapiro has spent twenty years in IRM and data processing. She is currently Associate Vice-President and internal data consultant for Chemical Bank of New York City. She has spoken at many conventions and conferences.

MANAGING ORGANIZATIONAL CHANGE AS DATA ADMINISTRATION IS IMPLEMENTED

SUMMARY OF REMARKS

Richard Lytle
Drexel University

Our culture holds hard science and technology superior to such 'soft areas' as information content.

An organization has both a formal and an informal structure. It is important to know both when introducing change. One should also know the attitude toward top management in the rest of the organization. In converting the data processing staff to data administration, selling planning is the hardest part. Try to find allies in line areas.

The 'soft areas' must be managed rationally. Staff attitudes towards change are seldom taken into account in business presentations. Commitment at the top, ideally at the very top, is essential.

The 'change agent' for the soft areas should be personally consistent with the corporate culture. Once the users catch on, they will be way ahead of the change agent, because they know their business.

FACTORS SELECTED FOR DISCUSSION

- CULTURE OF THE ORGANIZATION
- PERSONALITIES OF KEY MEMBERS OF TOP MANAGEMENT
- POSITIONS AND PERSONALITIES OF THE CATALYSTS OF CHANGE
- FACTORS SPECIFIC TO DATA ADMINISTRATION

CULTURE OF THE ORGANIZATION

- FORMAL AND INFORMAL STRUCTURE
- ATTITUDE TOWARD MANAGEMENT PER SE

- PLACE ON ACADEMIC VS BUSINESS SPECTRUM
- ROLE OF DATA PROCESSING IN THE ORGANIZATION
- INTERNAL CULTURE OF DATA PROCESSING
- ATTITUDE TOWARD PLANNING
- OPENNESS TO CHANGE OF ANY KIND
- AWARENESS OF DATA ISSUES SPECIFICALLY
- WILLINGNESS TO MANAGE "SOFT" FACTORS RATIONALLY

PERSONALITIES OF KEY MEMBERS OF TOP MANAGEMENT

- AS INDIVIDUALS
- INTERACTION AT TOP MANAGEMENT LEVEL

POSITION AND PERSONALITY OF THE KEY CHANGE CATALYSTS

- ADMINISTRATION POSITION
- PERSONAL RELATIONSHIPS WITH KEY PARTICIPANTS
- FIT WITH CORPORATE CULTURE

FACTORS SPECIFIC TO DATA ADMINISTRATION

- INHERENT DIFFICULTY OF EXPLAINING DATA ADMINISTRATION IN MOST ORGANIZATIONS: DATA ADMINISTRATION EQUAL MIND CONTROL?
- PERCEIVED IMPORTANCE OF THE TASK THAT IS THE OCCASION OF DATA ADMINISTRATION IMPLEMENTATION
- CONFUSION OF TECHNICAL AND DATA ISSUES
- CONFUSION BETWEEN DATA ADMINISTRATION AND DATA PROCESSING

- DATA PROCESSING VERSUS DATA ADMINISTRATION??

EXAMINE A FEW PRESCRIPTIONS FOR SUCCESS

- DATA ADMINISTRATION CAN SUCCEED ONLY WITH TOP MANAGEMENT SUPPORT
- DATA ADMINISTRATION MUST REPORT AT LEAST TO LEVEL THAT DATA PROCESSING REPORTS
- CONSTRUCTING DATA ARCHITECTURES MUST PROCEED TOP DOWN
- AUDIENCE-SUPPLIED PRECEPTS

Dr. Lytle is Dean of the College of Information Studies at Drexel University. Until 1987, he was Director of IRM at the Smithsonian Institution. He previously held positions at the Smithsonian, Rice University, and Washington University. He has a PhD in Information Science from the University of Maryland.

MANAGING ORGANIZATIONAL CHANGE AS DATA ADMINISTRATION IS IMPLEMENTED

Margaret Skovira
Bureau of the Public Debt

In the children's song, "The Farmer in the Dell," the farmer acquires a wife and child, a nurse, a dog and cat and in effect builds an organization. The farmer's last selection is the cheese, who is immediately abandoned to "stand alone" by the group into which he was so recently welcomed.

Is this scenario familiar to the data administrators in the audience? Was data administration introduced in your organization with a great deal of fanfare, which soon amounted to a crowd of people waiting to see what you would do? and how? and how soon?

Hiring a data administrator in an organization which has not had one introduces a change that goes beyond an organizational realignment. It does, of course, add a box to the organization chart, and cause some responsibilities to be shifted, but it does much more than that: the introduction of data administration implies a change in management emphasis. An organization without data administration is managing the <u>processing</u> of its data; an organization with data administration aspires to manage the <u>data</u> itself.

Why is the shift from managing processes to managing resources so significant? Consider, for a moment, an organization with no personnel office. Each manager could hire a staff, reward it as he or she saw fit, and retain employees only as long as the need was evident. In such a hypothetical situation the manager would have control over the personnel it took to do the job at hand. A personnel director with classification standards, hiring practices, pay scales, performance appraisals, productivity standards and so on would benefit the organization as a whole, but the individual manager would lose control over one essential resource to getting the job done. He would lose control to a function that would not consider only his office's needs as he would have considered them. A personnel manager might require activities for the benefit of the organization, or for the benefit of its employees, activities that might interfere with "getting the job done."

"Getting the job done" is the focus of operations managers and process managers. Ensuring that the resources are there to do the job is the focus of personnel directors, vice presidents of finance, and data administrators. "Managing data as a resource" is not my topic. But it is important in that context to recognize that data administration is more than a new office or a

further specialization of technical functions. Data administration has the potential to change the way an entire organization does business.

The problem for the data administrator becomes one of how to introduce this far-ranging type of change without adversely affecting productivity. Notice, I did not say "without disrupting the organization." As Dr. Menard has suggested, change will be disruptive when human organizations are affected by it. The challenge is to direct organizational energy into new avenues of endeavor as the traditional ways of doing things change.

The goal is to manage the way data is maintained, used, and discarded and not bring to a grinding halt those processes that depend on the availability of data.

The introduction of data administration has both a tactical and a strategic aspect. Tactically, the data administrator must establish an organizational base. The organizational base will be derived from the components of the old organization, and integrated with them. Strategically, the data administrator will use this organizational base as the springboard for the move to data management.

My experience with data administration is exclusively from within the data processing shop, and my thoughts on change management derive from that viewpoint. Inside the data processing organization, data administration is usually introduced deliberately. It will not appear to be technology driven, as end-user computing is, for example. It will appear to be the result of management commitment to an emphasis on the importance of data. It will not come about because someone is already performing the function, legitimizing what already exists in fact. Rather, the organization chart, the function statements, and the position descriptions will be carefully developed and the reorganization announced with much fanfare. Having established the requisite management commitment to data administration, the organization will stand back to observe how the first incumbent will perform. The cheese stands alone.

If the data administrator wishes to bring about the change she has been chartered to introduce, she will not allow herself to stand alone for long. A strategic data planning project will keep the data administration staff busy, and will have a long-term payoff, but it will not integrate data administration into the organization in the short term. And organizational integration means contributing to the accomplishment of today's work today. The benefits are twofold: first, a product in a user's hand is worth two in the long range plan. When it comes to establishing credibility there is nothing like a concrete accomplishment. Second, by devoting energy to meeting current work demands, the data administrator will uncover areas of

opportunity for the future, opportunities for improved data management; not theoretical opportunities, but real ones.

For example, in my organization the only interaction between users and data has traditionally been a custom built application. As a result, a certain amount of manual data manipulation was occurring when it was not cost effective to design and write programs. Data was being keyed into PC's from computer listings. The automated data access developed for users by the data administration staff was not elegant, but it made the user's job easier and removed an item from the system development backlog. And the data administration branch had a visible accomplishment.

The danger of this approach is the old problem frequently summarized as, "when you're up to your neck in alligators it is difficult to remember that your original objective was to drain the swamp." Achieving a balance requires not losing sight of the objective, in our case a data oriented approach to systems planning and management, and always evaluating the daily opportunities for how they will advance that objective. In providing a single user a means to access data, the data administrator could be laying the foundation for the development of an information center, for example. Another example of a double-edged opportunity is data base administration (DBA). In my experience, the data administration shop, if it has DBA responsibilities, can be overwhelmed in maintaining physical, single application, databases, to the detriment of the achievement of long-term objectives. On the other hand, in the course of time many opportunities will surface through the DBA function, opportunities for data sharing and for information driven systems.

Another aspect of organizational change to be managed by data administration is the introduction of new ways of doing work. The design and definition of data structures, once a by-product of system design, becomes a primary objective in a data administration environment. Entity analysis, data analysis, even a little business systems planning enter the methodology. Though some of these processes may have been occurring intuitively, with data administration comes the opportunity to formalize them into the system development methodology. A data dictionary becomes critical, and automated tools for data analysis, database design, and for information extraction are likely to be called for. The data administrator's task is to see that the tools and methodologies for the new ways of working are in place. But these tools and methodologies apply to the whole organization and cannot be established by fiat. The data administrator may be developer, facilitator and implementor - but not the "owner" of the new methodologies. And never merely a spectator on the sidelines. The introduction of these changes, so critical to long-term accomplishments, will require all the diplomatic and technical skill that the data administrator can muster.

The introduction of new tools and methodologies will not be universally welcomed. A dictionary full of metadata restricts the "creativity" of application developers, and requires additional steps in the system design process. Initially, at least, the systems analyst will not find his job made easier by the fact that data definitions are centralized and independent of applications. Instead, he will bemoan the need to coordinate his activities with other organizational components, and the inevitable errors, omissions and inconsistencies that require time and effort to resolve. As a data oriented system development methodology is introduced, the first attempts to do new tasks may be tentative, trial and error efforts in which no one is sure of the objective because "we've never done this before."

The challenge for the data administrator is to facilitate the implementation of these changes in the ways of doing work without retaining sole proprietorship of the processes. There is a certain amount of "selling" involved, as Dr. Menard has suggested. The selling may occur in the guise of training, or, better, during consensus building as new procedures are incorporated into existing methodologies. Eventually, the benefits have to be apparent to all involved, or the new approaches will be short cut.

What does the data administrator hope to achieve in the short term? The management commitment that was behind the original establishment of data administration must be bolstered with an organizational commitment: an organizational commitment to a data management approach to data processing, an organizational commitment to planning and developing systems from an information requirements perspective.

To be successful in the long term, data administration must change the way an organization views data. Data will be treated as the primary resource of systems, as the term "data processing" has always implied. Data will not be taken for granted but analyzed, understood, and its forms perfected before processing systems and user interfaces are designed. To be effective, this new view of data must be shared throughout the organization. It cannot be a vision held only by the data administration staff.

Data administration cannot be introduced in isolation, the cheese cannot stand alone. The first step is the realignment and integration of functions. The second is the modification of the infrastructure of standards and procedures, methodologies and tools, to facilitate a data management approach. As these changes are introduced, some immediate accomplishments are desirable to establish the utility and credibility of data administration, now, and its potential for the future.

When all this is done, data administration will have a base from which to address the entire organization's data and information requirements. Ultimately data management can become as normal as personnel management.

Ms. Skovira is Data Administrator for the Treasury Department's Bureau of Public Debt. She has spent twenty years in the Automatic Data Processing (ADP) field. She is currently Chairperson of AFFIRM.

TWO COMPLEMENTARY METHODS OF DEVELOPING DATA ARCHITECTURE

MODERATOR

Judith Newton
National Bureau of Standards

PANELISTS

Cathy Hirsh
Tom Turk

TOP-DOWN METHOD OF DEVELOPING DATA ARCHITECTURE

Cathy Hirsh
American Management Systems, Inc.

Peter Drucker has defined information as "data endowed with relevance and purpose" in his article, "The New Organization," published in the January-February, 1988 issue of the Harvard Business Review. Today, organizations have much automated data, in fact, too often, more data than can be used effectively. In addition, and unfortunately, much of that data cannot be applied with "relevance and purpose" because the data available are not the correct set of data needed or because they are of inadequate quality.

How can this problem be corrected and how can organizations position their information systems to be capable of delivering needed information, or data that can be endowed with "relevance and purpose", to support the goals of the enterprise? One of the answers lies with the development and implementation of an appropriate data architecture. This presentation addresses one way of developing a data architecture--the Top-Down Design Approach.

There are two key objectives of a data architecture:

- To be able to provide a capability that can deliver the information required by the organization over both the short and longer terms; and

- To provide a basis and aid for coordinating and implementing appropriate information systems facilities.

To fully satisfy these objectives, any approach that is to be used in designing a data architecture must have the following characteristics:

- The activities performed in designing the data architecture must be integrated and addressed together with other requirements, such as functional requirements, technology requirements, organizational requirements, etc. While it is desirable to utilize an approach that is "data-driven," identification of data requirements, absent of consideration for other needs, cannot be the sole focus.

- The approach must be capable of providing results that are timely, developed at a reasonable level of effort, and that address the appropriate level of detail.

- The approach must be capable of addressing requirements for data other than those arising from traditional, "institutional" systems. In the past, most of the automated support, and consequent requirements for data, have resulted from the attention focused on the needs for basic support of organizational operations. Today, capabilities to support the "professional" or knowledge worker must also be considered--capabilities such as requirements for decision support systems, executive information systems, and user support systems. In addition, capabilities to support "external" systems interfaces or potentially, even users of the organization's systems who are themselves not part of the organization (external to the organization), must also be addressed. As figure 4 shows, institutional systems represent only one aspect of the total information systems portfolio which most organizations require.

- The approach must also result in recommendations which are practical and feasible within the context of the organization's situation. Recommendations that are "ideal," but not practical for the organization, are not the "right" recommendations.

- The approach should be capable of being applied either at the top-level of the enterprise or to portions of the organization. Often, the scope of a study at the top-level of the organization would be too broad, or else, the organization is not yet ready to address the issues at that level. Hence, an effort for some important or "strategic" portion of the organization might be more appropriate.

- The approach must provide assurance that the recommendations are integrated with the organization's business or mission-support strategies.

The Top-Down Design Approach, if implemented properly, has all of these attributes.

The key steps and products of the Top-Down Design Approach are shown in figure 5. There is not time to discuss this process in detail. Instead, some examples will be used to highlight a couple of steps in the process.

The first example (fig. 6) shows an "information flow" model that is one of the products of the enterprise analysis phase. This diagram shows candidate groupings of information systems and data flows among them based upon a high-level analysis of organizational functions and their supporting information requirements. This diagram provides a high-level overview and context that can be used for further refinement to determine

additional detail about specific systems and data requirements and, in the case of potentially distributed processing environments, to determine possible levels and types of distribution.

The next example (fig. 7) presents a summary level information architecture. It identifies the major systems, the subject data stores and data classes, flows among the systems, and the location of the systems (i.e., mainframe, local workstation, external, etc.). While not obvious to one unfamiliar with the specific situation, it also incorporates the information technology approaches selected to satisfy identified organizational strategies. This diagram demonstrates the type of information necessary to the next stages of effort, the more detailed development of data, application, and distribution architectures and the development of an implementation plan.

Up to this point, the presentation has addressed the context within which a data architecture is developed utilizing a top-down design approach. The remainder of the discussion will focus specifically on the development of the data architecture and on the benefits of developing it in this fashion.

The process for developing a data architecture following a top-down design approach is the following. It begins, as one might expect in utilizing a top-down approach, at a high level with content progressively refined as more detailed requirements are determined. Entity-Relationship (E-R) Analysis techniques are used initially and, as further detail is developed, throughout all phases of development and implementation. As the design progresses, primary and secondary identifiers and groups of attributes, or domains, for each of the entities are identified. As identifiers and attributes are determined for each of the entities, the rules for first and second normal forms can be applied. At the logical database design level, the full set of attributes should be identified and the database design can be represented following relational data model conventions. The next graphic depicts the process just described. Note that further detail on the "data side" progresses only as further detail is determined on the "process side," and vice versa.

Figure 10 is a sample Entity-Relationship diagram. The scope covered by this diagram is intended to be that for a full personnel system, and it identifies all the entities for which support will be provided by the particular system. The diagramming technique used in this example is one of the techniques used in E-R diagrams. Other techniques, some of which show more detailed information, are also possible.

The E-R diagram, even though at a high level, displays much information for review and analysis - "business rules" that might otherwise not be visible but imbedded in program design

specifications, code, or omitted entirely from the system. For example, the diagram shows that for this particular organization, an employee can have multiple positions and that a given position can be held by multiple employees. For other organizations, this rule might be entirely different. A 1-to-1 relationship, for instance, is one some organizations have which means that an employee can only have one position and that a position can be held by only one employee. Similarly, the entire content of an E-R diagram can be reviewed to ensure full coverage of the necessary entities and that the relationships defined among them are proper for the organization. Such "rules," shown in diagrams of this nature, can be easily reviewed, analyzed, and verified by end users and top management to ensure that the system will be designed to meet the organization's requirements.

Optimally, it is nice to be able to begin data analysis and the preparation of E-R diagrams at the enterprise level with refinement occurring as further design work progresses to the system and project levels. At each successively lower level, the E-R diagrams are reconciled to the next-higher level to ensure that consistency of the design is maintained across all systems for the enterprise.

In some cases, as in the example, it is not possible to begin at the enterprise level, and work begins at the system, or even at the project level. Even in these instances, the use of entity analysis techniques is very beneficial, providing a full picture of the data at the particular design level and providing a basis for review and analysis as other systems are developed that will interface to or be integrated with the system.

Benefits of the top-down design and entity analysis approach are:

- It provides a very understandable, simple technique for exposing, determining, and verifying business rules and integrity rules related to the data.
- It provides an easy technique for building consensus as to what data requirements a system needs to support.
- It can be applied in a phased manner as the process and data designs are refined progressively.
- It can be applied to both new and existing systems and data stores.
- Numerous tools are now available to support the techniques with more refined tool support expected for the future.
- It provides a big picture look to avoid the "missing the forest for the trees" syndrome.

In conclusion, a data architecture, as well as all the other information system products must relate to and support an organization's strategies and management objectives. The top-down design approach, coupled with entity analysis techniques, by providing an overview process, provides an opportunity for development of products that do so. This approach works and, like any approach, can be adapted to meet the needs of a particular organization and situation.

Ms. Hirsh is Senior Principal at American Management Systems, Inc. She is currently manager of the information resources management practice area. She has supervised projects addressing a wide range of IRM issues. She has been a manager of many corporation-wide planning and implementation projects, and team leader for numerous system development projects and ADP studies. Previously, she was director of data administration for Government Employees Insurance Company (GEICO). She is past member of the board of directors for GUIDE International and currently manages GUIDE's Executive Forum for senior level executives. She is also the Vice-Chairperson for Programs for AFFIRM.

TWO COMPLEMENTARY METHODS OF DEVELOPING DATA ARCHITECTURE-- TOP DOWN DESIGN APPROACH

Cathy Hirsh, Senior Principal
American Management Systems, Inc.

Prepared for
DAMA Symposium
Data Administration: Management and Practice

May 17, 1988

ams

Figure 1

OBJECTIVES FOR A DATA ARCHITECTURE

- "INFORMATION IS DATA ENDOWED WITH RELEVANCE AND PURPOSE"
  - -- Peter Drucker, "The New Organization," Harvard Business Review, January-February, 1988
- OBJECTIVES IN BUILDING A DATA ARCHITECTURE
  - -- To be able to provide a capability that can deliver the information required by the organization over both the short and longer terms;
  - -- To provide basis and aid for coordinating and implementing appropriate information systems facilities.

ams

Figure 2

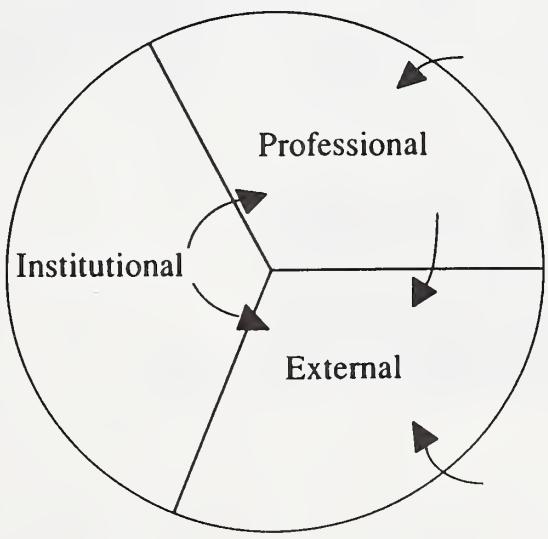
ANY APPROACH USED TO DESIGN A DATA ARCHITECTURE MUST HAVE THESE ATTRIBUTES

- Activities in building a data architecture must be integrated and addressed together with other requirements.
- Approach must result in timely recommendations with level of detail appropriate to objectives.
- Must address systems and requirements other than "institutional" systems.
  - -- Decision support systems
  - -- Executive information systems
  - -- User support systems
  - -- "External" systems/data
- Must result in practical, doable recommendations.
- Can be applied to portions of the organization as well as to the entire "enterprise."
- Must support the organization's strategies.

ams

Figure 3

Information Systems Portfolio



ams

Figure 4

TOP DOWN APPROACH
KEY STEPS AND PRODUCTS

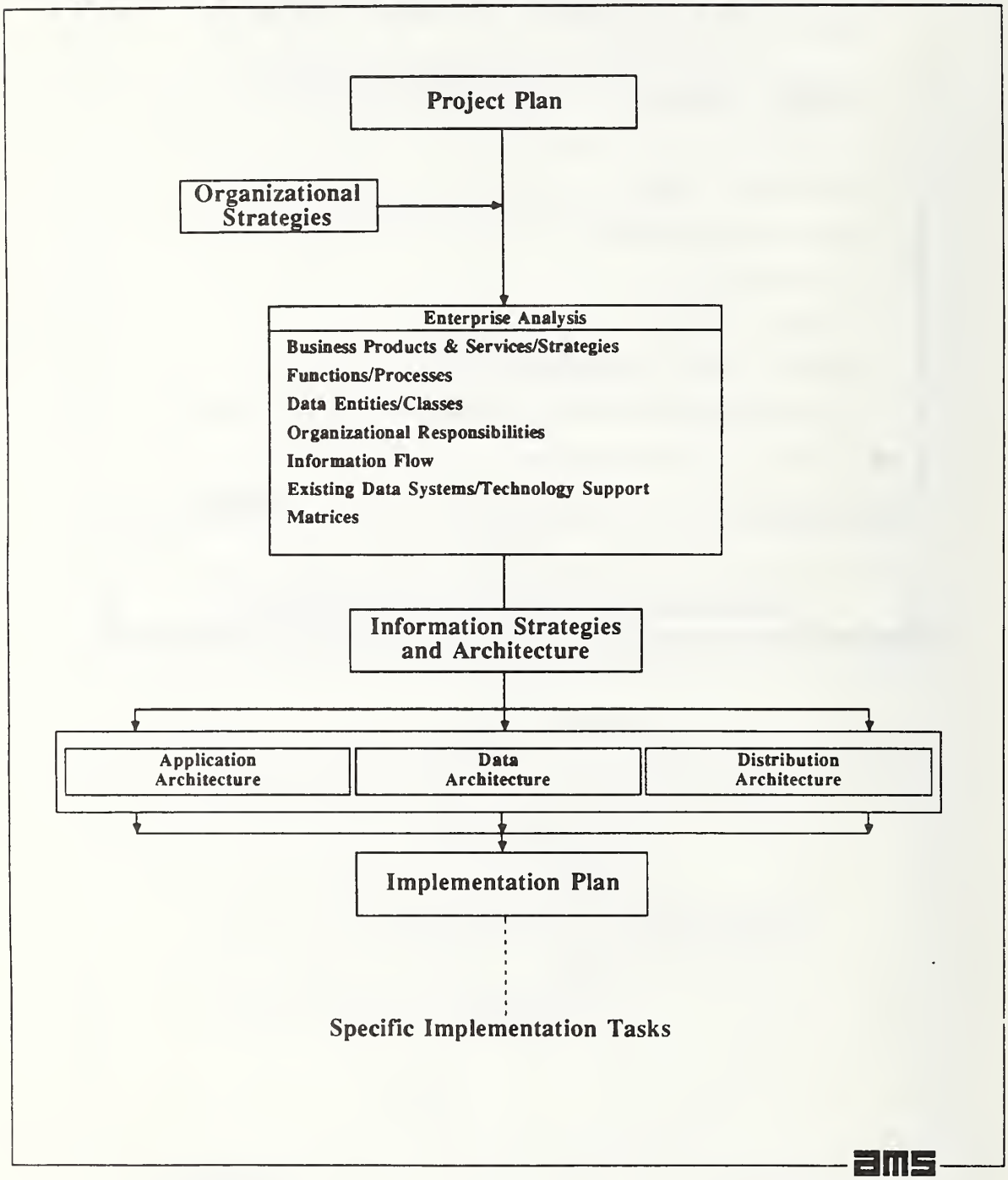


Figure 5

PROCESS EXAMPLE

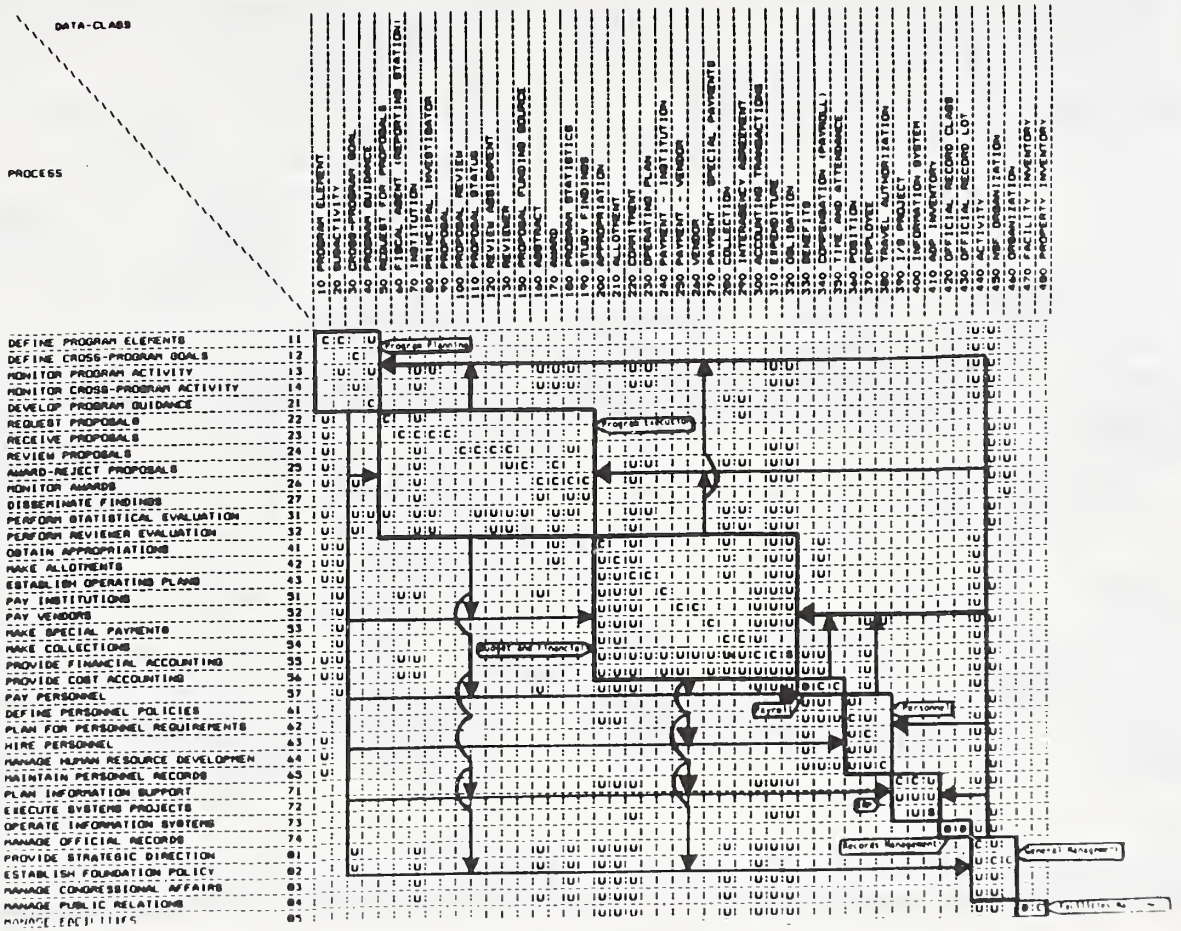


Figure 6

INFORMATION ARCHITECTURE

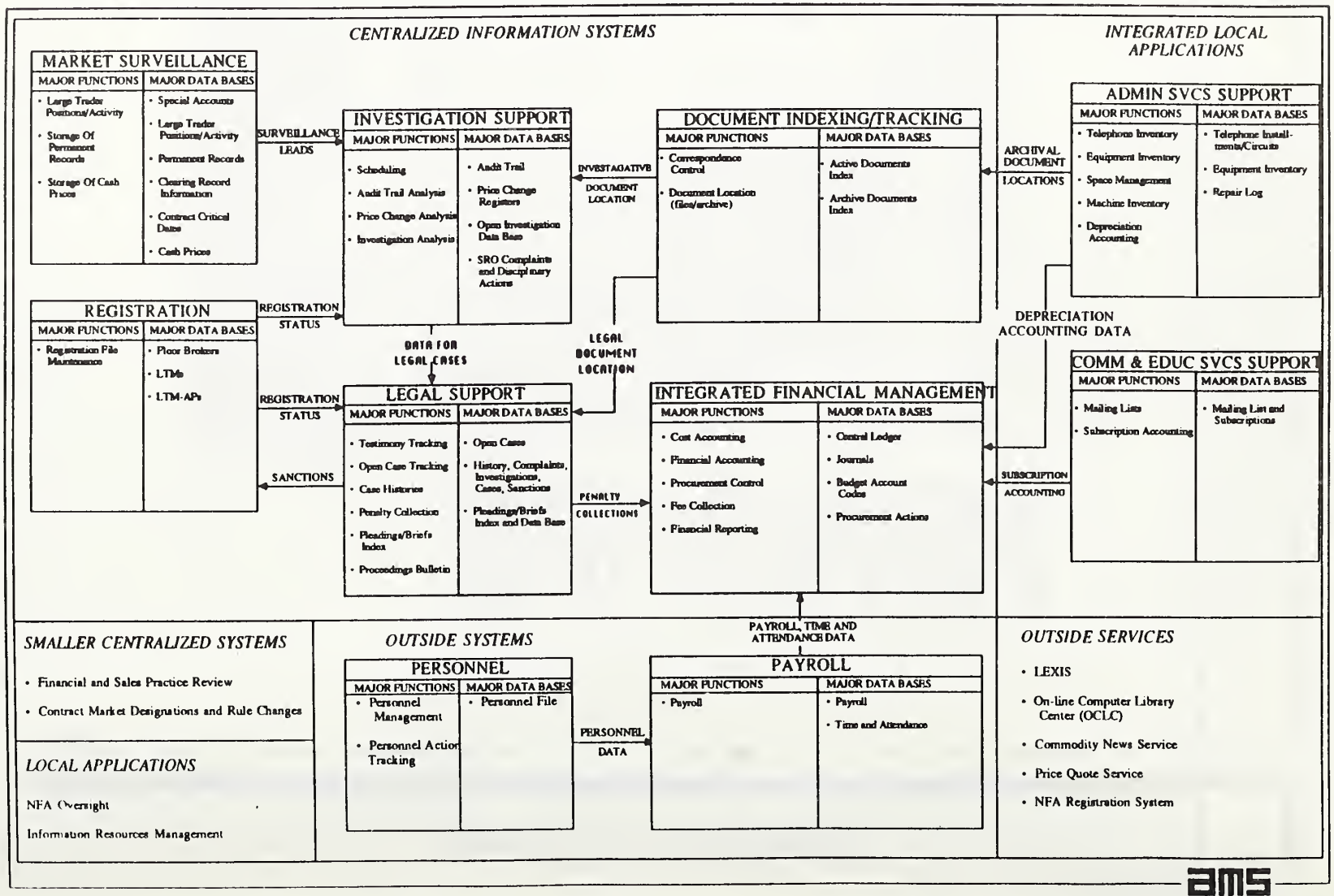


Figure 7

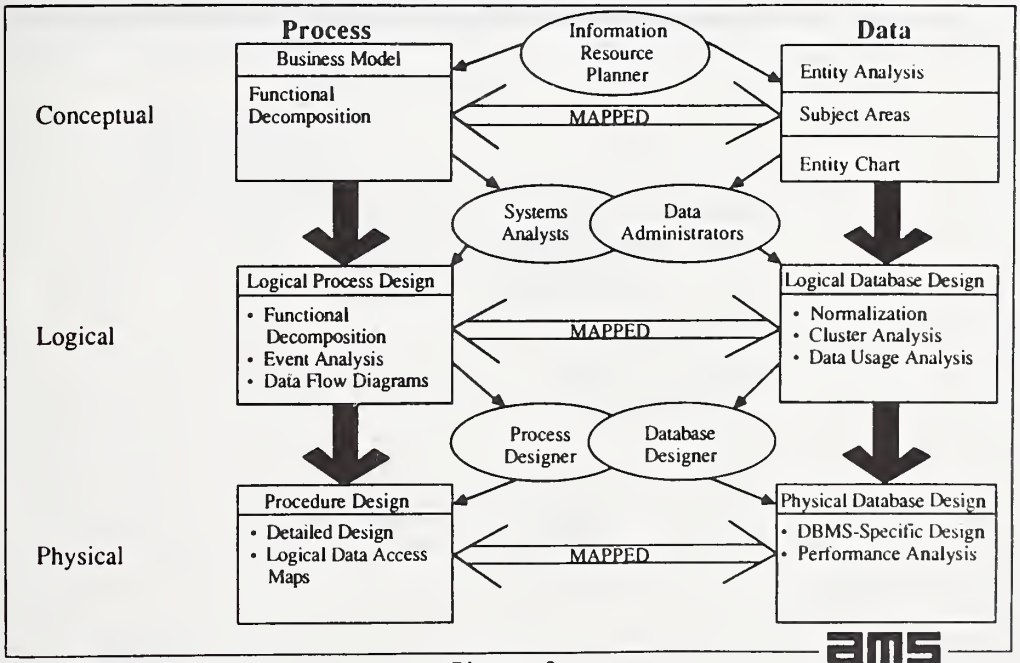
DATA ARCHITECTURE DEVELOPMENT PROCESS

- Begins at a high level, becoming more detailed as the "process" side is refined.
- Utilizes Entity-Relationship Analysis Techniques and diagrams throughout all phases of development and operation.
- As design progresses,
  - -- Identifiers and groups of attributes or domains are identified for each entity;
  - -- Rules of first and second normal forms are applied.
- At completion of logical data base design, all attributes are identified and design can be represented using relational data model conventions.

ams

Figure 8

DATA ARCHITECTURE DEVELOPMENT PROCESS



ams

Figure 9

A SAMPLE ENTITY-RELATIONSHIP DIAGRAM

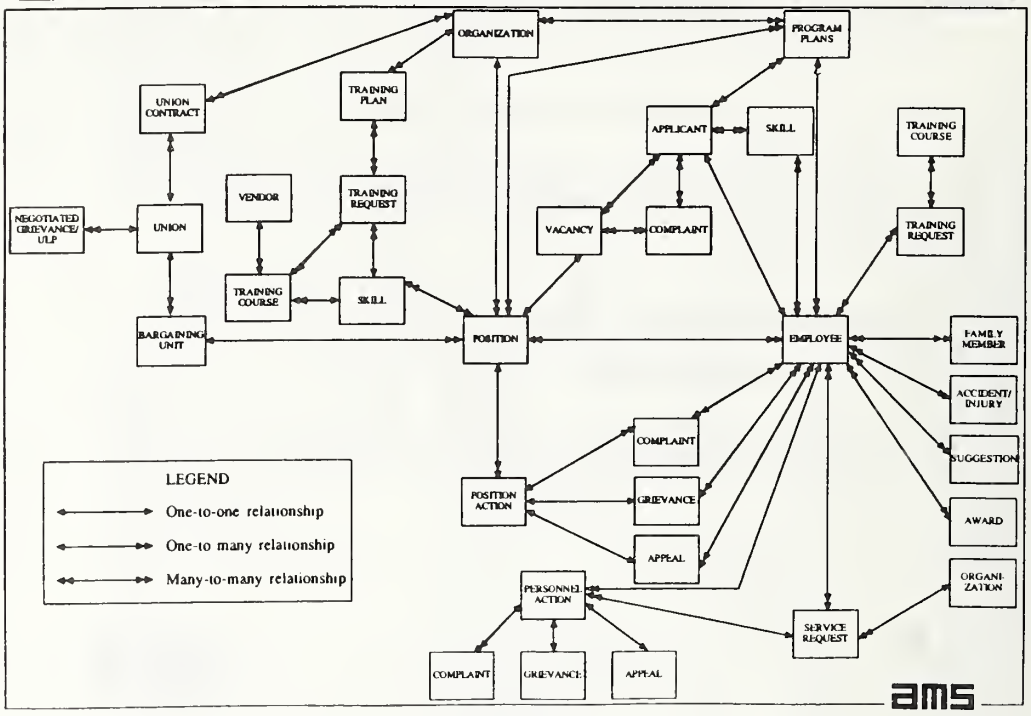


Figure 10

DATA ANALYSIS SCOPE

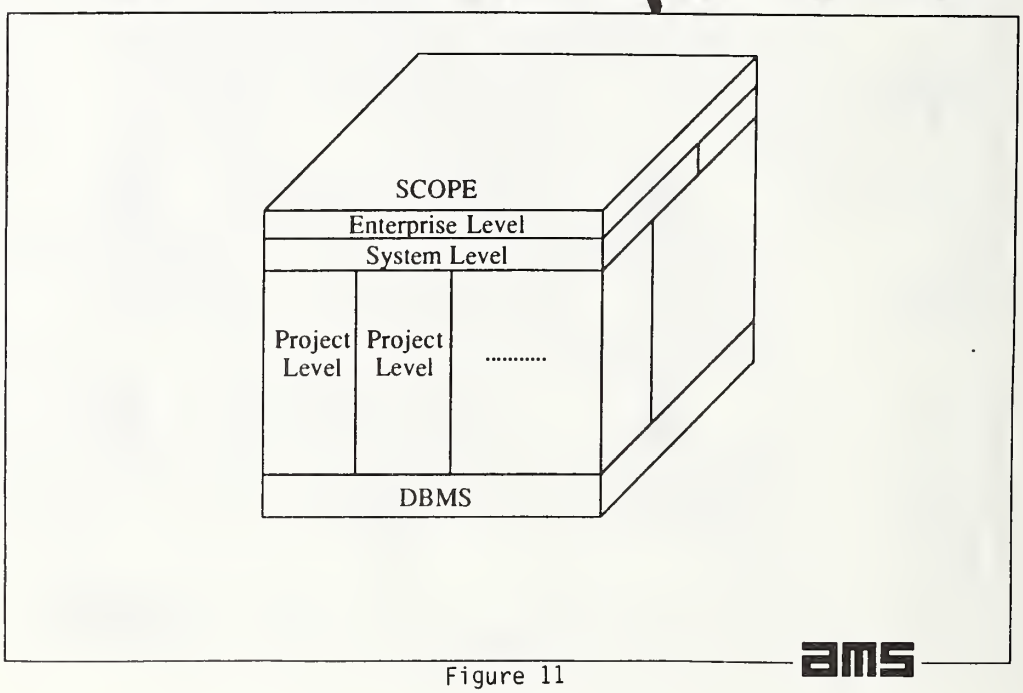


Figure 11

BENEFITS OF THE APPROACH

- An understandable, easy technique for exposing, determining, and verifying business rules and integrity rules.
- An easy consensus-building technique.
- Can be applied in a phased manner.
- Can be used for both new and existing systems and data stores.
- Automated tools are available and becoming better.
- Avoids the "Missing the Forest for the Trees" syndrome.

ams

Figure 12

CONCLUDING REMARKS

- Data architecture must relate to and support organizational strategies and management objectives.
- Top-Down approach and entity analysis techniques, applied and adapted appropriately, provide opportunity to develop needed products.

ams

Figure 13

DATA PLANNING METHOD OF DEVELOPING DATA ARCHITECTURE

Tom Turk
American Data Technologies, Inc.

Why Data Planning?

To support a multi-divisional company you must ensure that each division's data is compatible and meshes. When their data views are brought together, the inconsistencies show up very clearly.

The problem in today's database environment is not the lack of data, but the over-abundance of redundantly stored data in different files and databases, and data stored on different hardware devices. Making reliable data available for decision makers in this environment is nearly impossible or very expensive and time consuming.

What is Data Planning?

Data Planning provides a synergistic approach using normalization techniques to define and create a database environment which satisfies end-user needs, reduces system development expense, and enables more effective utilization of "user friendly" software. A synergistic approach defines the whole data environment first, then breaks it into its component parts for implementation. Data Planning does this by establishing the direction for database development throughout the company by creating an inventory of the company's data and by establishing a model of the data from which all databases and files are built.

Data Planning emphasizes front-end, top down analysis to establish the way the company structures its data (called a logical data model). The planning process integrates analysis of business functions and the data required to perform the business functions. Analysis of a system's functions occurs later, during system development. During system development the designers use the company's logical data structures to design the processing logic and the database.

The planning process is composed of six steps with their completion identifying or producing: an inventory of the company's data; a model of the company's data; dependencies of data within the data model; logical data structures used to design the physical database; and identified source systems for the data.

The seven steps of Data Planning are:

1. Determine the business model which depicts the business' functions, events (flows of data), and categories of data required to support the business.

2. Map current and proposed computer systems to the business functions which are supported by them.

3. Identify data items required by each business function and develop a normalized data model for it.

4. Determine the logical data structures used to design the physical databases.

5. Identify basic volume information for the normalized data.

6. Determine the dependencies that exist among the normalized entities and chart the dependency flow required for creating and updating entities.

7. Determine the business function which is the source for each normalized entity and which system should be the source system.

Database design can be done more quickly, accurately and effectively using data planning techniques. These techniques translate a theoretical model to a physical database which can be implemented in any of the current database management systems. A total of three data models (normalized, usage and structural) are developed during data planning. From these data models, a physical database can be designed and implemented prior to the development of the system's programs. The resulting physical database design can be for micro, mini, or mainframe computers.

Unlike most system, business, or data planning approaches (e.g., IBM's Business System Planning) this data planning approach collects detailed data items and stresses <u>data</u>, not systems, problem analysis, or the development of a high-level view of the entire company. The Data Planning approach results in a general business model of the major functional areas of the business with enough detail to ensure that a detailed data model can be developed. The data model is a logical design of the data that can be implemented in any database management system currently available.

Benefits

Some of the more traditional benefits of data planning are:

- More data sharing because the data structures are based upon the company's definition of data, not a specific system's.
- More consistent information because of standard terminology and definitions.
- Increased knowledge of data availability because of a centralized, mechanized inventory of information about data.
- Better systems planning as the approach fosters the integration of systems' functions based upon the commonality of data.
- Reduces costs by minimizing data analysis efforts, by providing initial physical structures earlier in the development cycle, and by reducing maintenance efforts with more common input/update processing.

Data Planning aids the company in achieving its goals and objectives by improving management's ability to more easily obtain timely data required to increase management's ability to make knowledgeable decisions.

When data planning is integrated with the company's tactical plans, data can be made available on a predefined schedule as part of Management Information Systems' (MIS) tactical plans. This enables system developers and end-users to be able to plan on the availability of data.

System development and maintenance costs can be reduced by sharing data and designing databases for multiple systems. Systems are designed to use the Data Planning developed databases rather than the databases design to satisfy only that system's needs.

During system analysis, the data plan identifies whether the data required to support the system is available or will be available based on current or planned systems. If the data is not available, the system's scope will have to be expanded to collect the data.

Mr. Turk is well-known in data administration circles. In his 18 years in data processing, he has held positions ranging from programmer to manager of data and database administration. As a consultant and head of his own company, American Data Technologies, Inc., he has assisted various companies in the implementation of data administration; developed and taught courses to both end users and MIS personnel; and conducted data planning projects and developed data models. His book, Planning and Designing the Data Base Environment, was published in 1985.

SYNERGISTIC PLANNING CYCLE

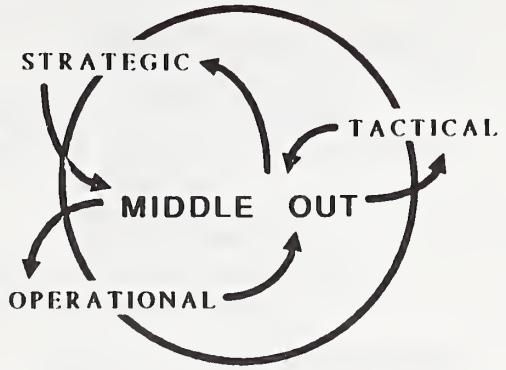


Figure 1

SYNERGISTIC DATA PLANNING PLUS USER FRIENDLY SOFTWARE

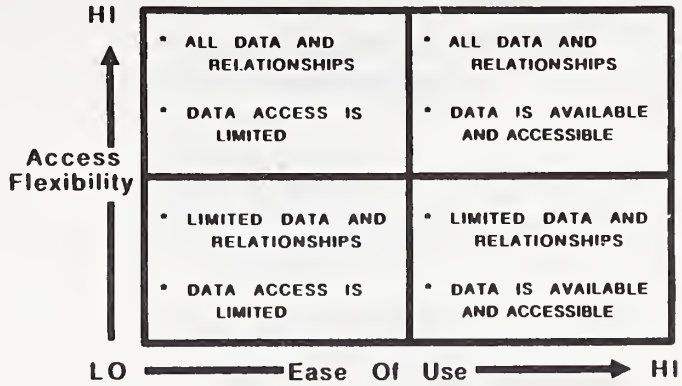


Figure 2

SYNERGISTIC DATA PLANNING

DATA PLANNING APPROACHES

- Business Model Development
  1. Joint Development Workshop with Users
  2. Joint Development Workshop with MIS
- Data Model Development
  1. Data Class Workshops within Business Function
  2. Data Class Workshop across Business Functions
- Information System Development
  1. Identify Development Opportunities
  2. Match System Data Needs to Data Model

Figure 3

SYNERGISTIC DATA PLANNING

BUSINESS MODEL DEVELOPMENT

1. JDW Workshops to Identify & Define:
   - Functions
   - Events
   - Interfaces
   - Data Requirements
2. Develop Functional Data Flow Diagram
3. For Functions, Identify:
   - Organizations Performing
   - Systems Currently Supporting
4. For Data Requirements, Identify:
   - Data Classes
   - Source Documents
   - Master Files & Data Bases Containing

Figure 4

SYNERGISTIC DATA PLANNING

JOINT DEVELOPMENT WORKSHOP

- Team: End–users & MIS
- Preworkshop:
  - Train Team Members
  - Identify Subject Matter Experts
  - Establish Initial Functions
  - Schedule Function Workshops
- Workshop:
  - Identify & Define Business Function Attributes
  - Develop Initial Data Flow Diagram
- Post–workshop:
  - Document Workshop Results
  - Establish "Functional" Data Flow Diagram
  - Review with End–users & Revise
- Management Review

Figure 5

SYNERGISTIC DATA PLANNING

DATA MODEL DEVELOPMENT

1. "Brainstorm" Workshops
   - Collect User Data Views
   - Identify & Define Data Items
   - Identify Data Sources
   - Collect Normalization Information
2. Collect Data Sources
3. Normalize Workshop & Source Data
4. Document:
   - Data Item Definitions
   - Entity Data Item Lists
   - Relationships
   - Linkage Diagram
5. Revise Based on User Reviews

Figure 6

SYNERGISTIC DATA PLANNING

DATA CLASS WORKSHOP

Day 1: Conduct Workshop

Days 2 – 6: Define & Document:

- Data Views & Data Items
- Source Document Data Items
- Normalize Data

Days 7 – 8: Users Review Documentation

Days 9 – 14: Define & Document:

- User Responses
- Current System Master Files/Data Bases
- Normalize Data

** Conduct Next Workshop **

Figure 7

AFTER THE WORKSHOPS

- Complete Normalization of:
  - Source Documents
  - System Master Files/Data Bases
- Review With Users:
  - Data Item Definitions
  - Normalized Entities
- Review With Management:
  - Identify Informational Data Base Opportunities
  - Identify Operational Data Base Opportunities
  - Establish Tactical Plan/Recommendations

Figure 8

SYNERGISTIC DATA PLANNING

USAGE DATA MODEL

- Users' Perception Of:
  - Data Input & Storage Volumes
  - Frequency of Data Input & Data Usage
  - Key Information Requirements
  - Data Access
- Users' Required Subset of Data Model

Figure 9

SYNERGISTIC DATA PLANNING

INFORMATION SYSTEM MODELING

1. Identify Development Opportunities:
   - Opertaional Systems
   - Informational Systems
2. Identify Subset of Business Model:
   - Functions to be Developed
   - Data Required
3. Identify Subset of Data Model:
   - Entities & Relations Required
   - Identify Data Dependencies
   - Identify Data Sharing Opportunities
4. Establish Prototype Data Base

Figure 10

SYNERGISTIC DATA PLANNING

PRODUCTS

- Function Definitions
- Events Content
- Data Stores Content
- Interface Definitions
- Data Item Definitions
- Entity – Data Item Lists
- Relation Definitions
- Linkage Diagram
- Data Dependencies Identified
- Logical Data Base Structures
- Key Product Lists

Figure 11

DATA & SYSTEM DEVELOPMENT LIFE CYCLE

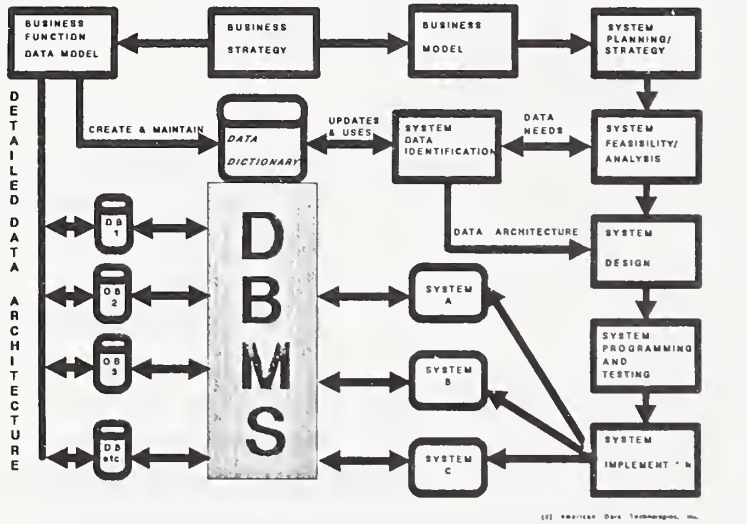


Figure 12

SYNERGISTIC DATA PLANNING

WHAT'S REQUIRED?

- User Commitment
- MIS Commitment
- 5 to 9 Months per Business Function
- Dedicated Core Team Members
- User Coordinators Dedicated

- User Coordinators Available

Figure 13

SYNERGISTIC DATA PLANNING

REQUIREMENTS FOR STARTING

- MANAGEMENT COMMITMENT:
  - Ensure Organizational Support.
  - Provide Full and Part-time Personnel.
  - Establish Data Stewards.
- METHODOLOGY:
  - Select Approach.
  - Train Personnel.
- PLANNING TEAM:
  - Core Team Must Be Full-time (80 to 90%).
  - Organizational Coordinators - 40 to 50%

Figure 14

SYNERGISTIC DATA PLANNING

PLANNING TEAM

- CORE TEAM COMPOSITION:
  - Data Administration
  - Data Base Administration
  - System Development
  - Subject Matter Experts (End-users)
- ORGANIZATIONAL COORDINATORS:
  - Identify Workshop Attendees
  - Identify Additional Subject Matter Experts
  - Coordinate User Participation
- DATA STEWARDS:
  - Responsible For Data Security
  - Authorized Access To Organization's Data

Figure 15

Synergistic Data Planning

IDEAL TEAM COMPOSITION

FULL TIME MEMBERS

One From Each User Organization

At Least One From Each Organization:

- Data Administration
- Data Base Administration
- System Analysis

PART TIME MEMBERS

Subject Matter Experts:

- 1 to 2 Days Per Workshop Per Manager Organization

Current System Subject Matter Experts:

- 1 to 2 Days Per Workshop Per Manager Organization

Figure 16

CONVERSION ISSUES

- Roles & Responsibilities
- Comprehensive Test Plan
- Controls for Parallel Tests
- Staffing
- Scheduling
- New Technology
- Integration of Old & New

Figure 17

SYNERGISTIC DATA PLANNING

1. Develop Functional Business Model
2. Identify Data Classes
3. Conduct Data Modeling Workshops
4. Define Current Master File/DB Data
5. Normalize Data
6. Identify Development Opportunities
7. Establish Tactical Plans

Figure 18

BUILDING THE DATA ADMINISTRATION FUNCTION

MODERATOR

Frankie E. Spielman
National Bureau of Standards

PANELISTS

Ronald Shelby
Elaine Hill

BUILDING THE DATA ADMINISTRATION FUNCTION

Ronald Shelby
American Management Systems, Inc.

Building an effective data administration function has proven to be a difficult, although not impossible, task during the 1980's. Corporations and government agencies have established data administration functions for a variety of reasons. Some organizations cite the need to reduce data redundancy, and others target the need to increase data sharing. While motivations for establishing data administration vary, a common problem transcends the initial efforts of most organizations; they underestimate the difficulty of establishing an effective data administration function. This presentation addresses the keys to building the data administration function, and reviews three organizations' approaches to building data administration.

If you look at the purpose of data administration carefully, it should be clear that its ultimate aim is to ensure that the data asset of an organization is managed appropriately. Every organization wants the right people to receive accurate data at the right moment in time to support their business, programmatic, or administrative duties. Few organizations anticipate the changes required to make this vision reality.

Usually, the demand for better data management results from changes in the role of information systems in an organization. As an organization's information systems are used to provide operational and management decision-making support, data quality and coordination problems become visible. Coordination and synchronization of data between two systems is an example of the problems organizations experience.

While these problems result from a significant change in the role of the information systems function, management's first impulse is to treat only the symptom (data problems). Organizations often create a data administration function after they experience clear data management problems, but before they understand the root causes of these problems. Most heads of data administration step into this situation on their first day on the job. Data administration is charged with responsibility for fixing the "data problem" before others recognize the breadth of change required to effect the solution.

If you face this, or a similar, situation in your organization, there are three keys to success that you should address to build a successful data administration function. These keys are organizational fit, credibility, and building a data management infrastructure.

Organizational Fit

First, stop for a minute to examine the relationship between the data administration (DA) function and its place in the organization. Does the DA function report to the individual who is its chief management proponent, or sponsor? If the head of DA does report to its sponsor, then it will be easier to gain support for the innovations required to implement an effective data administration function. If the head of DA doesn't report to the sponsor, DA could be deprived of the direct support it needs. Data administrators without a direct report to their sponsor should cultivate a "dotted-line" relationship by promoting a project that the sponsor favors.

Second, ensure that the functions assigned to DA match its charter, or mission. If DA is chartered to "manage the organization's data, ensure its accuracy, security, and availability to those who need it," but plays no role in information systems development, operation, and management, then the DA function can't expect to fulfill its charter. A charter that outlines a specific responsibility for managing an organization's data calls for a robust DA function with a significant role in information systems development and management.

Third, review the degree to which a DA function's authority and accountability correspond. The fit between authority and accountability should be viewed as an extension of the function/charter issue that has an even greater impact upon the chances for a successful DA implementation. If DA is accountable for establishing a policy and standards-driven DA program, as is often the case in government agencies, then the function must have the authority to issue policy and standards directives. Correspondingly, accountability for ensuring that future data requirements are met requires authority to establish and support a strategic data planning effort, and monitor the implementation of the plan during information systems development.

Credibility

Establishing and keeping credibility is the second key to success. Planning the data administration function's implementation, managing expectations, and managing innovation as you establish data administration will boost DA's credibility and chances for success.

Good planning is an excellent way to provide focus while building the DA function. A plan should be specific about the DA function's activities and deliverables, and it should be realistic. For instance, building a strategic data plan and delivering a data and systems architecture might be realistic

during the first year of DA at one organization, while being improbable at a second organization. Because the DA function will impact many parts of an organization, an initial plan should look at least three to five years ahead. Update the plan each year to keep it current.

Use planning to shape the expectations others have of the DA function. Planning, describing, and meeting clearly stated objectives is the best way to manage expectations. Another way to manage expectations is making clear, realistic promises during meetings, and delivering on these promises.

Saying you will "ensure we manage data as a resource" sounds vague and can create very different expectations in the minds of different listeners. Don't use this phrase too often. Instead, try to manage expectations by using clearer expressions such as "we will ensure management's reporting needs can be met by the new database." Break down management's expectations for your function into smaller, feasible tasks, then complete these tasks so that DA's credibility will grow.

Facilitating and managing innovation is also an important facet of building credibility. Desirable innovations include an increased emphasis upon data when information systems are developed, and changing the role of the information systems function in the organization. You should encourage increased end-user involvement in the system development process, and change the system development approach of your organization from a process-driven to a data-driven approach. Above all else, encourage broad participation in planning and implementing the innovations your organization will need. While innovation can be planned from the top-down, it must be implemented from the bottom-up. Build partnerships with management and floor-level staff to support the innovations DA will need to succeed.

Data Management Infrastructure

Building and operating a data management infrastructure is the third key to success for the data administration function. Whether data administration is implemented by building a strong data administration organization, by distributing functions between existing organizational units, or by taking a policy and standards-driven approach, you must establish a solid data **management infrastructure**: a substructure or underlying foundation for managing data appropriately. This infrastructure should include policy, staff skills, methodologies, and automated tools.

Data is clearly an important asset in the information age, and organizations should have clear policies based upon this premise. Data administration policies should address data planning and

acquisition, requirements definition, data stewardship, documentation, and quality assurance. Policies should specify improved organizational effectiveness as a key objective for the innovations data administration will propose.

Building new skills for information system, end-user, and data administration staff, and applying these skills in the workplace is another key to successful innovation. Data modeling, documentation, and database design skills are needed to develop information systems. Increasingly, the data administration function participates in building data-sharing systems that touch all levels within organizations. Building, maintaining, and using these systems requires significant staff skills.

New methodologies for designing and building information systems are an important aspect of building a data management infrastructure. Process-oriented development methodologies that view data as something that is a by-product of processes were effective ways of automating the back room functions addressed by information systems in the past. However, these methodologies have become ineffective in the electronic, information-rich organizations of today. Information engineering methodologies that rely upon data-driven techniques such as data modeling and event analysis, should be taught and used to improve the quality of the data your organization has available in the future.

After other elements of your infrastructure are in place, use automated software tools such as data dictionaries, database management systems, and CASE (Computer Assisted Software Engineering) software to manage data effectively. Like other significant resources (money, facilities and people are other resources that are usually the subject of special, organization-wide management attention), data is too pervasive and too important to manage without automated tools. Using these automated tools, the data administration function can participate in planning and building the systems and databases that will take your organization into the 21st century.

Automated CASE software, dictionaries, and database management systems are being used to do more than manage data. They are being used to build information systems that use data to strategic advantage. For example, many corporations analyze data describing existing customers to design new products tailored specifically to these customers. Other corporations can provide customer service worldwide by having data profiling their customers available to their offices around the world. Government agencies are eliminating fraud and abuse from many programs by matching and analyzing data. As the price of technology falls and organizations' data management skills increase, data's value as an asset will increase in most organizations. An effective DA function will help organizations to leverage the asset further.

Ensuring good organizational fit, maintaining credibility, and building a data management infrastructure are the keys to success in building an effective data administration function. Technology, especially data dictionaries and CASE software, has improved the short-term benefits for organizations that build an effective DA function. However, the keys to building the data administration function involve more than effective use of these technologies.

Mr. Shelby is Senior Principal at American Management Systems, Inc., working in data administration and systems life cycle management, strategic planning, database design, and information engineering for both government and private corporations. Previously, he worked in data administration for Travelers Insurance (Canada) and the Department of Interior's Office of the Secretary, and was a consultant with Applied Data Research (ADR). He has 10 years experience as a practitioner and consultant in the data management field. He is currently Vice-President of the International Data Administration Management Association.

DATA ADMINISTRATION MANAGEMENT

THE KEYS TO BUILDING THE DATA ADMINISTRATION FUNCTION

Ron Shelby
Gaithersburg, Maryland
May 17, 1988

ams

Figure 1

DATA ADMINISTRATION'S PURPOSE

Ensuring the data asset
is appropriately managed.

ams

Figure 2

BUILDING THE DATA ADMINISTRATION FUNCTION

KEYS TO SUCCESS

1. Organizational Fit
2. Credibility
3. Data Management Infrastructure

ams

Figure 3

KEYS TO SUCCESS

ORGANIZATIONAL FIT

1. Placement vs. Sponsor
2. Functions vs. Charter
3. Authority vs. Accountability

ams

Figure 4

KEYS TO SUCCESS

CREDIBILITY

1. Planning
2. Managing Expectations
3. Managing Innovation

ams

Figure 5

KEYS TO SUCCESS

DATA MANAGEMENT INFRASTRUCTURE

1. Policy
2. Staff Skills
3. Standards and Methodology
4. Automated Tools

ams

Figure 6

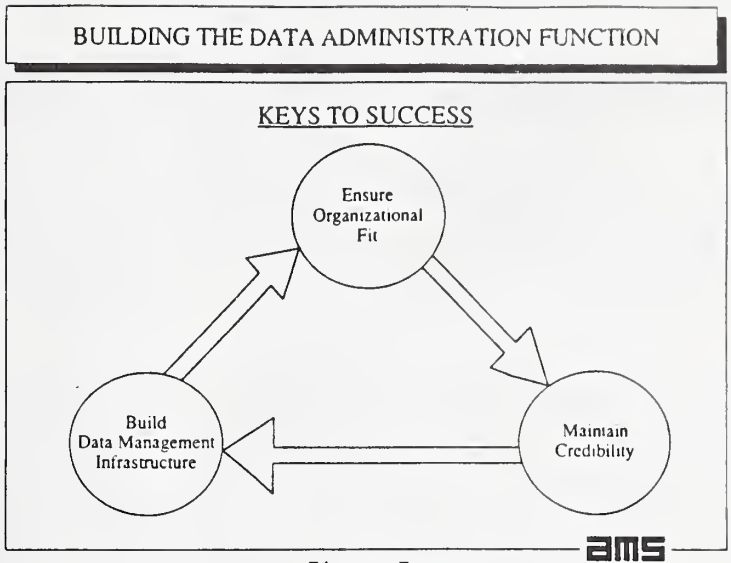


Figure 7

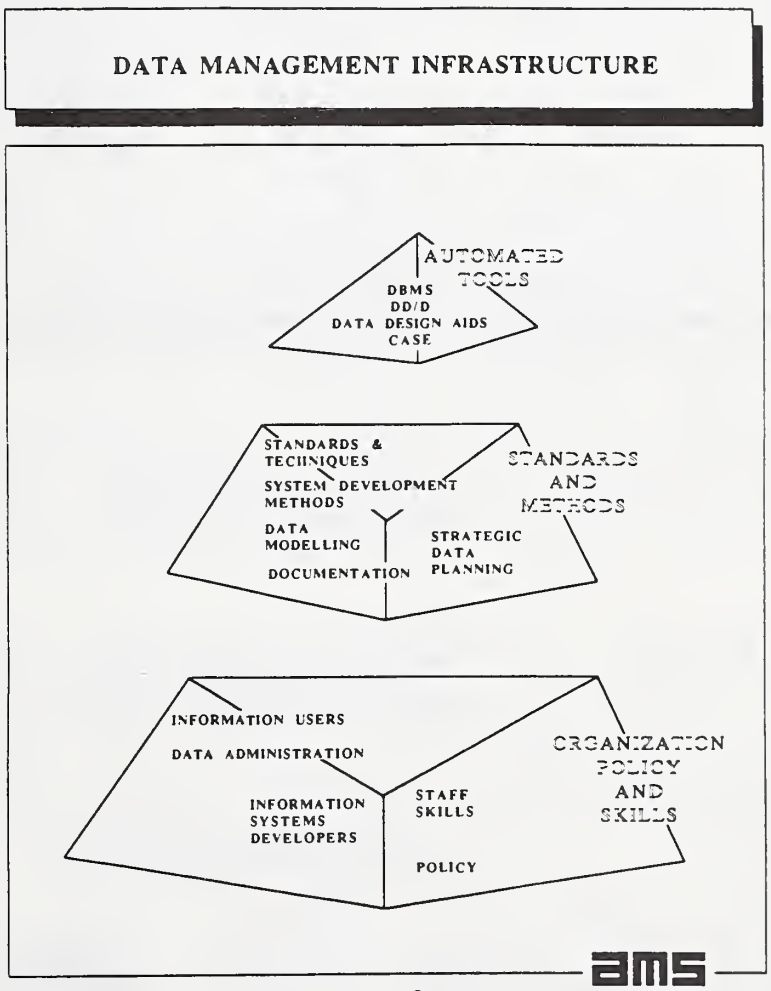


Figure 8

EXAMPLES

	ORGANIZATONAL FIT	CREDIBILITY	INFRASTRUCTURE
XXX Financial Services	• Reported to Sponsor • Narrow Functions/Broad Charter • Authority ≠ Accountability	• 5 Year Plan/Annual Revision - Incremental Approach • Shaped Expectations • Directly Managed Innovation	• Policy Lagged Implementation - Linked to Business • Staff Trained/Managed • Revised Standards/System Development/Change Management/Organizational Structure • Extensive Tool Use
ABC Food Stores	• Unclear Sponsor • Broad Functions/Charter • Authority Unclear/ Accountability Clear	• Annual Plan - High Impact Approach • Raised Expectations • Documentation = Only Innovation	• No Policy • Training Deferred - Only DA Staff Involved • Standards Deferred/Data Modelling Methods Addressed • DBMS/Dictionary
Agency A	• Reported to Sponsor • Broad Functions/Charter • Little Authority/ Accountability	• Annual Plan - Low Profile Approach • Constrained Expectations • Facilitated Innovation	• Policy-Driven Approach • No DA Staff/Other Staff Training Not Addressed • System Development Standards/Data Planning Policy • No Automated Tools

Figure 9

EXAMPLES - EPILOGUE

XXX Financial Services — Stable data administration practices from business planning through operational systems management.

ABC Food Stores — Eliminated data administration after two years. CEO re-started effort two years later to support strategic planning.

Agency A — Slow progress, but widespread interest in data management across the agency.

Figure 10

BUILDING THE DATA ADMINISTRATION FUNCTION

Elaine C. Hill
Chief of Naval Personnel

The Navy Manpower, Personnel, and Training (MPT) business includes policy, planning, and implementation for the acquisition, training, distribution, and career management for over one million Navy members. Other areas supported by MPT functions include military pay, financial management and civilian personnel management.

The MPT business is an information-dependent business. Our information management goals are to get the right data to the right people at the right time. The information processing environment includes over 150 automated systems running on diverse and geographically separated computers that support world-wide users. A number of information deficiencies have been highlighted over the last several years that have a direct impact on critical Navy management decisions and operations. These deficiencies have been caused by the development of stovepipe systems that supported individual user needs, and technology that was relatively inflexible and could not provide a shared data environment. As the Navy moved toward total force management and the information needs broadened, a spider web of systems interfaces grew. These interfaces were generally unmanaged as a whole and were poorly documented. The results were data redundancy and inconsistency, extremely high maintenance costs, and management decisions made on poor quality data.

The MPT community began addressing the data problem with the establishment of the Data Resource Management (DRM) Program in 1980. This program was to establish the policies and plans necessary to manage MPT data as a corporate resource. In December 1984, the Chief of Naval Personnel (CNP) Data Administration Office was established as a prototype organization to implement the DRM policies and procedures.

As the organizations have matured, we have come from an environment where technology and information systems drove the data to an environment today where data is driving the technology and the systems. Information Resources Management (IRM) is the framework we are using to manage the data and information, as well as the information technologies. Our goal is to reach a Chief Information Officer environment where data is truly managed as a resource, as are money and people.

The MPT IRM Program addresses three major areas: program management (plans, standards, and dollars), information management (includes DRM and Data Administration), and information systems and technology. A hierarchy of guiding

policies and instructions provide the impetus and the direction for this Navy-wide program. The key concepts of the IRM program are: that information has value and cost, that we must manage information and data first, and that we must manage systems and technology to meet requirements for information and data and to improve efficiency of functional processes.

The MPT policies and guidance are incorporated in the MPT IRM instruction issued in October 1986. The instruction establishes policies and responsibilities for managing MPT data throughout the data life cycle, focuses on Data Administrators as the execution arm of the DRM policies, and delineates the roles of the many players in the DRM world.

The DRM Program has realized many accomplishments over the last five years. For the most part, benefits have been seen in the functional world through standardization of data elements, integration of data, systems, and functions, and data and management issue resolution.

There are three major concepts being used to manage the MPT data resource: Data Administration, Database Administration, and Data Dictionary/Directory Systems. In trying to develop a workable data administration program, definite roles have been defined and refined. While the need for data administration has been slowly recognized, its value and benefits are now accepted. Data administration supports users, system developers, and management. The key elements that are making data administration work are the corporate data dictionary/directory system (the MPT Information Resource Encyclopedia), data standards and guidelines, and data planning. All these things are essential in a shared data environment.

Needless to say, implementing data administration has been a tremendous learning experience. Our thoughts on what standards were necessary three years ago are somewhat different from where we are today. It is important to be flexible enough to change and to take advantage of opportunities as they arise.

The MPT Information Resource Encyclopedia (IRE) is really the cornerstone of the DRM program. We use MSP's DataManager and DesignManager products to support the IRE and our data modeling efforts. The development of the IRE has been a major undertaking and a constantly moving target. Our main goals have been to support documentation of the MPT data, to support data modeling, and to support documentation of information systems and life cycle management.

Initially, the corporate dictionary was called the Data and Information Resource Directory. Data elements from about ten major master files and about 40 systems were documented in the dictionary and resulted in about 4000 data elements. The amount

of redundancy and inconsistency in the data caused us to reevaluate our approach, standardizing element by element. We decided to redesign our dictionary and to develop standards based on two new corporate databases being built, the personnel and jobs databases. These databases represent about 60 percent of the MPT corporate data.

Congress has directed that the Navy integrate its pay and personnel systems. We realized the importance of a common IRE to that effort and so we have designed and developed the IRE together with the Navy Finance Center in Cleveland, Ohio. We are also developing an interface between DataManager and Supra, our database management system, as well as a Dbase III interface that will allow for upload and download and manipulation of IRE data by a PC.

The DRM program has also been instrumental in outlining the corporate database strategy for MPT. Data architectures have been developed over the last five years leading to the design of corporate subject area databases. These corporate databases are the cornerstone of the data/technology strategy phase of our IRM program. Key areas offering challenges and opportunities include data integration, standards, introduction of new technology, and transition from the existing environment.

An area that both DRM and Data Administration staff members spend a lot of time and effort doing is analyzing and resolving data issues. These issues arise because of such things as new data requirements being levied on the Navy or by the Navy, erroneous processing by existing systems, or implementation of new systems or modifications to existing systems that must be coordinated across a wide range of systems and users.

I have highlighted throughout the briefing the need to manage data as a resource. One result of managing data should be improved data quality. In order to measure quality one must have standards and metrics. We have begun to develop guidance for data quality assurance. While there is a good bit of literature and guidance for quality assurance of systems, there is very little that addresses the data itself. Our focus is on the accuracy, timeliness, and completeness of the data. We expect to learn a great deal over the next year as we attempt to apply the concepts and criteria.

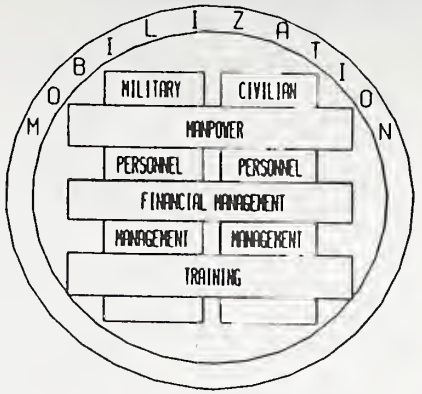
I have tried to cover where we have been over the last four to five years as we have developed the DRM Program. Our growth has been tremendous. In 1984, we had six people and today we have over 40 people in the DRM and Data Administration offices. We are about to embark on a reorganization that will put the error research and correction function under the data administration function. We are designating Associate Data Administrators in customer support centers and geographically separate commands

where corporate data knowledge exists. We have consistently spent more dollars than we budgeted because of the burgeoning interest in the program and the visibility of the issues we deal with.

In summary, a few lessons have been learned. To be successful, you have to get organizational commitment, from top management to the working level. You must continually build support. You must show successes. You must educate people and show them how you can make their life easier or better.

Ms. Hill is Deputy Director of Data Administration for the Chief of Naval Personnel. She was responsible for creating the Data Administration Office at the Navy Military Personnel Command which has served as a prototype organization for the Navy. She has spent 14 years in the Federal Government holding positions in systems operations, systems analysis and programming, functional analysis and database design, distributed processing, and database planning and evaluation.

THE NAVY MPT BUSINESS.....



.....IS AN INFORMATION DEPENDENT BUSINESS

Figure 1

9.87.03

MPT BUSINESS

- INFORMATION MANAGEMENT
  - 1/2 BILLION RECORDS SUPPORTING 2 MILLION PEOPLE
  - 150 AIS'S WITH 8,000 DATA ELEMENTS
  - DIVERSE AND GEOGRAPHICALLY SEPARATED CUSTOMERS
  - PRIMARILIY APPLICATION ORIENTED
  - PROLIFERATION OF PC'S
  - DECENTRALIZED/DISTRIBUTED DATA BASES
  - MULTIGENERATION/MULTIVENDOR HARDWARE AND SOFTWARE
  - INSTALLATION OF NEW TECHNOLOGY

Figure 3

9.87.04

MPT BUSINESS

- PERFORMANCE OF MPT-RELATED FUNCTIONS SUCH AS:
  - PERSONNEL ACQUISITION
  - TRAINING
  - DISTRIBUTION
  - CAREER MANAGEMENT
  - PAYMENT
  - FINANCIAL MANAGEMENT
  - INFORMATION MANAGEMENT
- MPT FUNCTIONS ARE INFORMATION DEPENDENT AND IT IS IMPERATIVE THAT THE:
  - ✓ RIGHT INFORMATION IS PROVIDED TO THE
  - ✓ RIGHT PEOPLE AT THE
  - ✓ RIGHT TIME

Figure 2

9.87.05

THERE ARE INFORMATION DEFICIENCIES IN THE MPT BUSINESS

- SEPARATE REGULAR AND RESERVE PERSONNEL DATA BASES DO NOT ADEQUATELY SUPPORT MOBILIZATION
- FRAGMENTED AND REDUNDANT TRAINING SYSTEMS RESULT IN POOR QUALITY INFORMATION
- LACK OF ACCURATE AND TIMELY DATA CAUSED DECERTIFICATION OF THE NAVY PAY SYSTEM BY GAO
- FLEET COMMANDERS CAN NOT MANAGE MANPOWER RESOURCES WITHOUT ACCESS TO MPT DATA BASES
- PERSONNEL EXPENDITURES AND END-STRENGTH LEVELS CAN NOT BE RECONCILED DUE TO INCONSISTENT DATA

Figure 4

9.87.01

TECHNOLOGY CREATED A PROBLEM

- DATA REDUNDANCY
- DATA INCONSISTENCY
- PROLIFERATION OF DATA FILES
  - Redundancy
  - High Maintenance Cost
  - Lengthy Processing Time
  - Downstream Impacts
- MULTIPLE INTERFACES

Figure 5

9 87 09

REAL PROBLEMS

- RECOGNITION OF THE VALUE OF INFORMATION
- UNDERSTANDING OF DATA
- DEFINITIONS
- FUNDAMENTAL STANDARDS (DATA ELEMENT NAMES)
- COMMONALITY OF DATA
- TRANSITION FROM FIRST GENERATION THINKING TO FOURTH GENERATION TECHNOLOGY

Figure 6

THE MPT IRM DIRECTION

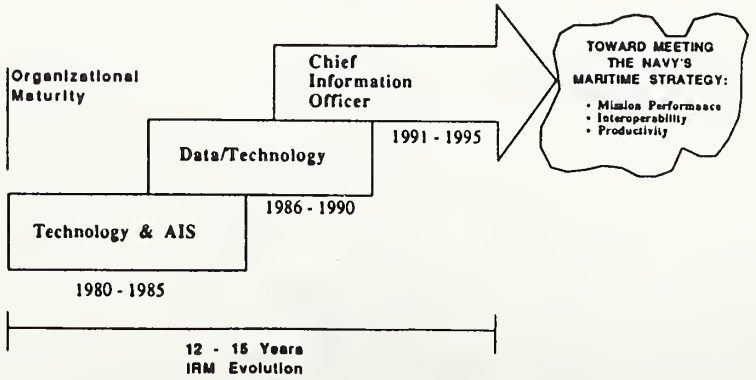


Figure 7

INFORMATION RESOURCE MANAGEMENT PROGRAM

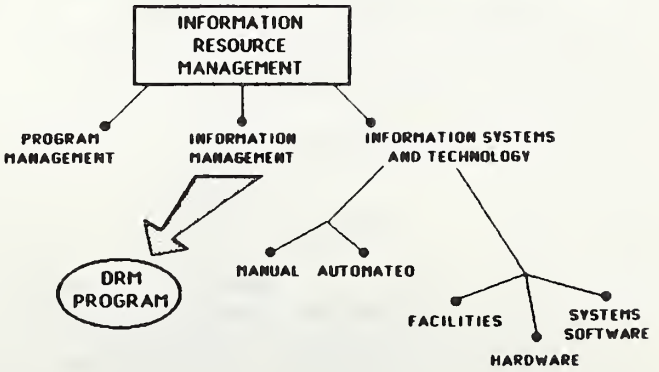


Figure 8

9 87-17

DRM ACCOMPLISHMENTS

1 NAVY-WIDE MPT MANAGEMENT ISSUES
- Active Duty End Strength
- Individuals Account Management
- Mobilization
- Dependents
- Central adjucation of Security Clearances
- Health Professional Program
- Standardization of NEC's
- CINCLANTFLT Retention Date
- Common Order Writing

2 FORMAL FUNCTIONAL INVOLVEMENT
- Personnel/Pay, Mobilization, Training, Distribution and Accounting Interface Conferences
- Military Personnel Data Integration Steering Committee (OP-01, NMPC, CNRC, OP-09R, CNRF, CNET, MEDCOM, NFC, NAVCOMPT, EPMAC AND NRPC)
- Architectures

3 OSD/DMDC LIAISON
- DOD DATA ELEMENT STANDARDIZATION
- Reserve Component Common Personnel Data System
- Active Component Personnel Data System
- Defense Enrollement Eligibility Reporting System
- New GI Bill

Figure 9

THREE KEY CONCEPTS IN MANAGING

THE MPT DATA RESOURCE

Data Administration

Plan, document, manage, and control the data resources of an entire organization

Database Administration

Physically implement and maintain the organization's data infrastructure

Data Dictionary/Directory System

A central repository of information on the data used by the organization and how the organization uses that data

Figure 10

ROLES IN THE

MPT DATA ADMINISTRATION ENVIRONMENT

Users

- *Define the data needed to accomplish their goals*
- *Accept ownership responsibility for individual data elements*
- *Define the logical structure of their data*

Data Administration

- *Provide standards for defining, documenting, and using data*
- *Assist users in planning for their usage of data*
- *Resolve data issues*
- *Maintain the data dictionary/directory system as a central repository of information about data*
- *Assist users in the logical design of new data requirements*
- *Manage the data resource*

Database Administration

- *Physically implement logical data base designs*
- *Support users in implementing applications*
- *Optimize physical implementation*

Figure 11

Data Administration Supports...

Users:
- *By maintaining a central repository of information about data (Data Dictionary/Directory System)*
- *By resolving data-related conflicts*

ADP Developers:
- *By providing a central source of information about existing data*
- *By automating documentation of data and systems*
- *By providing standards for data access, storage, and definition*

Management:
- *By performing the long range planning of the organization's data needs*
- *By ensuring that data needs are defined in step with organization's objectives*
- *By ensuring that technology is kept in step with organization's data needs*

Figure 12

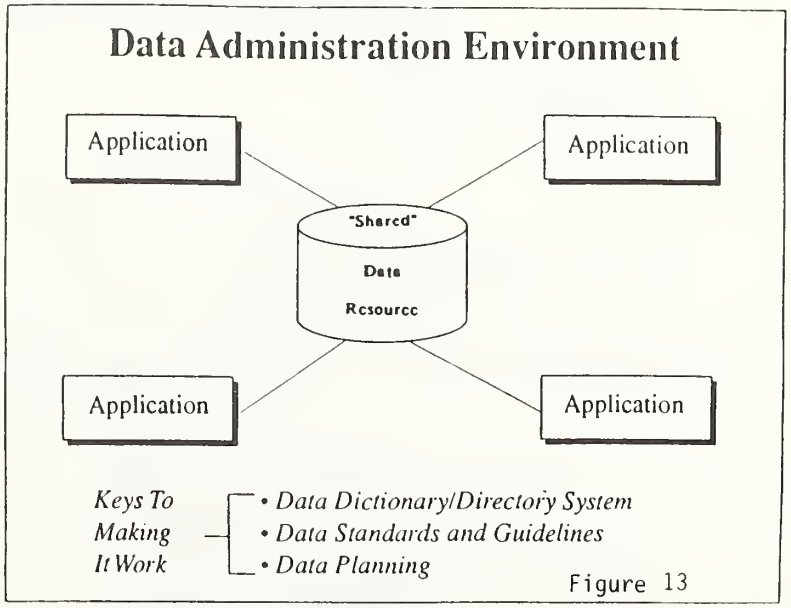


Figure 13

STANDARDS-

"Even if you're on the right track,
you'll get run over if you just sit there!"
Mark Twain

A Tough But Necessary Undertaking

Figure 15

STANDARDS-

"Even if you're on the right track,
you'll get run over if you just sit there!"
Mark Twain

A Tough But Necessary Undertaking

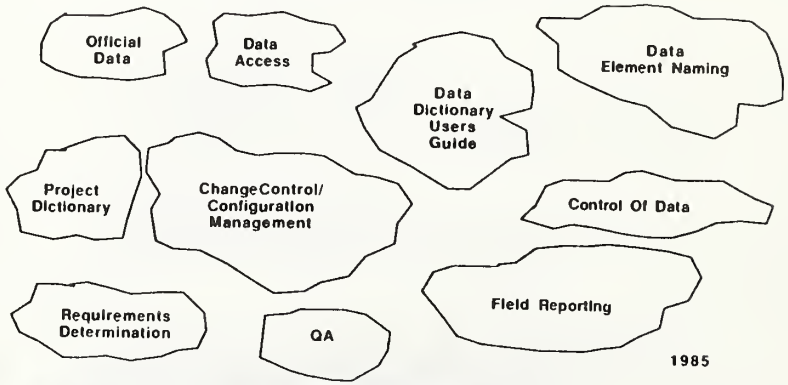


Figure 14

MPT INFORMATION RESOURCE ENCYCLOPEDIA (IRE)

Objectives

- Document Sources and Uses of Data
- Support Data Model
- Support Life Cycle Methodology

Other Functions

- Support Information Center
- Support Application Generator
- Support Distributed Processing
- Document Communications Network
- Maintain Hardware Inventory
- Support Data Security Planning
- Forms Control
- Change and Problem Control
- Standardize Coding and Tables
- Document Systems, Jobs, Programs, Files, Data Bases
- Standardize Names

Figure 16

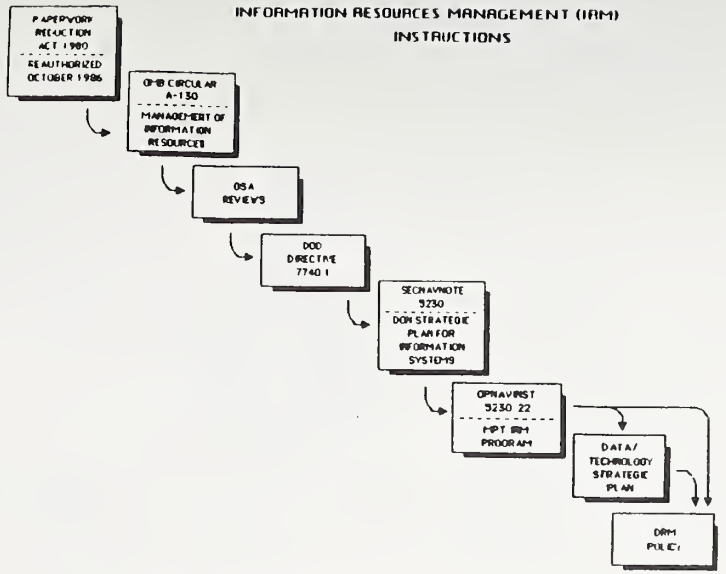


Figure 17

9 87-24

Information Resources Management (IRM)

Key Concepts:

1. Information has value and cost.
2. Manage information and data first.
3. Manage systems and technology to:
   - *Meet requirements for information and data;*
   - *Improve efficiency of functional processes.*

Figure 18

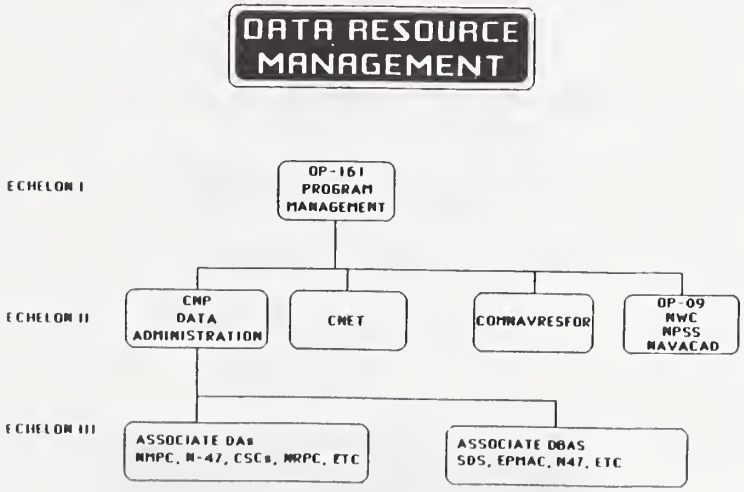


Figure 19

9 87 29

MPT DRM POLICY

- First effort to formally centralize guidance and policy for MPT data management
- Developed in support of the MPT IRM Instruction (OPNAVIST 5230.22, 6 Oct 1986)
- Draft policy (instruction) circulated for comments in February 1987
- Establishes policies and responsibilities for managing MPT data throughout the data life cycle
- Focuses on data administrators as the execution arm of the DRM piece of the IRM program
- Delineates the roles of
  - Data Resource Manager
  - Data Administrators
  - Data Base Managers
  - Data Element Sponsors
  - Functional Users
  - ADP Professionals

Figure 20

STATUS OF THE MPT INFORMATION RESOURCE ENCYCLOPEDIA PROJECT ...

Three steps in IRE Project:

1. Design and Build the Dictionary Structure
2. Populate the Data Dictionary
3. Maintain the Data Dictionary Metadata

Figure 21

MPT AND PAY DD/DS ENVIRONMENT 1985

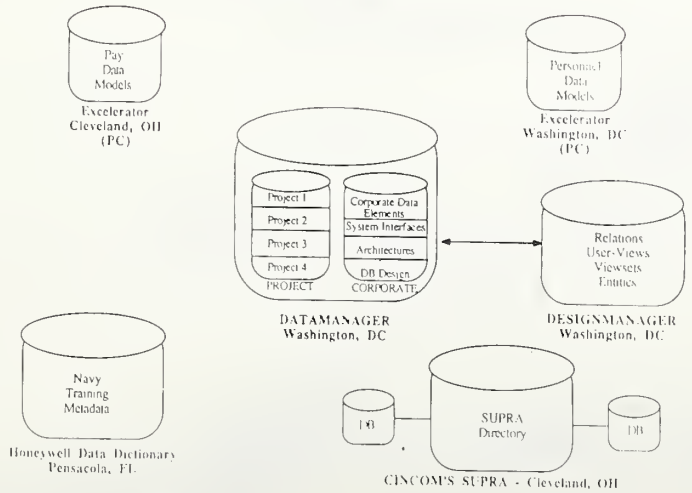


Figure 23

STATUS OF THE MPT INFORMATION RESOURCE ENCYCLOPEDIA PROJECT

1. **Design and Build the IRE Structure**
   - Finalized structure
   - Implemented structure (DATAMANAGER)
   - Testing structure
   - Finalizing guideline
2. **Populate the IRE**
   - Identified mandatory attributes for each Member-Type
   - Defined line numbering conventions for each Member-Type
   - Designed screens for metadata entry
   - Provide dictionary training
   - Currently supporting TFMMS, IMPDB, Functional Architectures, NAVMIPPS
3. **Maintain the IRE**
   - Designing interface to SUPRA DBMS
   - Drafted data element registration procedures
   - Developing User's Guide
   - Tested automated upload from floppy disks

Figure 22

CURRENT MPT AND PAY
INFORMATION RESOURCE ENCYCLOPEDIA (IRE) ENVIRONMENT

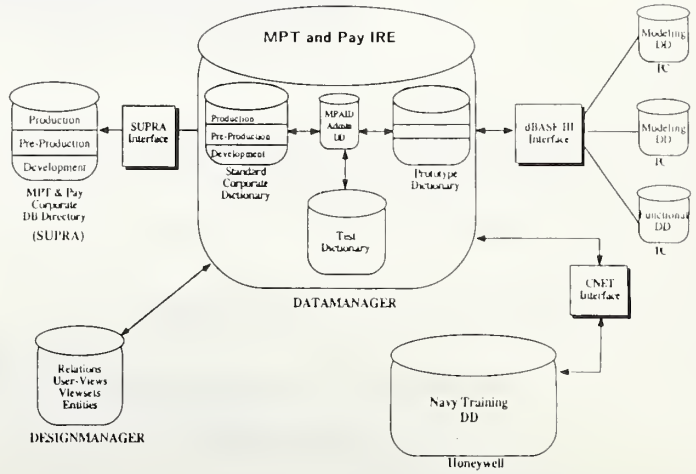


Figure 24

CORPORATE DATA BASE STRATEGY

- o Data integration
- o Standards
- o New technology
- o Transition

Figure 25

DATA RESOURCE MANAGEMENT ACTIONS

TRAINING
- NTIS Economic Analysis
- NEC Documentation
- Training Management Perspective
- Training Quota Management Data Flow

MANPOWER
- Activity Data Flow
- Corporate Activity File
- MANPOWER Data Base Files
- Mobilization
- TFMMS

PERSONNEL

PROGRAMS
- "Archiving" Automated Data
- Data Element Sponsorship
- IMPOB
- MOB Cookbook
- Military Locator System
- Nepers
- Officer File Disparity
- Promotion Board "Failsafe Plan"
- Promotion History Interface Active/Reserve
- PRT Officer Data "Autoscan"
- SDS MOB Event
- "1 Day" IRR Call-up
- CINCPACFLT Quality Detailing Issue

ISSUES
- AOC-NAVCAD Issues
- "Dependency" Data Issues
- Exception Reporting
- Medical Classification Codes
- Navy Sea College Program
- Officer Strength Formula
- Reserve Data Elements
- Retirement Programs Data Issues
- SSN-Verification SSA
- SPD Codes
- Joint Specialty Code
- VSBS Data Analysis

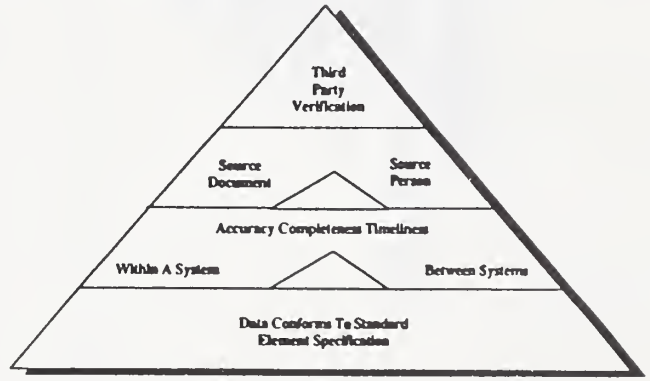
Figure 26

DATA QUALITY ASSURANCE
GUIDELINE FOR STANDARDS

- DATA MUST CONFORM TO ELEMENT SPECIFICATIONS
- DATA MUST BE ACCURATE, COMPLETE AND TIMELY (ACT)
- IRE SHOULD BE USED TO DOCUMENT
  - DATA SOURCE
  - DATA VALIDATION CRITERIA
  - DATASTORE/OFFICIAL DATASTORE

Figure 27

PROJECT APPROACH



Data Quality Depends On A Solid Foundation

Figure 28

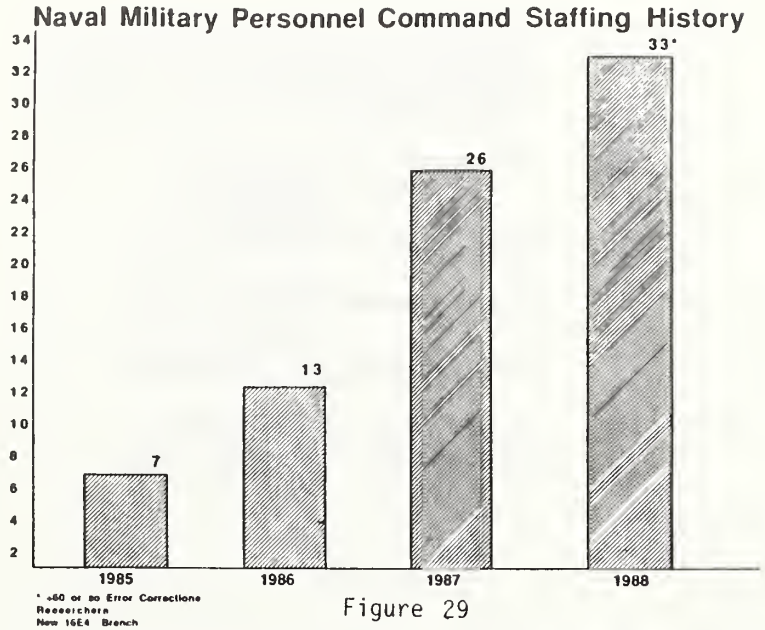


Figure 29

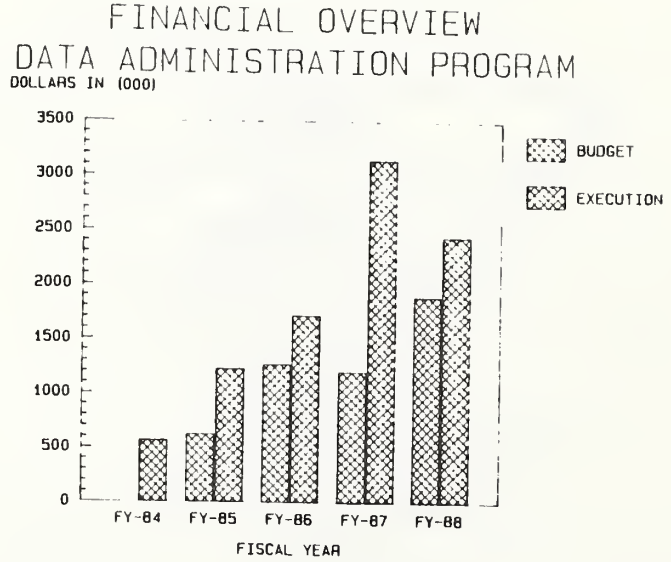
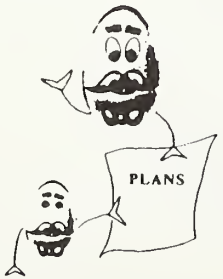


Figure 30

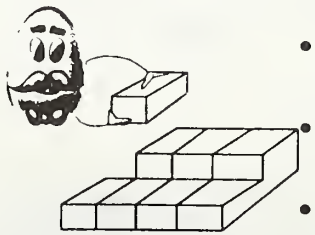
Get Organizational Commitment



- At all levels of management
- Gear presentation to audience
- Develop reputation for being right
- If at first you don't succeed, don't give up
  - *Show short term results*
  - *Collect success stories*
  - *Look for what works*

Figure 31

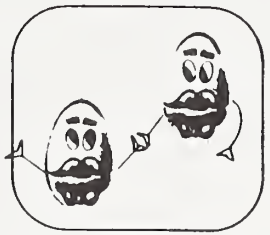
Keep Building Support



- Show immediate payoffs
- Produce useful results, tools
- Provide technical support
- Educate
- Keep doing your homework

Figure 32

Co-Opt The Opposition



"It 's better to have them inside shooting out than outside shooting in."

- Form working groups
- Listen to objections
- Have a court of last resort
- Start work on the easy parts while you resolve the hard parts
- Don't bring things up too early
- Have a solution ready

Figure 33

GENERAL SESSION

GAINING AND KEEPING MANAGEMENT COMMITMENT FOR DATA ADMINISTRATION

Herb Jacobsohn
Technology Information Products Corporation

GAINING AND KEEPING MANAGEMENT COMMITMENT FOR DATA ADMINISTRATION

SUMMARY OF REMARKS

Herb Jacobsohn
Technology Information Products Corporation

One of the big problems for Data Administration is getting the management commitment to manage the data resources. There are three primary interrelated areas addressed in this session: planning, business orientation, and implementation.

Planning is the key factor essential for managing information or data resources. The reason for current problems in justifying information planning is that we focus on solving them through technical solutions. Instead, the focus needs to be on the business issues and the crucial areas for investing the MIS dollars that will provide the best payback. It's difficult to develop a good information plan but it is important and can provide many benefits.

An example of one organization justifying the use of common data was illustrated by showing the cost benefits in developing generic code for their data elements. Initially, they selected four data elements and developed generic code, using it on 15 projects with a net savings of $62,000 for the first year. With this success, they then repeated the process on another four data elements resulting in a net savings of $108,000. This savings of $170,000 for the first year was the result of having one data element representing one thing or one definition. The savings for large databases over several years can be significantly large.

Another approach for driving the point across is to emphasize the dollars spent over a five-year period on MIS/DP. The cost can be staggering. Are managers happy with what they are getting? If not, then what can be done? If the way things are done is not changed, then it is likely that twice that amount will be spent over the next five years with similarly poor results. It is going to cost resources to develop systems anyway, why not add a little extra cost that can provide significant long-term benefits?

An MIS plan and Information Architecture supporting the business functions must be developed. The plan and architecture must address the sharable databases, integrated information systems, and distributed data processing. The cost may seem high but in relation to the total MIS budget it can be a miniscule amount. The payoff generally seems too far away and most benefits are long term. However, this can be offset by placing emphasis on

the short-term payoffs, for example, eliminating forms, identifying problems to be solved, and new ways of doing business.

The ideal approach to the development of information systems is top-down, completing the information architecture, data models, and normalized relations. The problem is that it takes too much time and costs too much money. On the other hand, bottom-up doesn't work because we end up with incompatible, incomplete, application oriented databases. The solution is to do some of both, completing 60-70% of the top-down approach on a business basis, then implementing bottom-up using the products developed during the top-down effort. While implementing on a bottom-up basis, the additional details can be completed, such as determining information usages, normalizing the relations, and developing accurate data models. Using this approach, significant results can be obtained in a short period of time as well as laying the framework for long-term benefits.

In summary, the business perspective must be the focus of attack for determining problems to be solved. Once these are determined then the technical problems can be resolved. Of course, from a central information resource or data resource management perspective, the data dictionary, encyclopedia, or repository is the key tool that aids in this effort.

Mr. Herb Jacobsohn is President of Technology Information Products (TIP) Corporation. He has been a leader in the data processing industry for 33 years. He has laid the groundwork for and influenced the establishment of software technology that has made data processing systems, information management and data management a reality. In 1981, he founded TIP, a company specializing in software technology for managing data throughout the systems life cycle.

WHY DO WE HAVE PROBLEMS JUSTIFYING INFORMATION PLANNING?

WE TRY TO DO IT FROM AN MIS PERSPECTIVE

WE USE TECHNICAL TERMS, E.G.,

- SUBJECT AREA DATABASES
- NORMAL FORMS
- SEMANTIC INTEGRITY
- CANONICAL SYNTHESIS
- ENTITY RELATIONSHIP MODEL
- 4GL'S

Figure 1

WHY DO WE HAVE PROBLEMS JUSTIFYING INFORMATION PLANNING?

"ITS HARD!"

PLANNING TAKES TOO LONG

PLANNING IS A DIRTY WORD

PLANNING IS A NO - NO

Figure 3

WE SHOULD TALK ABOUT

CRUCIAL AREAS FOR INVESTMENT OF MIS DOLLARS

SEQUENCE FOR INVESTMENT

HOW DO WE GET THERE FROM HERE?

REQUIREMENTS:

- HARDWARE
- SOFTWARE
- NETWORK

MANAGING PC'S

TAKING ADVANTAGE OF NEW TECHNOLOGY

Figure 2

IN REALTY, PLANNING IS VERY IMPORTANT

A GOOD ANALOGY IS . . .

ASK MANAGEMENT WHAT THEY DID THE LAST TIME THEY:

- BUILT A BUILDING
- BUILT A FACTORY
- MOVED INTO A NEW FACILITY

Figure 4

DID THEY:

- HIRE SPECIALISTS:

  ARCHITECTS?
  INTERIOR DESIGNERS?

- KNOW WHERE THE EXECUTIVE OFFICES WOULD BE?
- LAY OUT TELEPHONES AHEAD OF TIME?
- LAY OUT PRODUCTION LINES BEFOREHAND?
- KNOW WHERE THE REST ROOMS WOULD BE?

YOU NEVER NOT PLAN FOR A BUILDING

Figure 5

TELL MANAGEMENT

"YOU'RE GOING TO SPEND $N MILLION ON INFORMATION SYSTEMS IN THE NEXT X YEARS. I CAN'T BELIEVE YOU'RE NOT GOING TO DO THE SAME KIND OF PLANNING!"

IF THAT DOESN'T WORK,

AT YOUR NEXT JOB TRY . . .

Figure 7

BEFORE THE BUILDING IS CREATED:

ARCHITECTS
CONTRACTORS
MASONS
PLUMBERS
ELECTRICIANS
MOVERS

WORK TO A COMMON PLAN: (ARCHITECTURE)

(BLUE PRINT)

Figure 6

JUSTIFICATION FOR COMMON DATA

FIRST 4 DATA ELEMENTS

COST OF GENERIC CODE	$32,000
COST OF APPLICATION - SPECIFIC CODE (1 APPLICATION)	$25,000
SAVINGS (LOSS)	($7,000)

Figure 8

JUSTIFICATION FOR COMMON DATA

SAVINGS FOR FIRST YEAR - FIRST 4 DATA ELEMENTS

INDEPENDENT DEVELOPMENT COST (15 PROJECTS)	$97,000
COST OF USING COMMON ROUTINES	$28,000
SAVINGS	$69,000
LOSS ON GENERIC CODE DEVELOPMENT	($7,000)
NET SAVINGS	$62,000

Figure 9

JUSTIFICATION FOR COMMON DATA

FIRST YEAR SAVINGS:

FIRST 4 DATA ELEMENTS	$ 62,000
NEXT 4 DATA ELEMENTS	$108,000
TOTAL FOR 8 DATA ELEMENTS	$170,000
SAVINGS PER DATA ELEMENT	$ 21,250

Figure 11

JUSTIFICATION FOR COMMON DATA

SAVINGS FOR FIRST YEAR - NEXT 4 DATA ELEMENTS

INDEPENDENT DEVELOPMENT COST (15 PROJECTS)	$ 149,000
COST OF GENERIC CODE	$ 41,000
SAVINGS	$108,000

Figure 10

FURTHER JUSTIFICATION

THIS IS THE SAVINGS FROM HAVING COMMON DATA:

ONE THING

ONE DEFINITION

IMAGINE THE SAVINGS IF IT INCLUDED

ONE PLACE - SUBJECT DATA BASES

ONE NAME - DATA ADMINISTRATION (IRM/DRM)

Figure 12

WE'VE SPENT OVER

$XXX,000,000

IN THE PAST 5 YEARS ON

MIS/DP!

Figure 13

IF WE DON'T CHANGE THE

WAY WE DO THINGS,

WE'LL SPEND

2 TIMES

$XXX,000,000

IN THE NEXT 5 YEARS!

WE PROBABLY WON'T HAVE

ANYTHING MUCH BETTER!

Figure 15

ARE WE HAPPY

WITH WHAT WE'VE GOT?

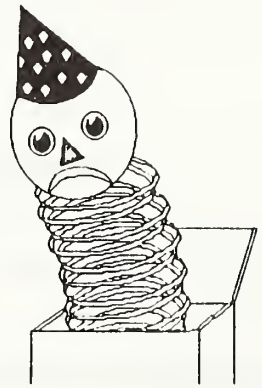


Figure 14

THERE ARE/WILL BE:

- TECHNOLOGICAL OPPORTUNITIES
  - RELATIONAL DBMS
  - 4GL'S
  - EXPERT SYSTEMS
  - ETC.
- COMPETITIVE PRESSURES
- FINANCIAL PRESSURES
- RESOURCE SHORTAGES

Figure 16

WE MUST DEVELOP AN EFFECTIVE MIS PLAN

- STRATEGIC AREAS FOR INVESTMENT
- INFORMATION ARCHITECTURE TO SUPPORT BUSINESS
  - SHARABLE (SUBJECT) DATA BASES
  - INTEGRATED INFORMATION SYSTEMS
  - DISTRIBUTED DATA & PROCESSING
- HARDWARE, SOFTWARE AND COMMUNICATION NETWORKS TO SUPPORT BUSINESS

Figure 17

COST: $250,000 - $300,000

TIME: 3 - 4 MONTHS!

.00333 OF YOUR $ XXX,000,000

Figure 18

PAYOFF

- COORDINATE DEVELOPMENT OF APPLICATIONS & DATA BASES
- REDUCE MAINTENANCE BY XX%
- REDUCE MACHINE COSTS
- REDUCE NETWORK COSTS

Figure 19

WHY DO WE HAVE PROBLEMS JUSTIFYING INFORMATION PLANNING?

PAYOFF IS ALWAYS TOO FAR AWAY

MOST BENEFITS ARE LONG TERM

SHARABLE DATABASES
INTEGRATED SYSTEMS
REDUCED MAINTENANCE COSTS

SHORT TERM PAYOFFS ARE NEEDED

Figure 20

PSYCHOLOGICAL ISSUES

- HISTORICAL "GET IT DONE TODAY" ATTITUDE (INSTANT GRATIFICATION)
- FEAR OF THE UNKNOWN
- FEAR OF THE KNOWN
- "WHAT'S IN IT FOR ME" SYNDROME
- WE'VE TRIED "THIS" BEFORE

Figure 21

PEOPLE ISSUES

- COMMITMENT
- RESISTANCE TO CHANGE
- IT'S HARD TO DO
- EASY TO GIVE UP
- IT'S NOT MY AREA
- "MY DATA" SYNDROME

Figure 22

DON'T TELL ME THAT I HAVE TO SHARE MY DATA.

Figure 23

MANAGE EXPECTATIONS

- STAGE BENEFITS
  - SHORT-TERM (3 MONTHS)
  - MID-TERM (1-2 YEARS)
  - LONG-TERM (2-N YEARS)
- BE REALSITIC

Figure 24

SHORT TERM PAYOFFS

EXAMPLES:

ELIMINATION OF OVER 100 FORMS (> 10% REDUCTION)

IDENTIFICATION OF OVER 180 PREVIOUSLY UNIDENTIFIED PROBLEMS

IDENTIFICATION OF POSSIBLE NEW WAYS OF DOING BUSINESS

NEW INFORMATION SYSTEM OPPORTUNITIES IDENTIFIED

Figure 25

NEW WAYS OF DOING BUSINESS

- IDENTIFY DUPLICATE BUSINESS FUNCTIONS
- IDENTIFY FUNCTIONS SUPERVISED/MANAGED BY MORE THAN ONE ORGANIZATION
- IDENTIFY FUNCTIONS THAT DON'T SHARE ANY INFORMATION WITH OTHER FUNCTIONS IN "SAME" BUSINESS
- IDENTIFY DUPLICATE GENERATION OF SAME INFORMATION

Figure 26

FORM ELIMINATION

- TRACE ALL FORMS THROUGH INFORMATION USAGE MODELS
- VERIFY USAGE (OR NON-USAGE!) VIA FUNCTION/STORE MATRIX
- VALIDATE NON-USAGE (IF REQUIRED) VIA DETAIL ACTIVITY MODELS
- CANCEL/ELIMINATE NON-USED FORMS
- CONSOLIDATE SIMILAR FORMS

Figure 27

NEW INFORMATION/SYSTEM OPPORTUNITIES

- IDENTIFY REDUNDANT "CREATES" OF SAME DATA
- IDENTIFY REDUNDANT "STORES" OF SAME DATA
- IDENTIFY DIFFERENT PLACES/ORGANIZATIONS WHERE PIECES OF DATA GROUPS ARE CREATED
- IDENTIFY EXCESSIVE HANDLING OF DATA
- IDENTIFY EXCESSIVE/REDUNDANT TRANSMISSIONS OF DATA

Figure 28

MID-TERM BENEFITS

- ELIMINATION/REDUCTION OF REDUNDANT SYSTEMS
- ELIMINATION/REDUCTION OF REDUNDANT DATA STORAGE
- IMPLEMENTATION OF SYSTEMS TO ADDRESS REAL BUSINESS PROBLEMS
- IRM IS INSTALLED

Figure 29

LONG-TERM BENEFITS

- SHARED DATA BASES
- SHARED APPLICATIONS
- MINIMAL REDUNDANCY
- ABILITY TO EASILY INSTALL NEW TECHNOLOGY

Figure 30

TOP DOWN

IDEAL:

COMPLETE TOP-DOWN DATA MODEL

NORMALIZED RELATIONS

PROBLEMS:

TAKES TOO MUCH TIME (12-24 MONTHS)

COSTS TOO MUCH (CAST OF THOUSANDS)

Figure 31

BOTTOM UP

DOESN'T WORK:

- PROJECT BY PROJECT
- APPLICATION DATABASES
- NO COMMON LANGUAGE
- NO PERSPECTIVE OF WHOLE BUSINESS
- MISS IMPORTANT THINGS/ASSOCIATIONS
- NEVER DONE ?

Figure 32

SOLUTION

DO SOME OF BOTH:

BOTTOM-UP, ON AN IMPLEMENTATION BASIS:

- SUBJECT DATA BASE
- APPLICATION/SYSTEM

DEVELOP LOGICAL/RELATIONAL MODEL:

- GATHER INFORMATION USAGES
- PRODUCE NORMALIZED RELATIONS
- VALIDATE AGAINST INFORMATION ARCHITECTURE
- CORRECT MODELS
- IMPLEMENT

Figure 34

SOLUTION

DO SOME OF BOTH:

TOP-DOWN, ON A BUSINESS BASIS:

- USE BUSINESS TERMINOLOGY
- OVERALL BUSINESS PERSPECTIVE
- USER DIRECTED
- 60-70% COMPLETE

Figure 33

BENEFITS

ACHIEVE SIGNIFICANT RESULTS IN A SHORT PERIOD OF TIME

FIND BUSINESS PROBLEMS

IDENTIFY OPPORTUNITIES

PRODUCE INFORMATION ARCHITECTURE:

- SUBJECT DATABASE PLAN
- SYSTEM DEVELOPMENT PLAN

Figure 35

BENEFITS

PROJECTS FIT IN OVERALL BUSINESS CONTEXT

PROJECTS IMPLEMENTED IN "PROPER" SEQUENCE

APPLICATIONS INTEGRATED

DATABASES BUILT ON SUBJECT BASIS

Figure 36

INTANGIBLE JUSTIFICATIONS

SHAREABLE DATA BASES

- DISTRIBUTE INFORMATION TO THOSE WHO NEED IT
  - EASILY
  - PROPERLY
  - EFFECTIVELY
- EXTRACT INFORMATION FOR:
  - INFORMATION CENTERS
  - PC NETWORKS
- DISTRIBUTE PROCESSING

Figure 38

SHORT TERM PAYOFFS:

- COMPLETE UNDERSTANDING OF
  - BUSINESS OPERATIONS
  - INFORMATION REQUIREMENTS
- IDENTIFICATION OF REDUNDANCIES:
  - FUNCTIONS
  - INFORMATION USAGES
  - INFORMATION STORES
- STREAMLINING BUSINESS OPERATIONS
- OPTIMIZE DOLLARS BEING INVESTED IN MIS

Figure 37

INTANGIBLE JUSTIFICATIONS

STANDARDIZATION OF INFORMATION

BUSINESS DESCRIPTIONS
- DOCUMENTED
- AGREED ON
- WELL-DEFINED

CAN EFFECTIVELY USE CENTRAL INFORMATION RESOURCE
- DATA DICTIONARY
- ENCYCLOPEDIA
- REPOSITORY

Figure 39

NBS-114A (REV. 2-8C)

U.S. DEPT. OF COMM. **BIBLIOGRAPHIC DATA SHEET** (See instructions)	1. PUBLICATION OR REPORT NO. NIST/SP-500/159	2. Performing Organ. Report No.	3. Publication Date October 1988

4. TITLE AND SUBTITLE

Data Administration: Management and Practice

Proceedings of the First DAMA Symposium

5. AUTHOR(S)

Judith J. Newton and Frankie E. Spielman

6. PERFORMING ORGANIZATION *(If joint or other than NBS, see instructions)*

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
(formerly NATIONAL BUREAU OF STANDARDS)
U.S. DEPARTMENT OF COMMERCE
GAITHERSBURG, MD 20899

7. Contract/Grant No.

8. Type of Report & Period Covered

Final

9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS *(Street, City, State, ZIP)*

National Capital Region of the Data Administration Management Association (NCR DAMA)

Federal Data Management Users Group (FEDMUG)

Association for Federal Information Resources Management (AFFIRM)

10. SUPPLEMENTARY NOTES

Library of Congress Catalog Card Number: 88-600586

☐ Document describes a computer program; SF-185, FIPS Software Summary, is attached.

11. ABSTRACT *(A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here)*

This publication constitutes the proceedings of a one-day symposium at the National Bureau of Standards on May 17, 1988. It was jointly sponsored by the National Capital Region of the Data Administration Management Association (NCR DAMA), the Federal Data Management Users Group (FEDMUG), and the Association for Federal Information Resources Management (AFFIRM).

The symposium provided attendees with an opportunity to share the insights of leaders in the Data Administration field. Special emphasis was given to the factors which contribute to successful implementation of Data Administration management policies and practical methodologies.

12. KEY WORDS *(Six to twelve entries; alphabetical order; capitalize only proper names; and separate key words by semicolons)*

administration; data; data administration; data architecture; information; information asset management; life cycle; management

13. AVAILABILITY

☒ Unlimited

☐ For Official Distribution. Do Not Release to NTIS

☒ Order From Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.

☒ Order From National Technical Information Service (NTIS), Springfield, VA. 22161

14. NO. OF PRINTED PAGES

122

15. Price

ANNOUNCEMENT OF NEW PUBLICATIONS ON COMPUTER SCIENCE & TECHNOLOGY

Superintendent of Documents,
Government Printing Office,
Washington, DC 20402

Dear Sir:

Please add my name to the announcement list of new publications to be issued in the series: National Institute of Standards and Technology Special Publication 500-.

Name _____

Company _____

Address _____

City _____ State _____ Zip Code _____

(Notification key N-503)

U.S. Department of Commerce
National Institute of Standards and Technology
(formerly National Bureau of Standards)
Gaithersburg, MD 20899

Stimulating America's Progress
1913-1988