DATA BASE SYSTEMS DEVELOPMENT

EVOLUTION TO STRATEGIC MANAGEMENT

INPUT

About INPUT

INPUT provides planning information, analysis, and recommendations to managers and executives in the information processing industries. Through market research, technology forecasting, and competitive analysis, INPUT supports client management in making informed decisions.

Continuous-information advisory services, proprietary research/consulting, merger/acquisition assistance, and multiclient studies are provided to users and vendors of information systems and services (software, processing services, turnkey systems, systems integration, professional services, communications, systems/software maintenance and support).

Many of INPUT's professional staff members have more than 20 years' experience in their areas of specialization. Most have held senior management positions in operations, marketing, or planning. This expertise enables INPUT to supply practical solutions to complex business problems.

Formed as a privately held corporation in 1974, INPUT has become a leading international research and consulting firm. Clients include more than 100 of the world's largest and most technically advanced companies.

INPUT OFFICES

North America

Headquarters
1280 Villa Street
Mountain View, CA 94041-1194
(415) 961-3300
Telex 171407 Fax (415) 961-3966

New York
959 Route 46 East, Suite 201
Parsippany, NJ 07054
(201) 299-6999
Telex 134630 Fax (201) 263-8341

Washington, D.C.
1953 Gallows Road, Suite 560
Vienna, VA 22182
(703) 847-6870 Fax (703) 847-6872

International

Europe
Piccadilly House
33/37 Regent Street
London SW1Y 4NF, England
(01) 493-9335
Telex 27113 Fax (01) 629-0179

Paris
52, boulevard de Sébastopol
75003 Paris, France
(33-1) 42 77 42 77 Fax (33-1) 42 77 85 82

Tokyo
Saida Building
4-6, Kanda Sakuma-cho
Chiyoda-ku, Tokyo 101, Japan
(03) 864-0531 Fax (03) 864-4114

DECEMBER 1989

DATA BASE SYSTEMS DEVELOPMENT

INPUT®

1280 Villa Street, Mountain View, California 94041-1194

(415) 961-3300

Published by
INPUT
1280 Villa Street
Mountain View, CA 94041-1194
U.S.A.

Information Systems Program (ISP)

Data Base Systems Development

UMDA • 301 • 1989

Abstract

The data management function of today's information systems organization is facing new issues and challenges. It is evolving from a technological support function to organization-wide administration of the information network's underlying definition and structure.

In this report, INPUT considers the current state of the data management function, the tasks it performs, how its responsibilities are changing, the impacts of new technology on it, and its shifting role.

This report considers changes that have taken place over the past year, ascertains progress and problems, and provides a basis for assessing the speed at which users are adopting relational and distributed data base technology. Where appropriate, direct comparisons are made to last year's research.

The goal of this report is to raise awareness about the changing role of this function, to suggest and support objectives and priorities for the next few years, and to indicate how the function itself is changing due to the accelerating shift to relational data base technology.

INPUT believes that the data management function needs to become more important, and with it the individuals who perform this critical task.

This report contains 77 pages and 39 exhibits.

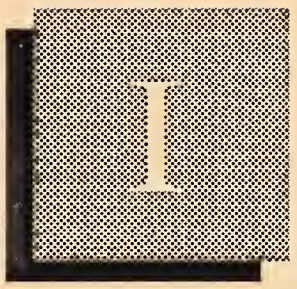
Table of Contents

Table of Contents (Continued)

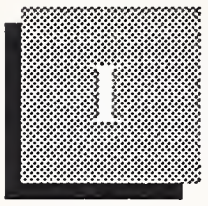
Table of Contents (Continued)

Exhibits

Exhibits (Continued)



Introduction



Introduction

A Overview

After years of serving as a technical controller for data base management system (DBMS)-based applications, data management is on the edge of a major change.

Fundamental DBMS technology is changing, those who develop DBMS-based applications are changing, and the information network of most organizations is quickly becoming distributed over multiple computing levels.

The task of doing data base design and other things a data base administrator has done are quickly becoming only a small part of the data management process:

- The installed base of IBM's relational data base management system, DB/2, has expanded significantly over the past two years and is approaching the installed base of IBM's IMS and DL/1 hierarchical data base products.

- The fastest growing software companies include those whose entire product line is based on relational products.

- It is becoming increasingly common to place relational DBMS systems on departmental minicomputers. PCs and workstations that are under the control of users and were used primarily for file transfer are beginning to implement relational systems.

- Acceptance of relational systems is fostered by the adoption of a standard language, SQL, which standard data bases did not have. Portability has become mandatory. Oracle's relational DBMS runs on almost any minicomputer—under DOS on personal computers, and under MVS on mainframes.

- Using data definition processes inherent in relational technology, it is possible to distribute data bases across multiple computers. As distributed data bases become more common, the entire process of data management will undergo major changes.

It is against this background of major change, and with the belief that data management is one of the most important obligations of the central information systems organizations, that INPUT provides this update to its report on the data management function.

The focus of the report is to assess the current status of development and provide an assessment of progress over the past year in dealing with problems and trends that were noted. With this as a base, users and vendors can identify the speed at which changes are taking place and can spot potential problem areas.

This report concludes that the data management function will require the concentrated attention of senior information systems executives over the next five years.

B

Scope

This report looks at the data management function of information systems from organizational and process standpoints. It considers problems and trends discovered during INPUT's 1988 research and notes changes that are taking place. It identifies any new problems or trends that have emerged over the past year.

The overall objectives are to characterize the current state of the data management function and to set a framework for information systems managers to use in directing this function over the next few years.

1. Objectives

The objectives of this report and the related research are to:

- Identify underlying trends and issues in data management relative to technology, responsibilities and resources
- Track progress with relational DBMS technology and its use by information systems and users
- Understand the magnitude of change that is impacting the data management function
- Set objectives for the data management function and its administration for the 1990s

2. Definitions

The following terms and definitions will be used throughout the report:

- Data Management—The broad, overall process of providing a definitional basis for the information network of the organization. The data management process incorporates the elements of data architecture, data administration, and data definition in a technical and nontechnical sense.

  The term data management will also be used to refer to the organization that addresses the overall data management process.

- Data Administration—The more traditional of the data control processes that addresses the technical definition and control tasks required by DBMS technology—in particular traditional data base management systems.

- Data Base Administrator (DBA)—Both the task and the specific job assignment of the individual who establishes the detailed definition logic required in programming a DBMS application.

- Relational Data Base Management System (RDBMS)—A data base management system that is based on the SQL language standard and that provides a flexible, table-like view of data relationships.

- Traditional Data Base Management System (DBMS)—A traditional data base management system that is hierarchical or uses the Codasyl standard for specifying data relationships.

- Three-Tiered Computing Network—INPUT uses a three-part (tier) structure to define the computing/information networks of most organizations. Tier one is the central or mainframe tier. Tier two is the middle level and is usually minicomputer-based. Tier three is the workstation, which now more commonly is a computer, rather than a terminal.

C Methodology

To gain an understanding of the state of the data management function, INPUT undertook, in 1988, a significant interview task. One hundred data base managers were interviewed, using a comprehensive questionnaire designed to:

- Assess current status of the data management organization. (What are the major issues and how is the data management function changing?)

- Determine the current DBMS environment, who is being served and where DBMSs are being supported

- Identify the level and type of application development activity with relational and distributed DBMS technology

- Learn whether, and how, the end user is using DBMS technology on mainframes, minicomputers and microcomputers

In addition, a few data managers were interviewed using a shortened version of the questionnaire shown in Appendix A. The purpose of the additional research was to obtain a preliminary indication of the impact of new data base technology.

For this report, INPUT expanded last year's secondary questionnaire to include questions from the extensive survey document. This was to ensure that the baseline information was comparable and would provide a basis for comparison.

The key focus of this year's research was to assess changes that have taken place, to assess progress toward objectives, and understand the speed at which movement is being made toward relational and distributed data base technology. The questionnaire also included questions designed to discover whether data managers are able to measure their success.

1. Research Demographics

The focus of the research was the data management of larger, Fortune 500-sized information systems organizations.

The targeted interviewee was the data manager within the central or corporate information systems organization. INPUT wanted to understand data management from the central point of view, not from that of an operating division that may have its own data administration staff performing DBA functions.

Where possible, INPUT interviewed the same individual who participated in the previous research.

The data managers came from the ten industry sectors listed in Exhibit I-1.

EXHIBIT I-1

Research Demographics

- Ten industries
  - Discrete manufacturing
  - Banking and finance
  - Transporation
  - Utilities
  - Retail
  - Process manufacturing
  - Insurance
  - Medical
  - Services
  - Wholesale distribution

2. Other Research

Other research included a search of current published information to identify any noted changes in the industry or trends that had emerged over the past year.

This report also draws upon INPUT's 1988 report, *Developments in Data Base Technology*.

D

Report Structure

The remaining chapters of this report address the following topics.

- Chapter II—Executive Overview: a summary of the findings and conclusions of the report.
- Chapter III—Data Management: Current Environment: a status report on the central data management function.
- Chapter IV—Impact of Relational Data Base Technology: a status report on the adoption of RDBMS technology and the implications for data management.

- Chapter V—Impact of Other Technologies: a look at distributed data base technology, the impact of networks, the data dictionary, and the relationship to computer-assisted systems engineering (CASE).

- Chapter VI—Data Management: Future Environment: INPUT's proposal for data management's role in the information systems function of the 1990s.

- Chapter VII—Conclusions and Recommendations: INPUT's conclusions from research and recommended priorities for senior information systems management regarding the data management function for 1989 and the first few years of the next decade.

E
Related Reports

The following INPUT reports were used in this study and provide additional information for the reader.

- *Distributed Data Base Management—An Early Look*
- *Information Systems Planning Report, 1988*
- *Computer-Assisted Systems Engineering, Markets and Opportunities*
- *Workstation Strategies*

III

Executive Overview

II

Executive Overview

As indicated in the introduction, the primary purpose of this year's report on the state and future direction of the data administration and management process was to assess the state of the industry one year later.

Following on the heels of two reports, *Distributed Data Base Management—An Early Look* and *Developments in Data Base Technology*, INPUT was interested in learning whether data administration was beginning to respond to a changing industry or whether, once again, management would lag behind technology development.

A fundamental conclusion of the report is that not much progress has been made. Acceptance and implementation of relational technology is progressing at an increasingly rapid rate, but management processes, needed to ensure maximum benefit from corporatewide information systems, are lagging behind.

To be effective in meeting the needs of the 1990s, the data administration process must undergo a fundamental reorientation due to technological change and the need for a tightly-integrated yet distributed information network. The historic technical and mainframe orientation will not support the information network needed and expected in the 1990s.

A Changing Environment

As indicated in Exhibit II-1, there are a number of reasons why data administration must undergo change if it is to meet future needs adequately.

- The shift to relational technology is rapidly outmoding the traditional data administration function. Growing acceptance of relational technology by end users and implementation on multiple platform tiers is necessitating a shift from a centralized, technical orientation to a distributed, consultant orientation.

- Implementation of multiple data bases is placing greater pressure on data administration to take on a management role so as to ensure coordination between data base systems.
- With data bases being implemented on multiple platform tiers, significant coordination is required. Management planning is necessary to ensure effective use of multitiered systems.
- End users are increasingly interested in access to corporate data, irrespective of its location. Data access and sharing is needed both vertically and horizontally, increasing the need for comprehensive data definitions.
- Additional DBMS technology will be available very quickly. However, organizations that continue to wait for the latest technology before changing their management process will lose significant benefit in use of corporate data as a competitive asset.

EXHIBIT II-1

Data Administration—Changing Environment

- Shift to relational DBMSs
- Multiple DBMS environments
- DBMSs at all three tiers of computing
- End-user access to data managed by DBMSs
- Additional DBMS technology on the horizon

B

Effectiveness of Data Administration

Whether as a result of the shift to relational technology or a general lack of attention to the importance of managing corporate data, more than half the IS managers consider the effectiveness of their data administration to be no greater than average.

As indicated in Exhibit II-2, at least one-quarter believe that their performance is below average or even poor.

The extent to which the administrative function is effective is highly subjective, however. More than half (56%) of the managers indicated that they have no measures to indicate the performance of their administration function, as illustrated in Exhibit II-3. Approximately one-quarter (28%), indicated that they have informal measures. The majority of these are based on monitoring inquiry response time and end-user "bitch level".

EXHIBIT II-2

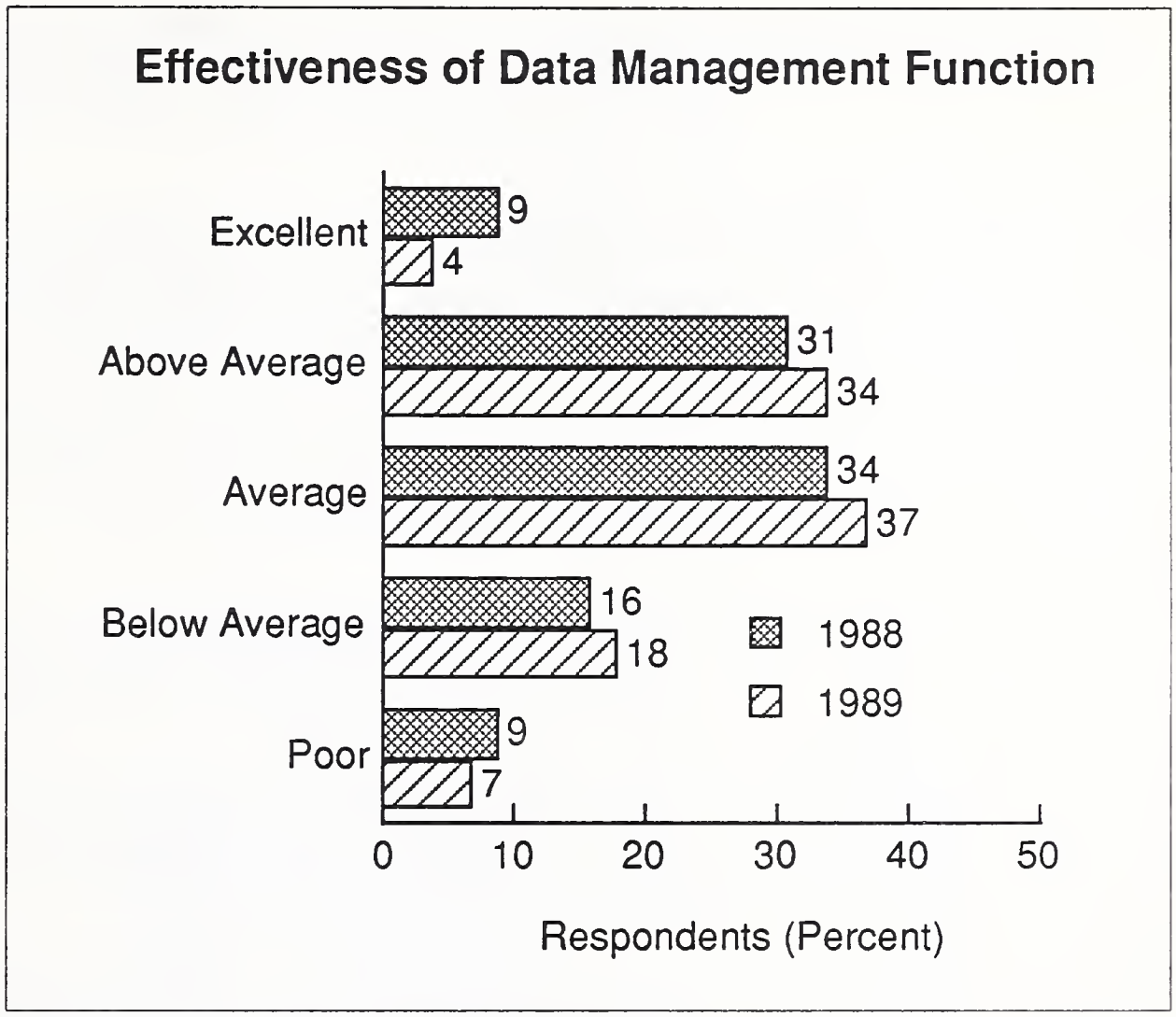
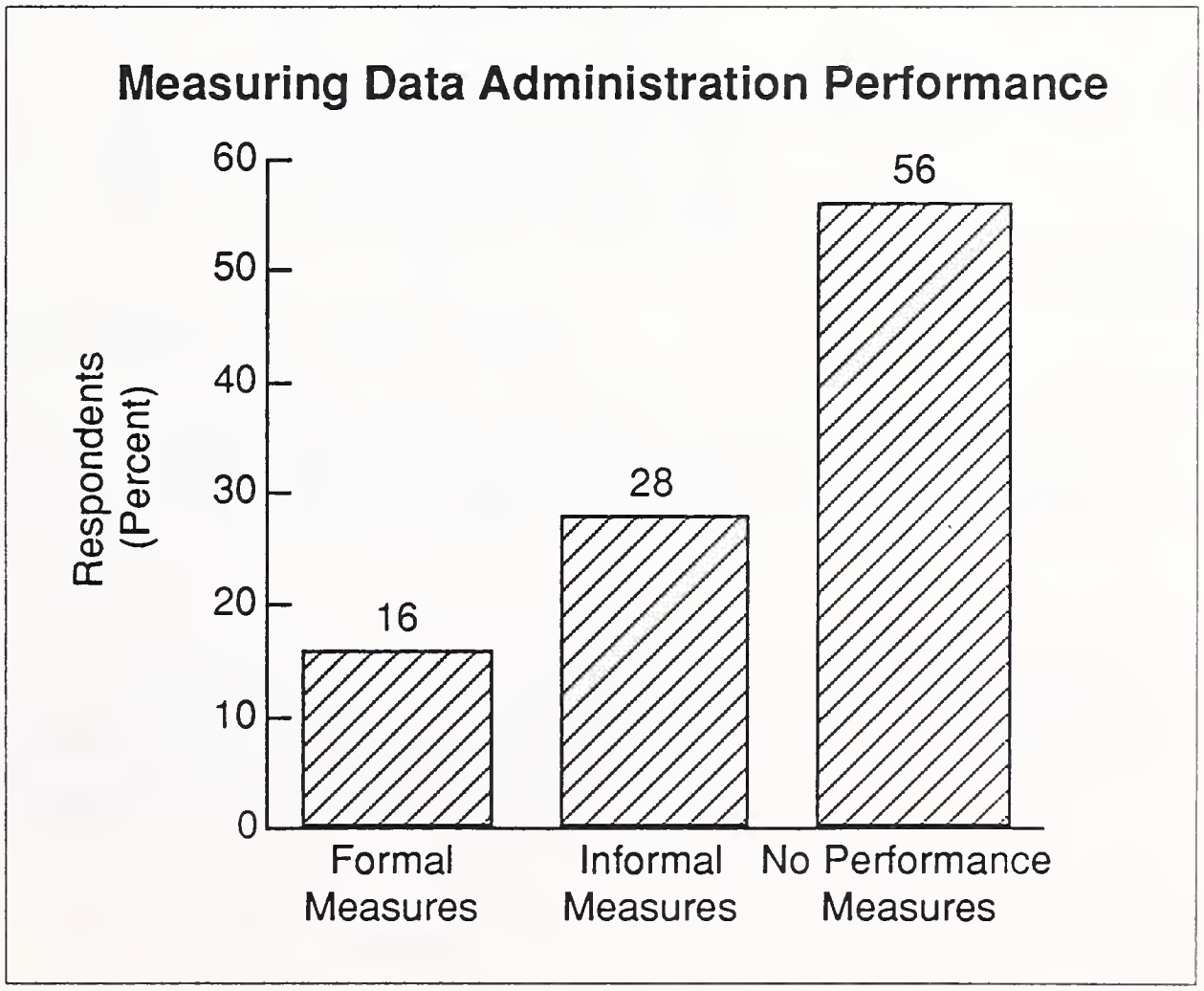


EXHIBIT II-3



Although a dramatic shift in effectiveness rating would not be expected over the course of one year, the consistence of the responses indicated that IS managers continue to face perplexing issues about how best to meet future needs.

C
Key Issues

Information systems managers continue to voice a number of issues that are key to the continued development of DBMSs. The most frequently mentioned issues included the following:

- Strategy and Direction—Managers have a continuing concern about the strategy and direction of information systems, and DBMS in particular, as it relates to the organization.

- Integrity and Security—Most information systems managers acknowledge the emergence of distributed data base systems and have concerns about the integrity and security of distributed data files.

- Technology—Considering the speed with which technology is changing, many managers are concerned about the correct time to make a major investment. This is a growing problem, as long-term funding becomes more difficult to obtain.

- Data Access—Access of data is also a growing problem. As local- and wide-area networks expand, there is increasing demand for access to corporate data bases, irrespective of their location.

Among the key issues identified, strategy and direction were the concerns voiced most frequently, by nearly a two-to-one margin. Exhibit II-4 identifies a number of sub-issues related to strategy and direction.

EXHIBIT II-4

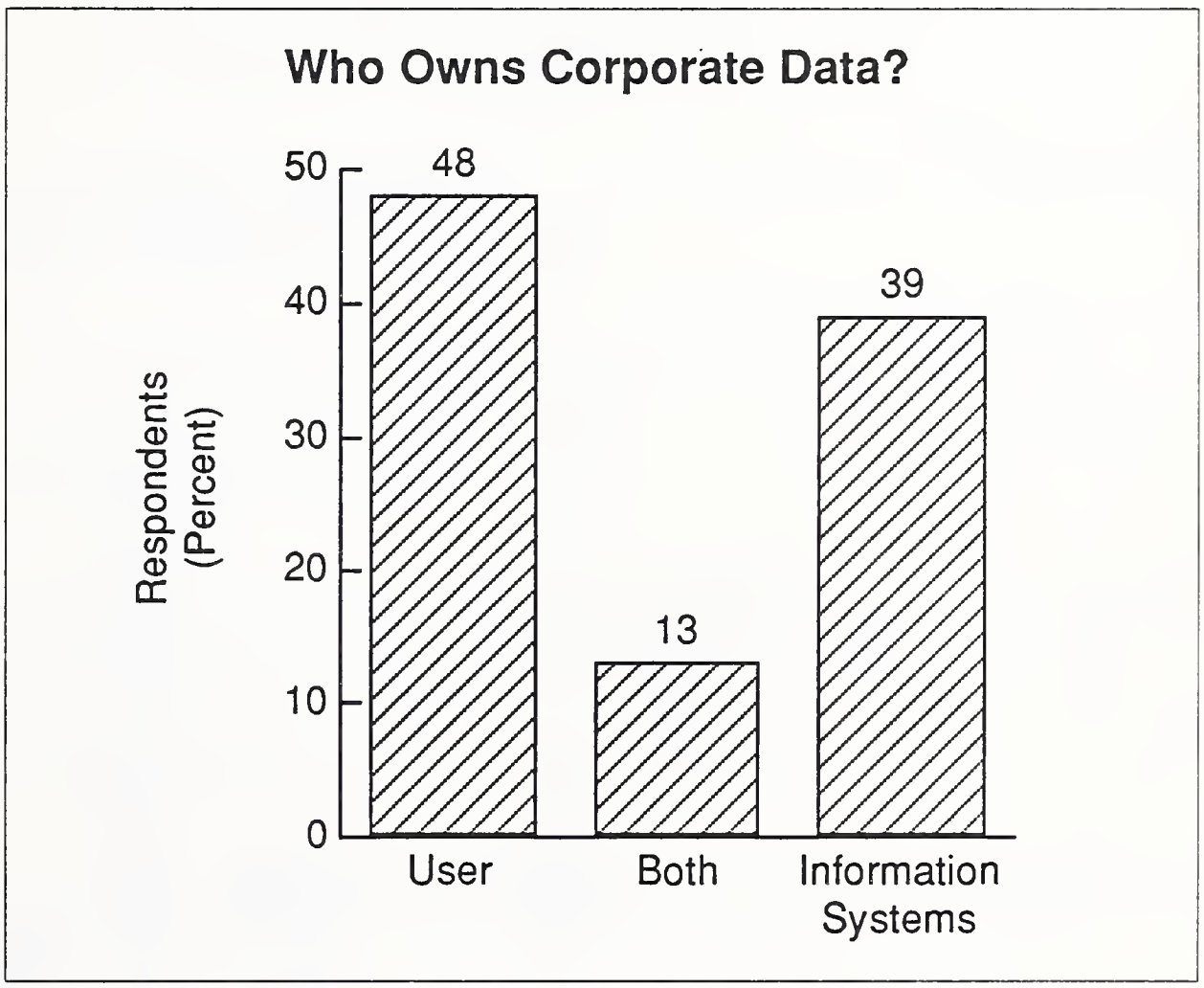
Key Issues—Strategy and Direction

- Managing distributed data
- Ownership—user versus IS responsibilities
- Managing growth and technology
- Planning for new technology
- Management support for data management process

Considering the issues related to strategy and direction, it is interesting to note that the majority of them relate directly to the relationship between information systems and users.

With the advent of relational technology, particularly when available on PC systems, data ownership has emerged as a major issue. While information systems staff have traditionally pursued policies that tended to maintain tight control over corporate data, it is clear that they can no longer claim the ownership of the data. As indicated in Exhibit II-5, it is clear that the user is increasingly the owner of data.

EXHIBIT II-5



Data ownership, access and use are key driving forces behind the need to change from data administration to data management. Only corporate-wide architecture and data management policies will forestall the fragmentation of data.

D

RDBMSs and Data Administration

In a period of two years, relational DBMSs have risen from an experimental technology to become the dominant data base technology. In the past two years, RDBMSs have become the data base of choice in the majority of firms.

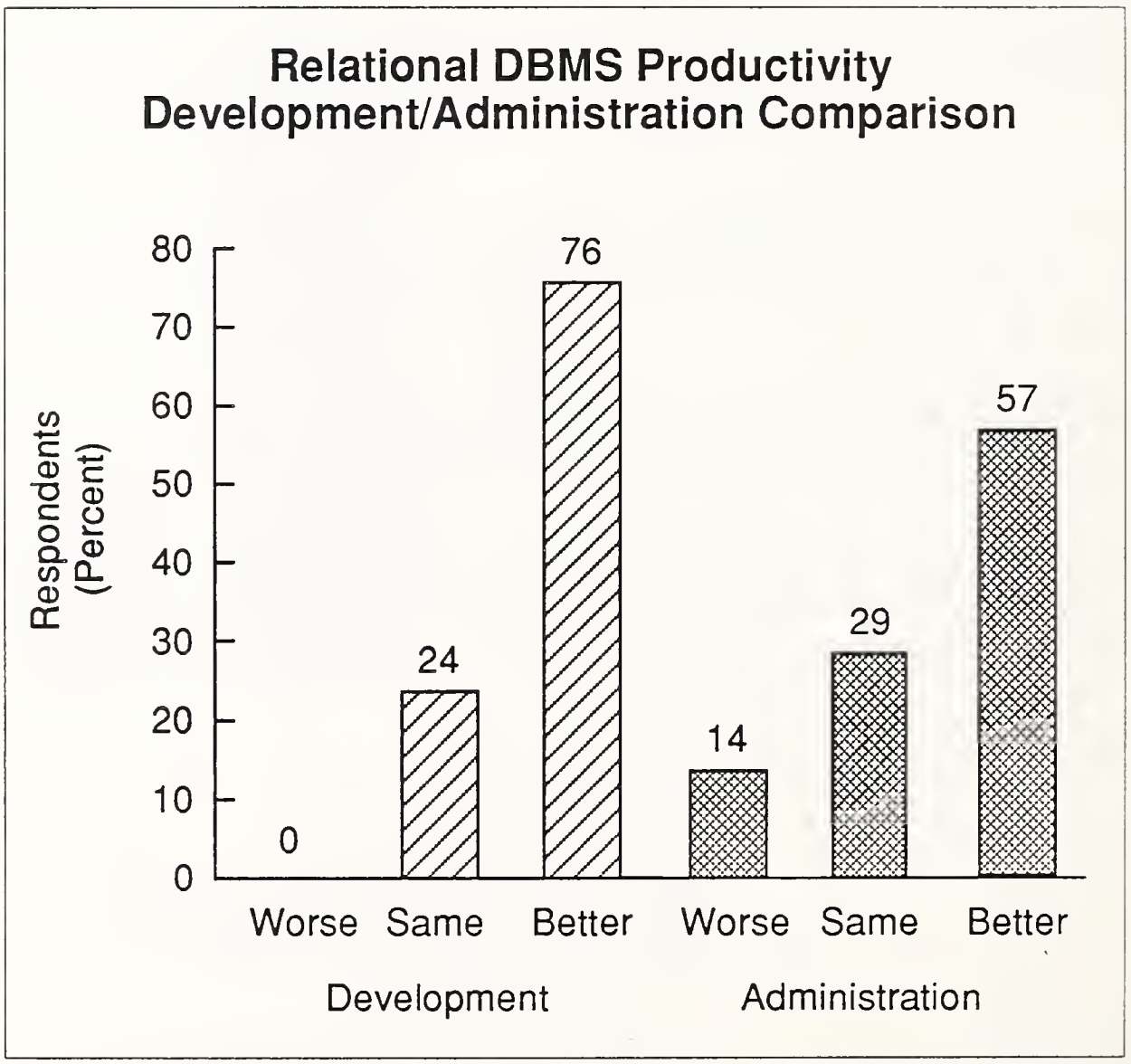
- More than half of the new data base applications are being developed using relational technology.

- Relational technology, coupled with the growth of PCs and workstations, is fostering a resurgence in the development of executive information systems.

The move to relational technology has started a revolution in data administration—a revolution that many information systems managers are not prepared to support.

As indicated in Exhibit II-6, RDBMSs have significantly improved the productivity of development staff. Of development staff interviewed, approximately 75% indicated that relational technology is resulting in development productivity improvements. Although a few indicated little change, no one indicated a worsening of productivity.

EXHIBIT II-6



Of equal importance are the responses indicating that the effect of RDBMSs on data administration is not yet as positive as on development. Only slightly more than half of the interviewees (57%) indicated that relational technology has improved the productivity of the data administration process. There are several reasons to account for such responses:

- In the short term, data administrators are faced with having to learn an additional system, compounding their already busy schedules.
- The trend toward implementing two DBMSs adds a level of complexity, as administrators have to consider how to integrate multiple (old and new) systems.
- Implementation of RDBMSs by users is forcing administrators to place increased emphasis on the needs and requirements of end users.
- With RDBMSs available on multiple processing tiers, administrators are having to give increasing consideration to distributed systems.

The extent to which productivity improvements will be achieved in the administration process is highly dependent on the rate of change from an administration to a management process.

While improvements will be noted as data administration staffs become more familiar with the technology, substantial improvements will not be seen until information systems groups develop an overall architecture, establish corporatewide (data) standards, and redefine the administration role as a management role.

E

Other Data Base Technologies

In addition to the benefits afforded by improved technology in RDBMSs, other technologies are having an impact on the evolution of the data administration process.

1. Distributed Data Base Management Systems

The concept of distributing data bases is not new. Numerous attempts have been made to implement distributed technology during the last fifteen years.

However, the advent of multitier relational data base products, coupled with the increased use of powerful minicomputer and PC systems, is fostering a resurgence in efforts to distribute data throughout an organization.

Distributed capabilities offer significant control benefits and also help the end user gain access to needed information. It is reasonable to assume that distributed capabilities will be put into use more frequently in the future.

2. CASE

Introduction of computer-assisted systems engineering (CASE) technology, along with the "repository", a major step beyond the data dictionary, will significantly enhance the ability to develop complex systems.

The success of CASE will depend heavily on how the data administration function accepts the capabilities and disciplines of CASE. CASE and the repository concept will not be successful if they do not encompass the data management process.

F

Future Environment

The process of managing corporate data in the 1990s will be considerably different than in the 1980s.

Robust relational technology, a full three-tier computing environment, and the increasing involvement of the end user in the decision and execution processes indicate that the data management process must change to become more attuned to planning, direction, and coordination (Exhibit II-7).

EXHIBIT II-7

Data Management Process for the 1990s

Tier	Description	Functions
1	Infrastructure	Information architecture Data network modeling Technology selection
2	Definition	Definition standards Definition review Coding schemes
3	Execution	Data element definition Design execution DBMS installation and support

The individual who directs the administration process in the 1990s must be more a manager (generalist) than an administrator (technician). The most important attributes will be a business (generalist) orientation and a focus on strategically-oriented architecture. Focus on design, detailed definitions, and execution will become the province of system users.

G

Recommendations

INPUT's overall recommendation is to begin immediately to reassess the data administration function and begin the migration from administration to management.

Exhibit II-8 provides a series of short- and long-term recommendations. The short-term recommendations are oriented toward understanding the existing environment and establishing a base from which to grow.

The long-term recommendations are oriented toward establishing an environment in which the user is a valued participant in the development and execution process.

EXHIBIT II-8

Data Management Objectives

Short-Term

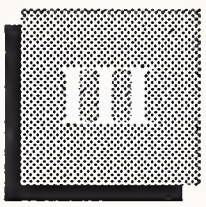
- Conduct data administration audit
- Define and gain acceptance for a charter
- Select DBMSs for PCs
- Reemphasize data dictionary
- Introduce data management generalist

Long-Term

- Broaden data management to all network levels
- Train the user
- Develop corporate data architecture
- Market data management process
- Shift administrative tasks to end user
- Explore impacting technologies (i.e., CASE)
- Experiment with DDBMSs

INPUT believes that the environment of the 1990s will necessitate strong user support. Multitiered data bases, distributed processing, expanded networking and greater user access will require direct user involvement.

Data Administration: Current Environment



Data Administration: Current Environment

A Background

With the advent of data base systems in the early 1970s, information systems were provided with significantly expanded tools for data storage and applications development. However, with the new technology, there was a new requirement for an individual who had special technical skills—the data administrator.

Prior to DBMS technology, the process of defining data files and structures was accomplished by the systems analyst developing an application. File structures and naming conventions were unique to a given application. With the new DBMSs, a specialist was necessary. The specialist had responsibility for defining data file structure and common file definitions that were used by all analysts and applications.

The first DBSs (data base specialists) were application or systems programmers. The data management function started with "data administration" and was staffed by a technical person. The data administrator frequently had little knowledge of the application or the business purpose. Following earlier trends in data processing, the administrator usually had no knowledge of the overall business or the relationship between various business elements.

By the early 1980s, many organizations had installed a DBMS and were managing a data administration function. Throughout this period, most had a single DBMS and were managing a data administration function that was viewed primarily as a technical function.

During this period, there were efforts to implement control mechanisms such as data dictionaries. While recognition was growing that data definitions needed to be standardized for corporate use, efforts were not intense and frequently met with only modest success.

By the mid-1980s a number of events had begun to cloud the job of data administration and raised a number of higher level issues. A transformation of data management was beginning, but few recognized it or the impact it would have on the organization. Among the events were the following:

- The majority of the applications in many organizations were DBMS-based, and the interaction between these applications gave rise to the corporate data base concept. In a file management environment, the data was "owned" by the department served by a particular application. In a data base environment, data is not "owned" by any single application or individual. The data was organized for the benefit of the technology and not to support the organization. The data base was viewed as belonging to information systems, not the user.

- The information network of most organizations was becoming distributed with the use of minicomputers. To support this distribution, file transfer systems were developed and a new era of data redundancy emerged. Information systems considered the central data base to be their property and carefully guarded its use. The user began to depend on the distributed data on the minicomputer and, increasingly, the PC. The information network and its data foundation had become distributed without a process to control it.

- With organizations becoming more decentralized and the demand for distributed data increasing, there was increasing need for connectivity. With increased connectivity through LANs, WANs and national and multinational networks, the need for data access increased.

- The long-sought relational model became available and, by 1987, was in early use by many organizations. The technological basis for the data administration process began a major change that will be underway well into the 1990s.

- End users began to develop production applications and make extensive use of the corporate data base through fourth generation languages. End users found relational technology something that they could understand as they became more skilled in computer use.

While there are major changes taking place, they are happening slowly. This chapter and the next consider changes over the past year and confirm that change will continue. This change will be evolutionary, not revolutionary.

B

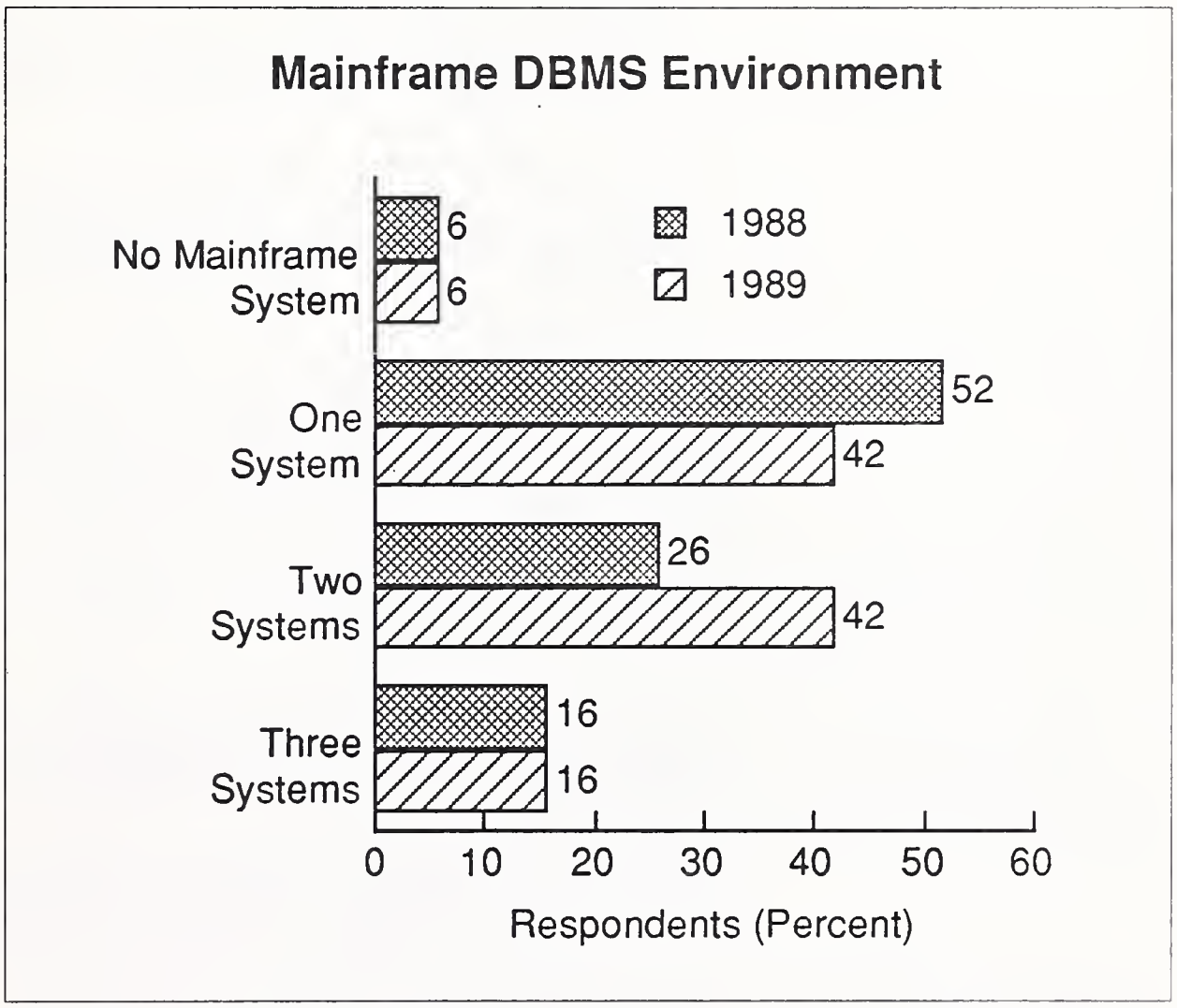
Changing Environment

Virtually every organization, without respect to size, is in the process of significant change relative to its data management systems environment. Emerging from environments with only a single DBMS, an increasing number of users now have multiple systems. In addition, with the advent of PC-based data base systems, DBMSs are now common on all levels of platforms.

1. Mainframe DBMSs

Exhibit III-1 indicates the extent to which multiple DBMSs are becoming more commonplace in many organizations.

EXHIBIT III-1



During the 1988 research program, it was noted that 42% of respondents indicated that they had more than one mainframe DBMS. During the same research, respondents indicated that 16% had three mainframe DBMSs.

Between 1988 and 1988, changes have begun to occur. In 1989, the number of respondents indicating that they have at least two mainframe systems increased from 26% to 42%. The number with three systems remained the same. The shift indicates several possibilities:

- The most likely possibility is that organizations that had new systems under consideration in 1988 (30% indicated that they did have addi-

tional systems under consideration) have now acquired and implemented the systems.

- Since the percent with three systems did not change, consideration must be given to the level of complexity in implementing and managing multiple systems.

  - The technical challenge of learning and supporting two or three mainframe DBMSs places significant strain on the resources of the data administration function.

  - In addition to the technical challenges of a new system, the task of transferring data between systems increases the complexity of the entire systems design and maintenance effort.

- A portion of the change in two mainframe systems (26% versus 42%) can be accounted for by a smaller sample size in 1989. However, considering that the respondents were the same as last year, movement to multiple systems cannot be overlooked as a definitive trend.

Since INPUT learned during the previous research that only 28% of those with more than one system had a relational DBMS as the second or third system, and that the level of interest is highest for relational systems, there are strong indications that future acquisitions will be relational.

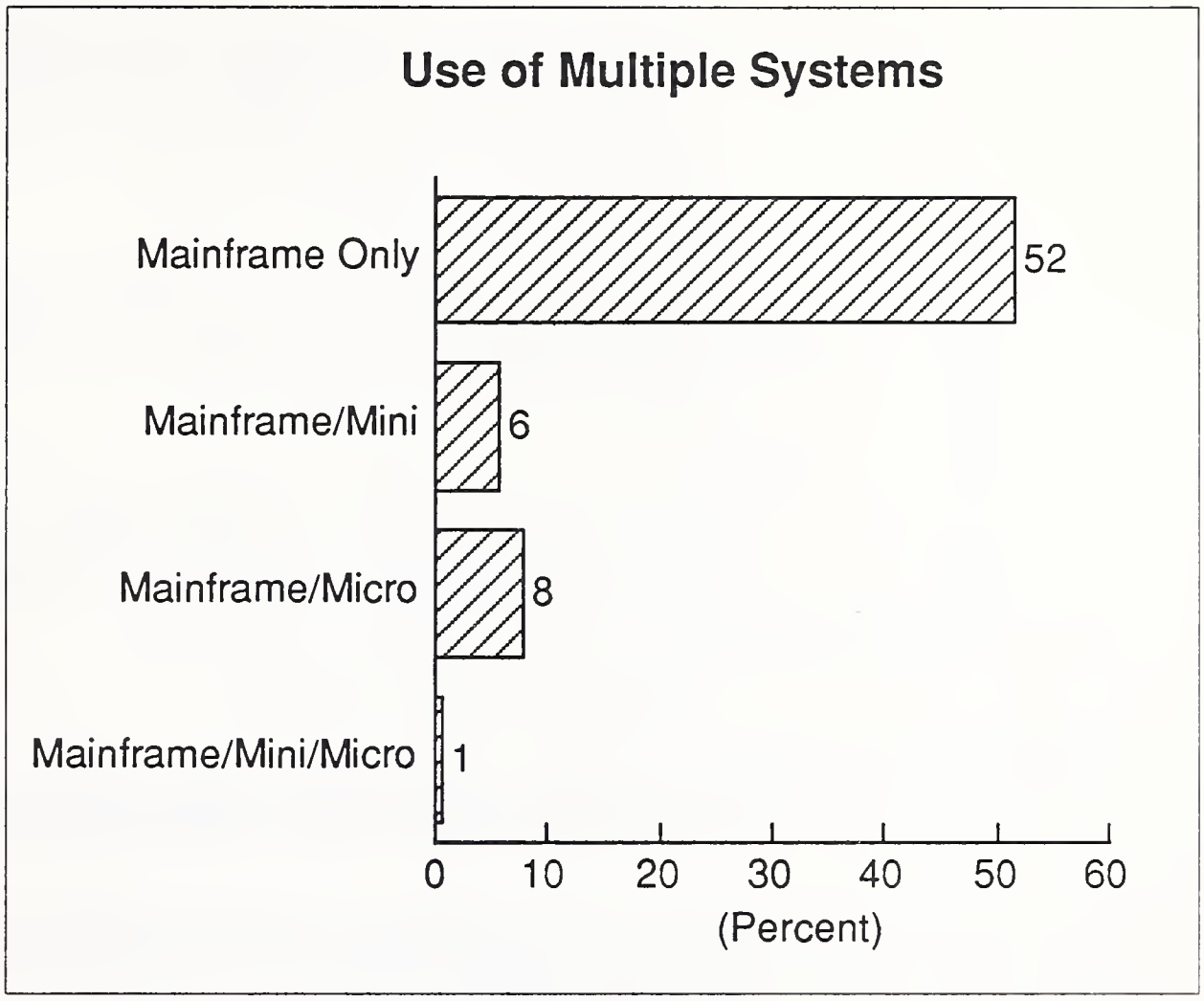
2. Multiple System Use

As part of the current year research program, INPUT considered the mix of data base systems. The purpose was to identify whether there is a predominance of systems on mainframes, minis or micros.

The mix and placement of systems can be extremely important, since the greater the number of platforms supporting DBMSs, the greater the level of complexity for the data administration staff.

As indicated in Exhibit III-2, the mainframe-only environment is currently the most prevalent. While this is not expected to change in the near term, there are several additional considerations.

EXHIBIT III-2



- The number of systems running on three platforms (mainframe, mini and micro) is quite small. INPUT believes that this reflects the level of complexity in managing a multiple platform DBMS environment.

- While the percent of DBMSs running on minis and micros is comparable, the comparatively low percent reflects the continued focus of information systems to successfully manage corporatewide mainframe data systems. (Note that this data excludes micros implemented by user departments solely for their own purposes).

- The nearly even split between mainframe/mini and mainframe/micro systems also reflects the differences in the way organizations conduct business. Many have highly autonomous divisions that have their own mini systems with their own administration. Many others rely on a single, large corporate data base for all corporate processing. These differences pose unique problems in the way data base systems are managed.

  - Distributed systems are frequently directly linked through telecommunications networks to central processing facilities. The mini data base is frequently a "distributed" data base with tight central control necessary to ensure coordination with the corporatewide data bases.

  - Micro-based data base systems are generally not directly linked to a corporate host and do not have replicated (distributed) copies of

corporate files. Microcomputer users generally extract data from corporate files for department-level analysis, or develop their own departmentwide data bases for their own purposes. Micros and workstations are also used increasingly as primary means of creating input for corporate processing.

Decisions to manage DBMSs on two or three platforms is a significant strategic decision, since the level of effort needed to ensure data accuracy, integrity and security increases geometrically as the number of connected systems increases.

INPUT believes that, while the data is not definitive, organizations will generally elect to adopt a mainframe/mini or a mainframe/micro approach as their primary strategy, and few organizations will elect to manage a three-platform environment.

3. Personal Computer DBMSs

Data base use on personal computers continues to be dominated by dBase III; the primary use is personal or departmental productivity.

However, this is beginning to change. The move to relational DBMSs is underway with the porting of their DBMSs by both Oracle and Relational Technology.

INPUT expects that implementation of RDBMSs on PCs will increase, and that this will cause increased interest in transferring data across tiers of the information network. This will create increased pressure on data administration staffs to define corporatewide formats and operating procedures. SQL and RDBMS will be the foundation for such transfers.

4. Considering New DBMSs

In INPUT's previous research, 30% of respondents indicated that they were considering new data bases. As indicated in Exhibit III-3, even with the level of implementation that has taken place over the past year, an estimated 30% of users are still considering additional data bases.

The 30% figure raises a couple of questions. With the major changes that are taking place in DBMSs why is the figure not greater? Considering the level of activity over the past year, are users looking for third and fourth systems or are they looking for improved systems?

EXHIBIT III-3

Considering New Data Bases

- 30% considering new DBMSs
- All are relational
- Most frequently mentioned:
  - DB2
  - Oracle
  - INGRES

- INPUT believes that the 30% who are considering additional systems are doing so for the purpose of expanding their data base services to additional levels of the organization.

  With this in mind, the primary focus is not the acquisition of a third or fourth system that will perform essentially the same functions, but a system (or systems) that will provide added functionality (primarily ease of use) and connectivity for the end users.

- INPUT also believes that the percentage of those looking for additional systems is actually higher than indicated by the respondents in information services. The figure could be as high as 60-70%, but the search is an informal, wait-and-see investigation.

  With current staffing, information systems have more than an ample workload. While working to meet current commitments, data management staff are looking for the right opportunity to implement the next generation of relational data base systems.

- The focus of many information service organizations is on systems that will be standard across multiple platforms and that will meet the need of vertical and horizontal connectivity and data sharing with standard definitions, data organization and corporatewide security.

C

Effectiveness of Data Administration

A key element of the 1989 research program was to assess changes that have, or have not, taken place to increase the effectiveness of data administration.

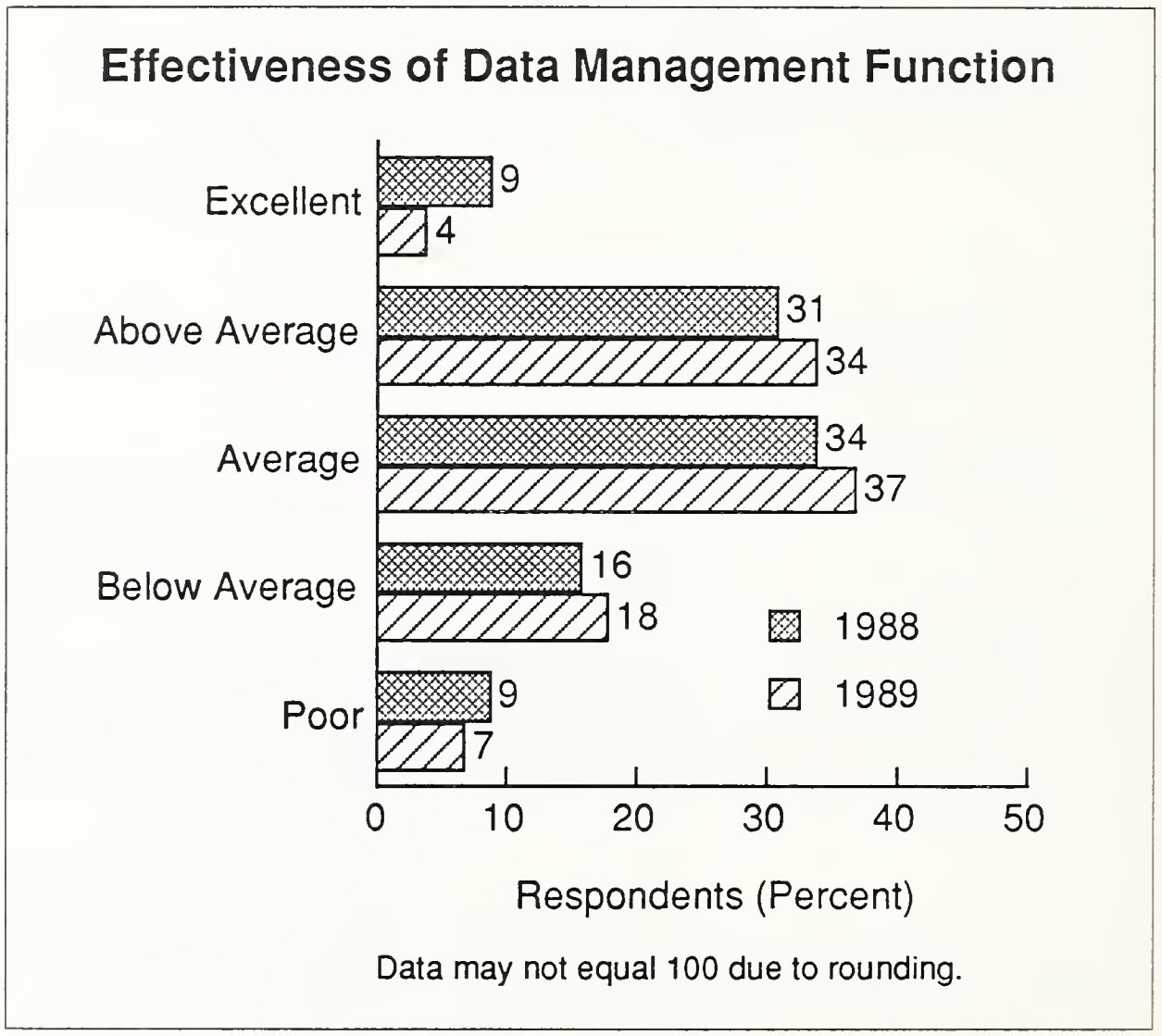
With this in mind, users were asked to rate the effectiveness of their data management functions and of their use of data dictionaries.

1. Effectiveness of Data Management Function

Having accepted that changes are taking place and that the role of the data management function will take on increasing importance, INPUT sought to identify the extent to which managers thought that the data management function was effective.

As indicated in Exhibit III-4, there has been little change between last year and this year. Fully 25% of the IS managers think that the effectiveness of the data management function is below average.

EXHIBIT III-4



While there was a slight increase in the percent of respondents that rated the performance average and above average, these changes are not believed to be significant. Essentially, there has been no change over the year.

To try to understand whether changes to this level of performance are likely, INPUT asked users to indicate what actions they were taking to improve overall effectiveness.

The predominant response indicates that there are few plans to improve overall effectiveness. Those with specific plans indicated the following most frequently:

- Selling and educating corporate management, users and applications developers
- Greater involvement in the systems' life cycle process
- Increase staff
- Don't see significant improvement until technology such as CASE and the repository become readily available

In all, the predominant theme is that more time is necessary for the data management function to be accepted as a necessity within an organization. Acceptance is needed in two key areas:

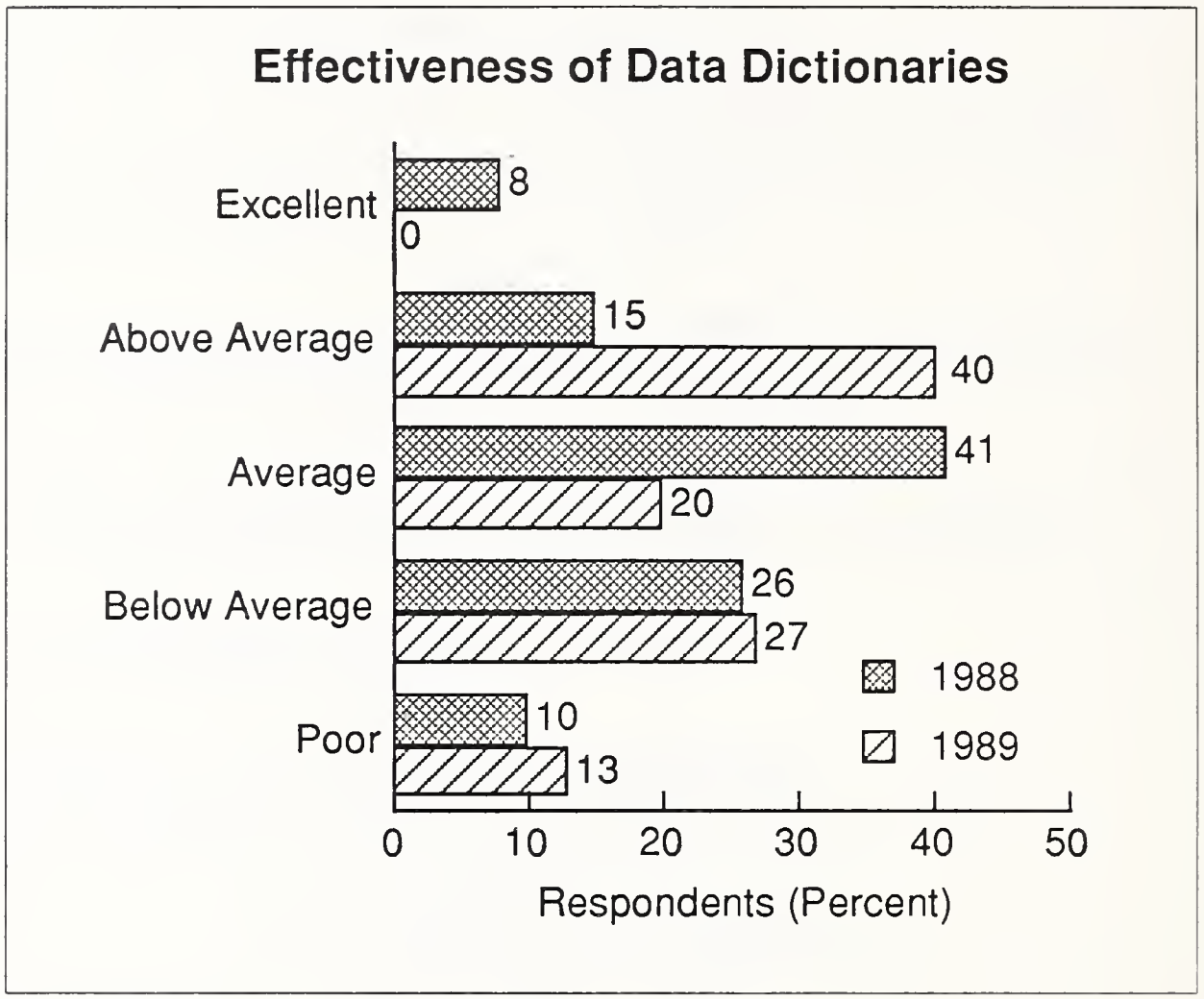
- Acceptance is needed by system developers who see themselves as equally capable of defining data requirements and managing data flow.
- Corporate management needs to understand that this is not an additional overhead expense with limited value. In addition, management is influenced by user groups who see data management as another attempt to control their environment.

These findings indicate the lack of a strong base from which to expand into new responsibilities and new DBMS technologies.

In conjunction with questions related to overall effectiveness, users were asked to rate the effectiveness of their data dictionaries.

As indicated in Exhibit III-5, the extent to which information systems managers consider the use of data dictionaries to be effective has shifted over the past year.

EXHIBIT III-5



During the current research process, there was a definitive shift in users' perception of the effectiveness of data dictionaries.

- While the percentage that consider their data dictionary effectiveness to be below average remained essentially the same (40% versus 36%), there was a decline in the percent that consider their effectiveness to be excellent (8% versus 0%).

- Over the same period, there was a shift in the number that consider their performance to be above average and those that rate performance as average. The average category shifted more toward above average.

- INPUT believes that there are two key reasons for these shifts.

  - First, users are gaining greater experience with relational systems, placing greater reliance on data dictionaries, and generally becoming more proficient in the application of DBMS technology.

  - Second, users are becoming more realistic in assessing their overall performance. This accounts for the decline in the number of excellent responses.

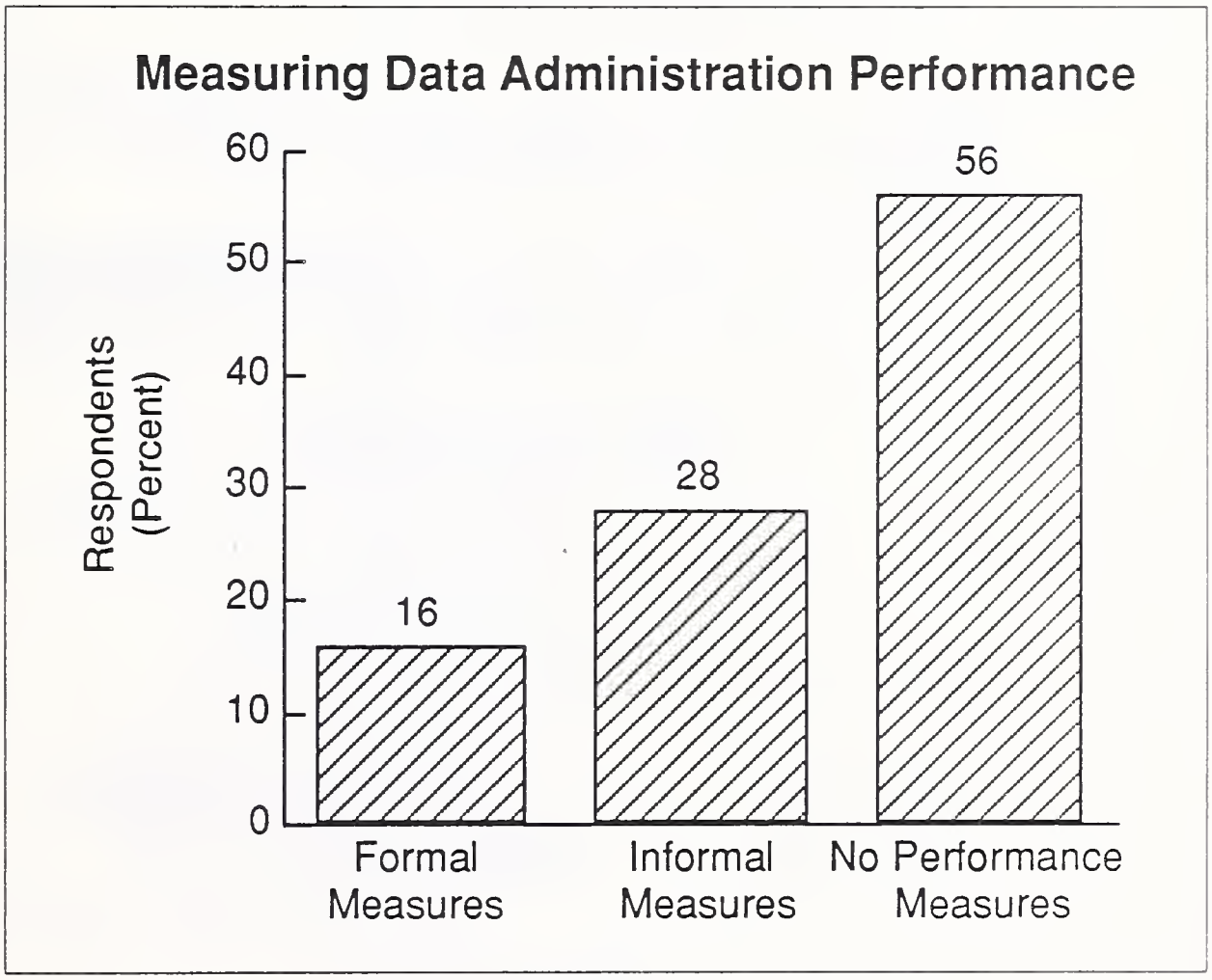
In general, users believe that their use of data dictionaries is increasing, but that a considerable amount of work remains to be done to make them truly effective.

2. Measuring Performance

The ability to measure performance is increasingly important in all organizations. While establishing performance measures can be difficult for a new technology, INPUT was surprised to find that there are few measures of the data administration function.

As indicated in Exhibit III-6, only a minimal number of organizations have formalized methods to measure the effectiveness of the data administration function. More than 50% indicated that they are not able to measure the performance of the data administration function.

EXHIBIT III-6



When questioned about their ability to measure the performance of data administration, users generally indicated the following:

- They have generally not been able to identify adequate measures.
- Response time and user satisfaction, as indicated by verbal interaction, were the key measures.
- Many were measured on the basis of their ability to interact successfully with the corporate applications development group(s).
- Formal processes centered around periodic, structured user surveys.

INPUT believes that measures of performance will become increasingly important if data administration is to successfully meet the needs presented by the data base environment of the 1990s.

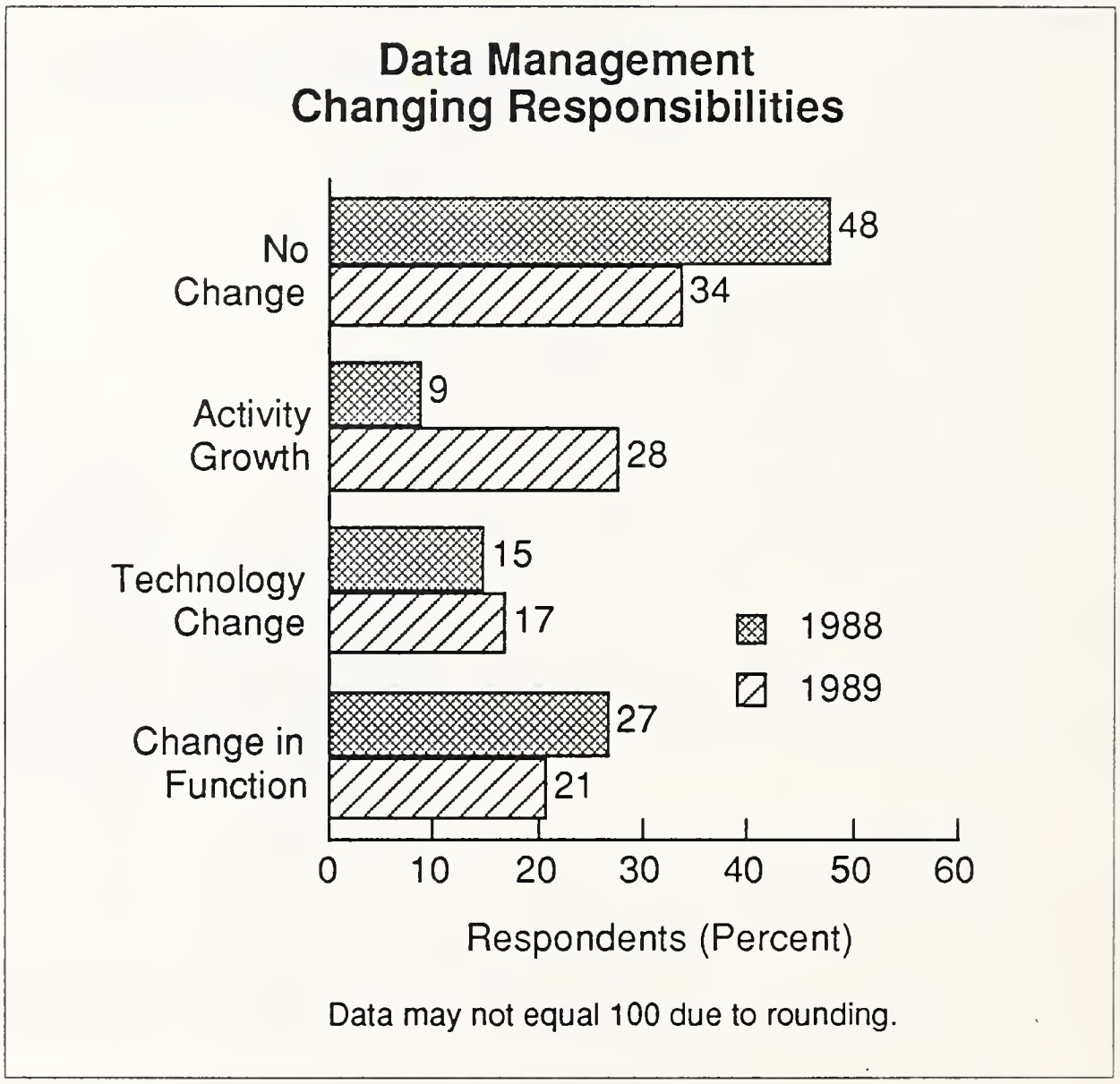
D

Changing Role

To determine the extent to which the role of data administration is changing, INPUT asked managers to indicate how their roles had changed over the previous year.

Figure III-7 reflects a somewhat different environment than existed last year, generally reflecting increased activity.

EXHIBIT III-7



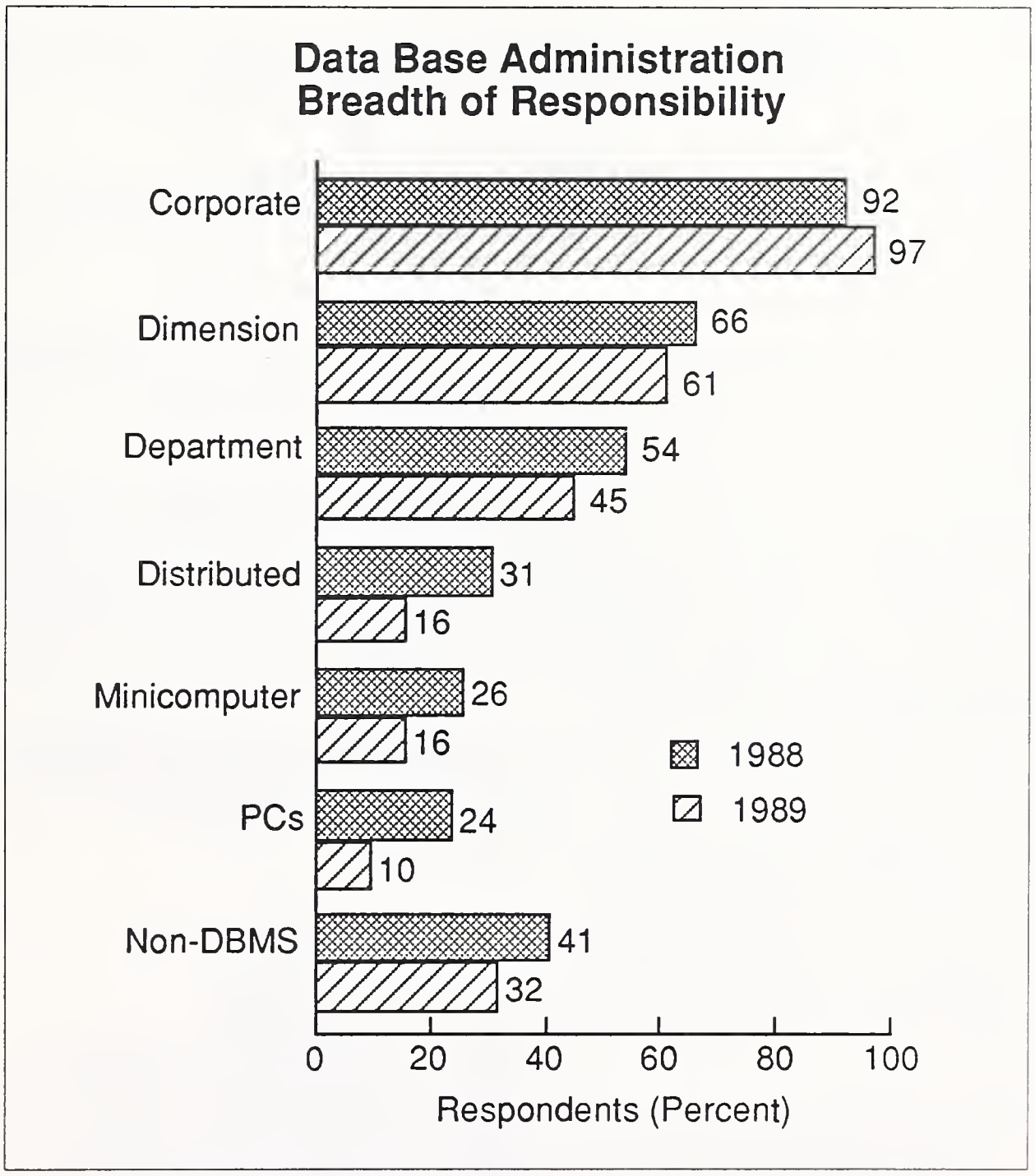
- As indicated by the reduction in the no change category, and an increase in the activity growth category, the year showed a considerable increase in activity.

- There was some change in the areas of technology and function, and an increased level of effort to implement new systems that were previously under consideration.

- INPUT believes that the limited change in technology indicates that, while work still remains, many users have become more settled in their approach to developing and implementing DBMSs.

With a couple of exceptions, a year-to-year comparison of the breadth of responsibility of data administration showed little change. Though not dramatic, there are changes indicated in the degree of responsibility for administering data on distributed, minicomputer, and PC systems (Exhibit III-8).

EXHIBIT III-8



At the corporate, division and department level, the role of data administration showed almost no change and is not expected to change. However, a reduction in responsibility for distributed, minicomputer and PC systems is indicated.

- Survey results indicate that there is a distinction being made between organization responsibility and platform responsibility. While information systems will remain responsible for administering data across organization lines, they will have less responsibility for administering data on individual systems.

- The decrease in IS responsibility is a logical progression, as users assume an increased role in defining their own needs and gain a better understanding of relational systems.

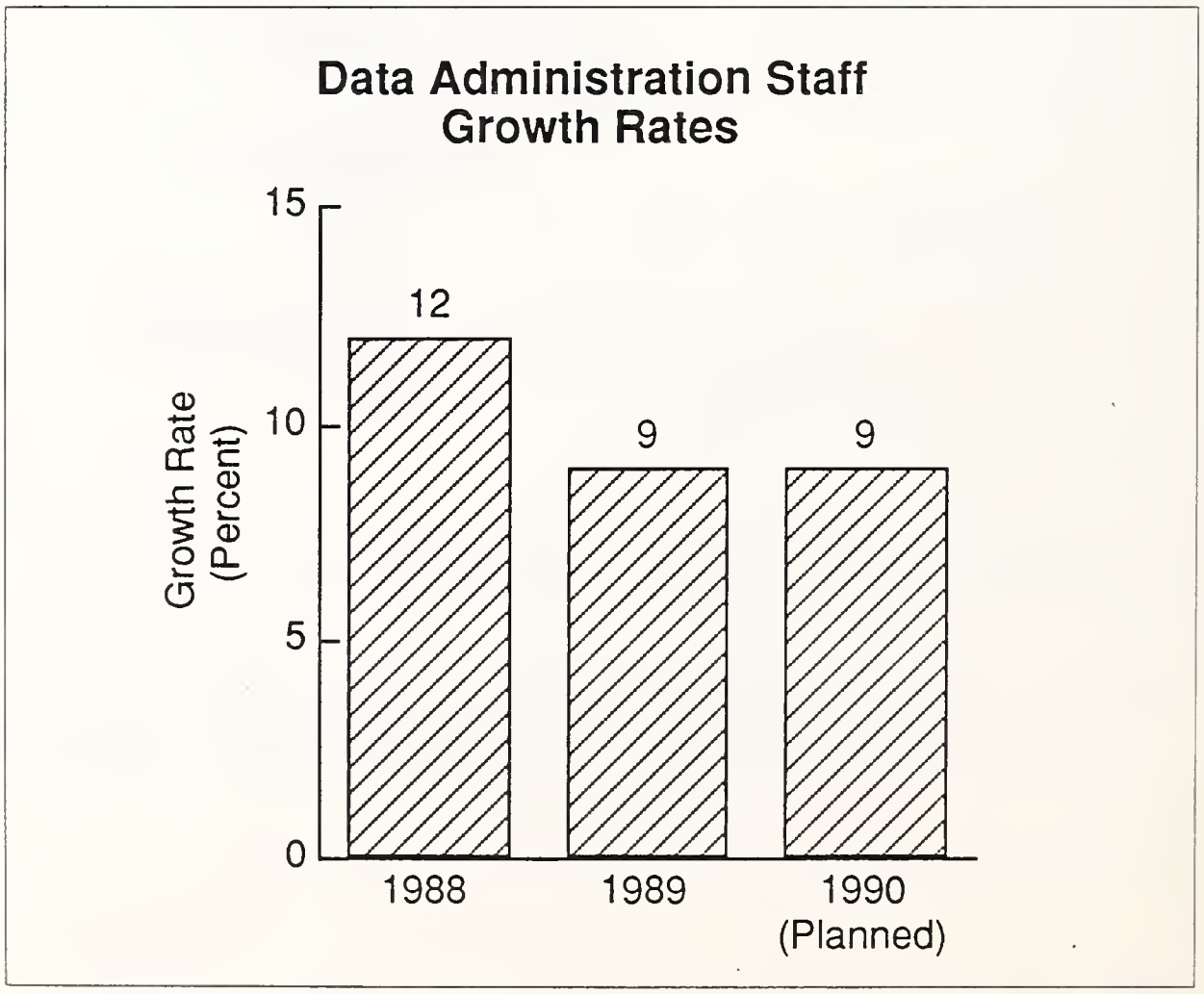
  The changes indicated are not without potential for significant problems, which increases the importance of dictionary use and corporate-wide data administration strategies.

- Assumption of greater responsibility by users could result in a wide variety of unshareable data bases and management data that reflects different results.

- With users assuming a greater role, the data dictionary becomes increasingly important as the single point of coordination for corporate-wide data.

As indicated in Exhibit III-9, there has been little change in the growth rate of staff to manage corporatewide data. In 1988, users indicated that their staff grew an estimated 12%. They projected growth of approximately 9% for 1989. Users indicate that the projected rate for 1989 has been realized, and that similar growth is expected for 1990.

EXHIBIT III-9



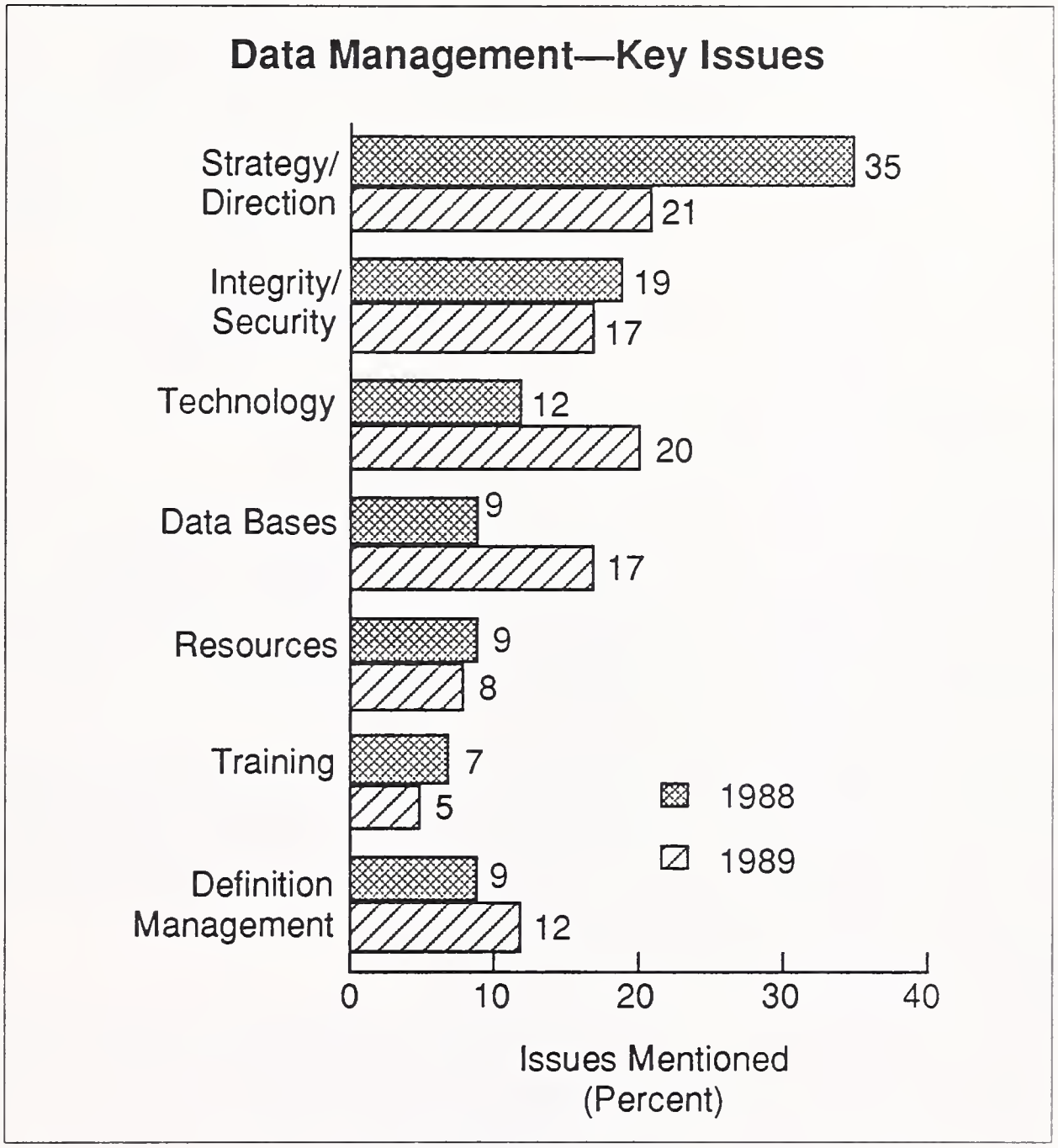
E

Key Issues

Exhibit III-10 identifies issues reported by managers of corporate data.

Overall, key issues with managers remain the same as last year. Strategy and direction remain at the top of the list, followed by data integrity/ security and technology. However, there are several changes that should be noted.

EXHIBIT III-10



1. Strategy and Direction

Strategy and direction remain the top concern of managers of corporate data. This reflects an ongoing concern over how DBMSs can be most effectively utilized in the organization.

However, a reduction in the frequency of mention indicates that managers are beginning to settle on strategic directions. The settling results from the level of effort expended over the past few years. It also reflects resolution of issues about multiplatform systems that have been major concerns over the years. Relational systems that are increasingly portable and easier to understand have helped remove obstacles to expansion.

Overall, managers indicated several areas of concern regarding strategy and direction. Frequently mentioned areas of concern are identified in Exhibit III-11. It's interesting to note that these have not changed since last year.

EXHIBIT III-11

Key Issues—Strategy and Direction

- Managing distributed data
- Ownership—user versus IS responsibilities
- Managing growth and technology
- Planning for new technology
- Management support for data management process

2. Integrity and Security

Integrity and security of data remains a key concern to most managers. This concern is not expected to subside and may well increase. The fundamental elements of the integrity and security issue are found in the growth of distributed and personal computing and the flow of data across multitiered computing networks.

As data managers make progress in defining corporatewide data structures, and users progress in implementing departmentwide data bases, data sharing will become increasingly important. With the sharing will come increased concern over integrity and security of data.

3. Technology

Over the last year, there has been increased interest in the direction of technology. The availability of portable, relational technology has resolved some previous issues. However, users looking to the future are now concerned about what technology will become available to address control of distributed systems and corporatewide connectivity.

4. Data Access

With the implementation of relational DBMSs on minis and micros, an increasing number of users are interested in having greater access to enterprisewide data. Data managers are now faced with demands for increased access to a wider variety of data.

Although several areas of change that have been indicated, most managers of data administration do not see significant expansion of their responsibilities. They are struggling to keep up with current short-term demands.

Impacts of Relational Data Base Technology

IV

Impacts of Relational Data Base Technology

This chapter considers a number of impacts on data administration from implementation of relational data base technology. Included are considerations of gains (or losses) in productivity, the need for training, the use of modeling, and whether relational systems are having an effect on the use of data dictionaries.

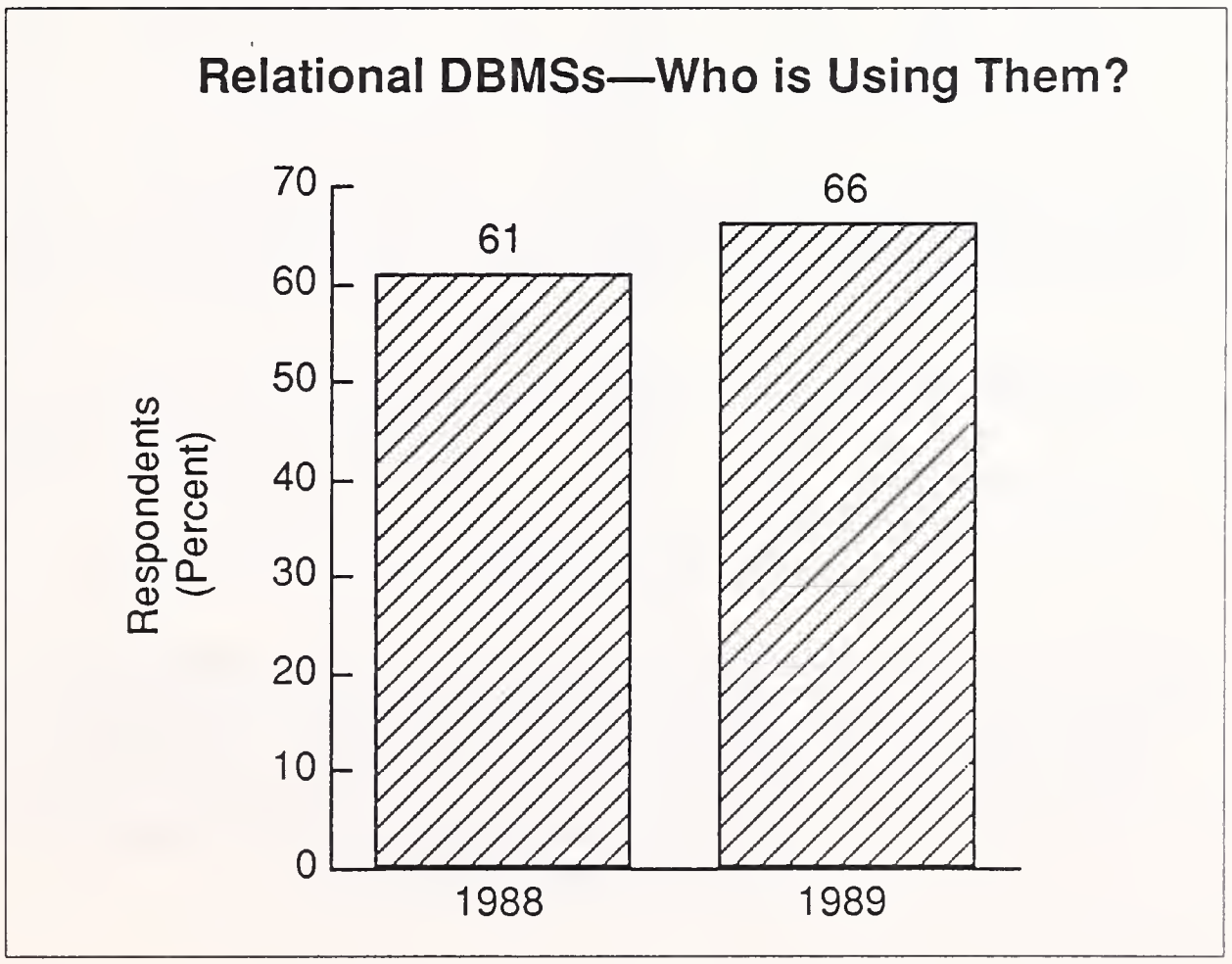
A

Status of Relational DBMS Use

1. Trend Toward Increasing Use of RDBMSs

As indicated in Exhibit IV-1, the use of RDBMSs is progressing steadily. In this year's research, over 65% of respondents indicated that they use RDBMSs, up from approximately 60% last year.

EXHIBIT IV-1



While the change in percent is not major, the level indicates that the move to relational systems has taken hold and that use will continue to grow.

The level of use also indicated that information systems staff are seeing benefits of relational DBMSs over other types of systems and are prepared to continue to develop systems using the new technology.

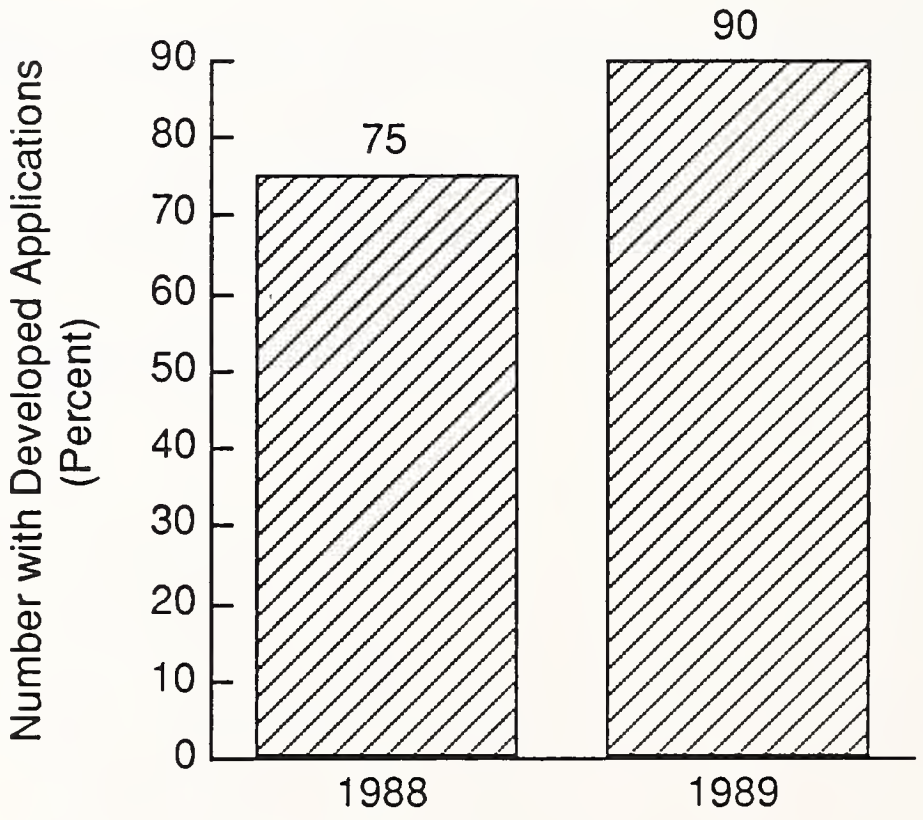
2. How It Is Being Used

Consideration of the number of users that have developed relational DBMS applications on their mainframes (Exhibit IV-2) further substantiates that use of relational systems is progressing rapidly.

In last year's research, approximately 75% of users indicated that they had developed mainframe applications. This year the percentage had increased to nearly 90%.

EXHIBIT IV-2

Relational DBMS Applications
Mainframe Development Progressing



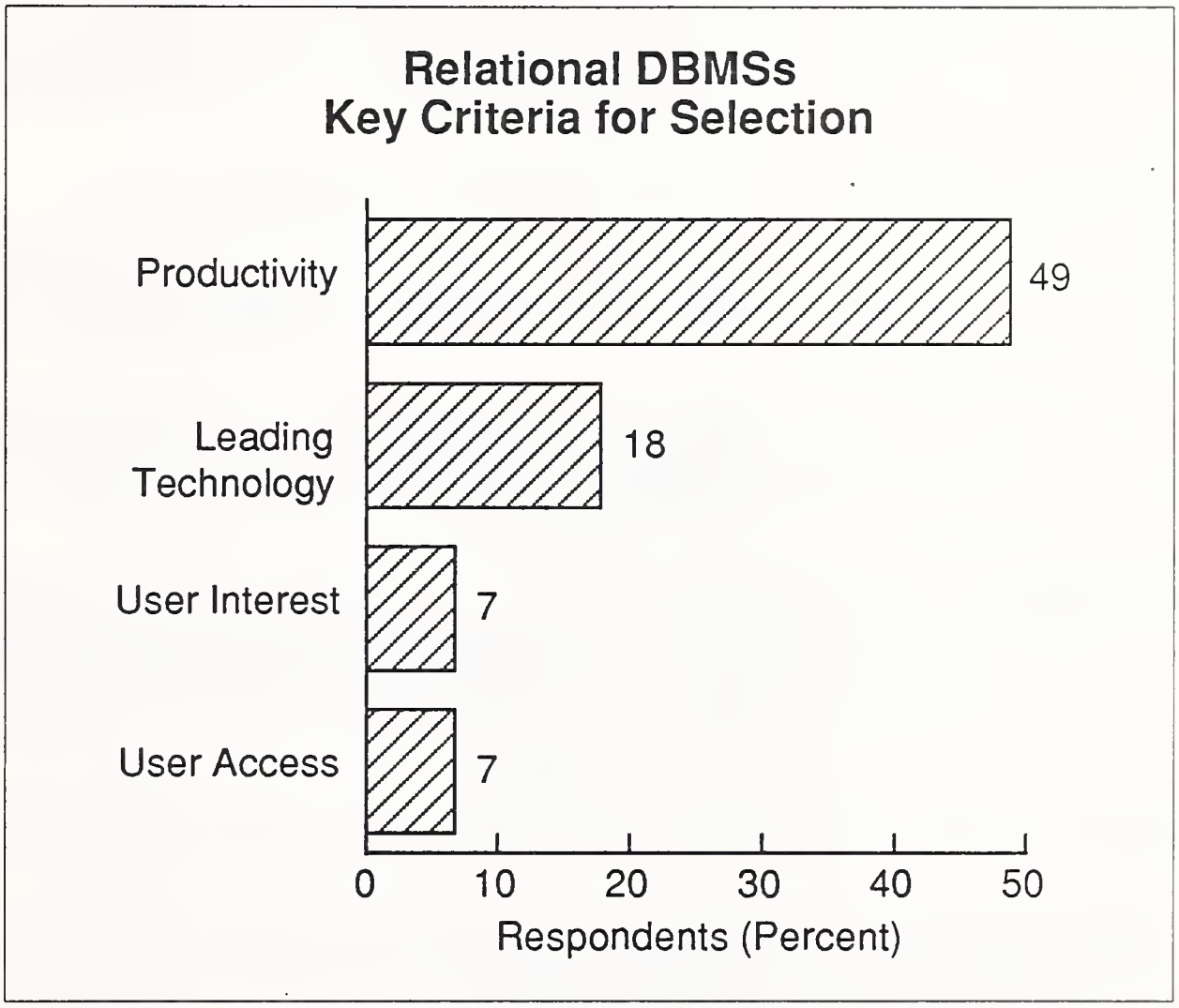
As compared to other platforms, the mainframe is currently the most often used platform for RDBMS applications. This is not surprising, since the survey was conducted with central information service executives. However, note should be made that an estimated 15% of respondents indicated that PC-based applications had been developed. This trend is expected to continue.

3. Criteria for Selecting an RDBMS

As part of the research, INPUT sought to identify key criteria for selecting an RDBMS over a conventional DBMS.

Though, in one respect the results are not exceptionally surprising, they do provide an indication of future trends, as Exhibit IV-3 illustrates.

EXHIBIT IV-3



- Nearly 50% of the respondents indicated that the key criteria for selecting an RDBMS was that they believed that it would increase productivity.
  - As a part of productivity considerations, flexibility was the most frequently cited reason.
  - The ability to make changes more easily was the second most frequently cited reason for increasing productivity.
- Though productivity gains were the most frequent reasons for selecting an RDBMS, being on the leading edge of technology was also considered important. In a number of cases, users indicated that they were influenced by IBM marketing which indicated that relational technology would be IBM's focus of the future.

- Although the percent of responses were not exceptionally high, it is interesting to note that (taken collectively) user interest and user access rated nearly as high as the fact that RDBMS is a leading technology. It is equally interesting that of the four major categories of criteria, user considerations were among the strong influencers.

- INPUT believes that the trend toward increased user influence will grow. As relational systems become more accepted in the user environment, decisions regarding system structure and operation will be increasingly influenced by how users need to conduct their business.

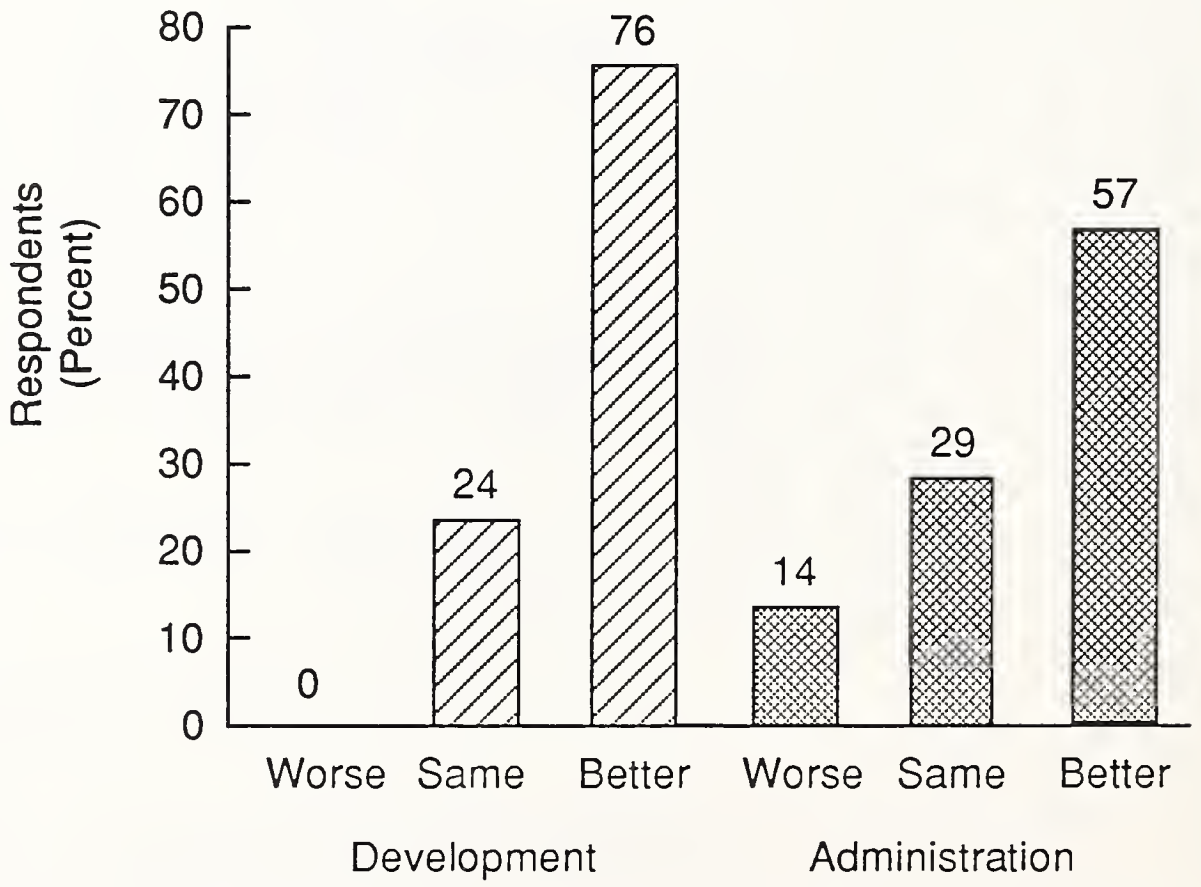
4. RDBMSs Improving Productivity

With improved productivity a key reason for selecting a relational system, INPUT sought to identify the extent to which improvements have been realized in systems development and in data administration.

As indicated in Exhibit IV-4, productivity improvements have been realized in both development and administration. However, some differences can be noted.

EXHIBIT IV-4

Relational DBMS Productivity Development/Administration Comparison



- Respondents indicated that development productivity improvements have been realized in nearly 76% of the cases. None of the respondents indicated that the RDBMSs had resulted in a worsening of productivity.

- For data administration, the results were somewhat different. While the percent reporting the same level of productivity was comparable to development, 14% indicated that the RDBMS has resulted in a loss of productivity. However, only 57% indicated that productivity of the data administration function had improved. Responses indicated several reasons.

  - The RDBMS has added workload to the data administration function, resulting in reduced concentration of effort.

  - Since RDBMSs are still somewhat new, there is a learning curve to overcome.

  - RDBMS administrators are having to spend more time assisting users. Given the technical orientation of many DBAs, user interface is sometimes considered unproductive.

The extent to which data administrators believe that they are realizing productivity improvements is expected to shift to the better category as they become more familiar with the technology and more oriented to meeting the needs of the end user.

B

Impacts of RDBMSs on Data Administration

1. Dual DBMS Environment

While the increase in the number of dual data base environments would suggest that this will continue, few believe that administering such an environment will be easy. Recent efforts by vendors such as IBM to move people to relational systems suggests that this trend should change over time.

Substantiation for this trend can be found in the percentage of information systems organizations that are moving quickly to do the majority, if not all, of their development on relational DBMSs. Knowing that much of this is IBM mainframe-based indicates that the dual environment is not well-accepted and will speed the move to a full RDBMS environment.

2. Training

The perception of the level of training required for successful implementation of RDBMSs varies considerably. The responses about training levels generally fell into two categories:

- The first was those who recognize that relational technology is considerably different from traditional DBMS technology, and that extensive training is necessary.

- The second was those who suggest that the introduction of relational technology has had little impact on the level of training in their organization. They indicated that they will obtain the needed information "on the job" and that training will take place "over time".

Though the results are inconclusive, a review of respondent comments indicates the following:

- Those who believe that training is a significant consideration are moving ahead aggressively, have implemented or are implementing a data dictionary, and are supportive of providing user support.

- Those who believe that training can be approached more slowly are implementing relational systems for the benefit of the corporate information systems staff, and are relying on systems analysts to learn about relational systems as part of the systems development process. Many of these do not have data dictionaries.

3. Data Modeling

The extent to which users have embraced data modeling varies. A number of users indicate that while data modeling has raised new issues within their organization, it has proven to be highly useful. They indicate that modeling has improved their overall understanding of their organization.

An equal number of users indicate that data modeling has not resulted in substantive change or that they have not placed any significant emphasis on modeling. However, note should be made that the majority of those that have not experienced appreciable benefit from data modeling are not using data dictionaries.

4. Data Administration/Data Dictionary

As indicated in Exhibit IV-4, many users indicate that relational systems have improved their data administration process.

Of key interest is that users indicate that implementation of a relational system has induced them to either begin to use a data dictionary or has increased their attention to improving the one they have.

While the majority reported that a relational system increased the number of tasks involved in administering the system, they were generally positive about the benefits that would result.

5. New Support Tools

Most data administration has survived without a fully functioning data dictionary or other administrative tools. Some administrators are using data modeling techniques, but to support the traditional DBMS environment.

The dynamic nature of the RDBMS environment, the use of multiple DBMSs, and the increased linkage of data networks and end-user computing technologies will all require tighter administration. In addition, many of the newer RDBMSs (Oracle, Sybase, Ingres) include an integrated data dictionary.

As distinct from last year, research this year indicates that users are beginning to make greater use of the support tools. This trend is not only expected to continue, but will become a primary focus of data administration activities in the years to come.

C

RDBMSs and the End User

In its current research, INPUT notes that the primary use of RDBMSs among users continues to be for management reporting, as shown in Exhibit IV-5. Secondary uses are in development, marketing and sales, and accounting.

EXHIBIT IV-5

How End Users Are Using RDBMSs
Top Four Applications

- Management reporting
- Development
- Marketing/sales
- Accounting

1. Mainframe Use

Use of the mainframe RDBMS by end users is usually tied to management reporting activities. Users are using 4GL technology to access and extract data, analyze corporate data and prepare management reports. Data is frequently moved to department minis and micros for analytical purposes.

2. PC Use

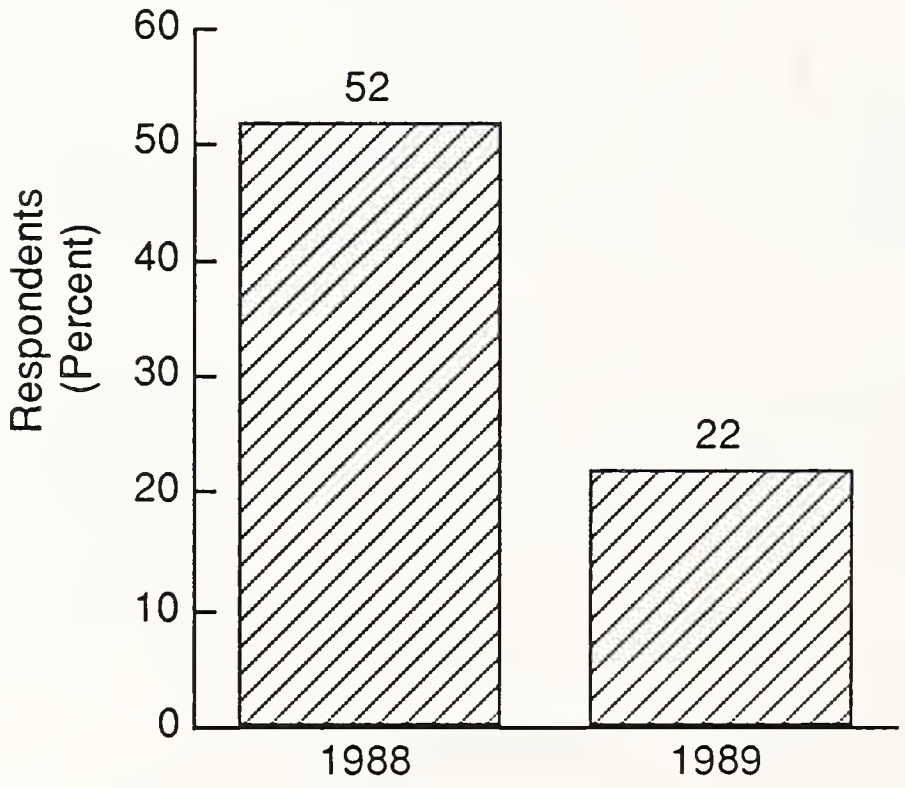
In previous research, respondents indicated to INPUT that more than 50% of users had developed PC-based relational DBMS applications. However, INPUT believes that this figure is excessively high, due to a general lack of understanding on the part of users of exactly what a relational system is.

Many users think that dBase or similar PC DBMSs are relational, but they are not. Examination of the applications that end users have developed with such tools indicate that the file structures are standard sequential files.

To gain a better understanding of the end user, PC application environment, INPUT asked whether end user applications had been developed using relational data base systems. Exhibit IV-6 indicates that the total is closer to 20%, as compared to the previously reported figure of 52%.

EXHIBIT IV-6

Relational DBMS Applications
PC Application Use



This difference is not surprising. While the question posed was similar to last year's, the respondent was encouraged to think of development using specific RDBMSs rather than development using traditional (dBase) systems. This contributed to the lowering of the figure.

In addition, INPUT believes that as true relational systems become more accepted by the end user, the user is becoming increasingly knowledgeable and able to provide more accurate answers. The lower number is expected to grow substantially over the next year.

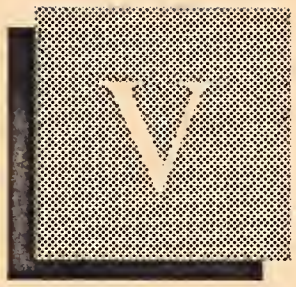
3. Implications

There are numerous implications for information systems of the growing use of relational technology by the end user. Several of these implications could have significant impact on the data management function.

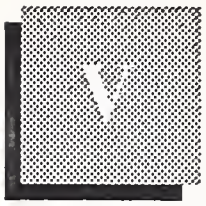
- From recent activity, it is increasingly apparent that users can understand relational technology. With more simplified technology, users will be more inclined to use relational DBMSs than previous technology. This will create increased pressure on data administration to provide corporatewide data organization. A little (uncoordinated) knowledge can be dangerous!

- As RDBMSs are increasingly accepted by end users, there is a risk that users will become more knowledgeable than information systems staff about the capability and potential benefit. Unless the data administration staff become knowledgeable, they will be in a position of losing their ability to control the overall strategy and direction of the company.

- Information systems staff must take on a more supportive role in working with end users. As end-user knowledge grows, they must be able to fit within an overall strategy or development will be fragmented.



Impact of Other Technologies



Impact of Other Technologies

While it may not be readily apparent, growth of DBMSs (specifically RDBMSs) and the data administration function is being impacted by growth of technology in several areas.

The data administration function will not only have to deal increasingly with multiple data bases, but must also be prepared to successfully handle distributed data, manage data at the workstation, support a revolution in application development processes, and provide significantly increased levels of access to an increasing number of users.

This chapter discusses some of the changes in tools and technologies that will affect the data administration function.

A

Distributed Data Base Technology

Simply defined, distributed data base technology permits the replication of all or part of a data base in one or more systems located at a distance from a central processing facility.

1. Use of Distributed DBMSs

From INPUT's report, *Distributed Data Base Management—An Early Look*, Exhibit V-1 provides a graphic example of an application that uses DDBMS technology and includes replicated (redundant) data. The distributed portion of the DBMS provides the technology to:

- Track data relationships across platforms on which the data base resides
- Permit and maintain redundant data where required for processing efficiency
- Assure that a transaction which updates data stored on more than one platform does so on all affected platforms (or no platforms), thus maintaining data integrity

EXHIBIT V-1

Distributed Data Base Management
A Replicated Data Example

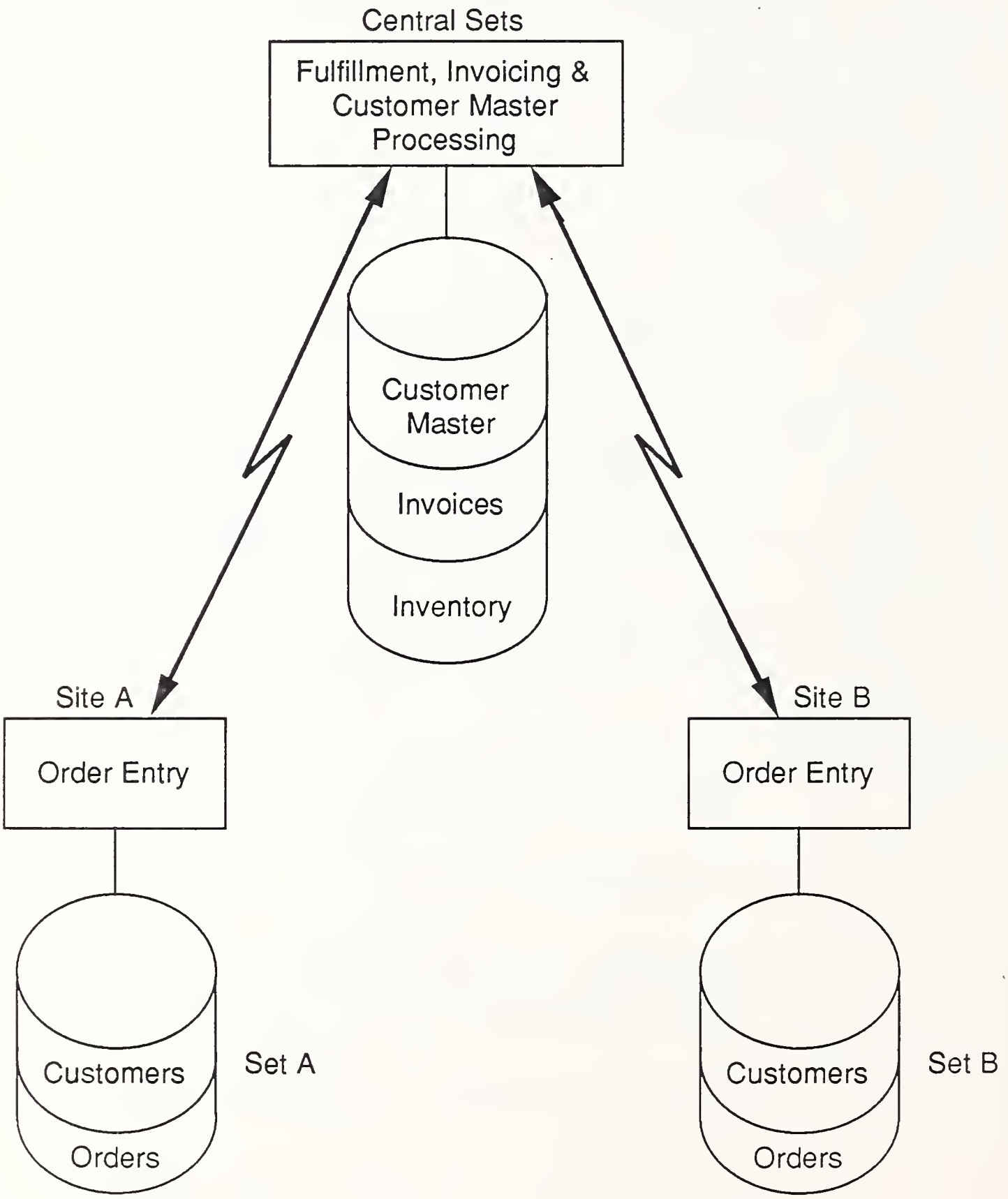
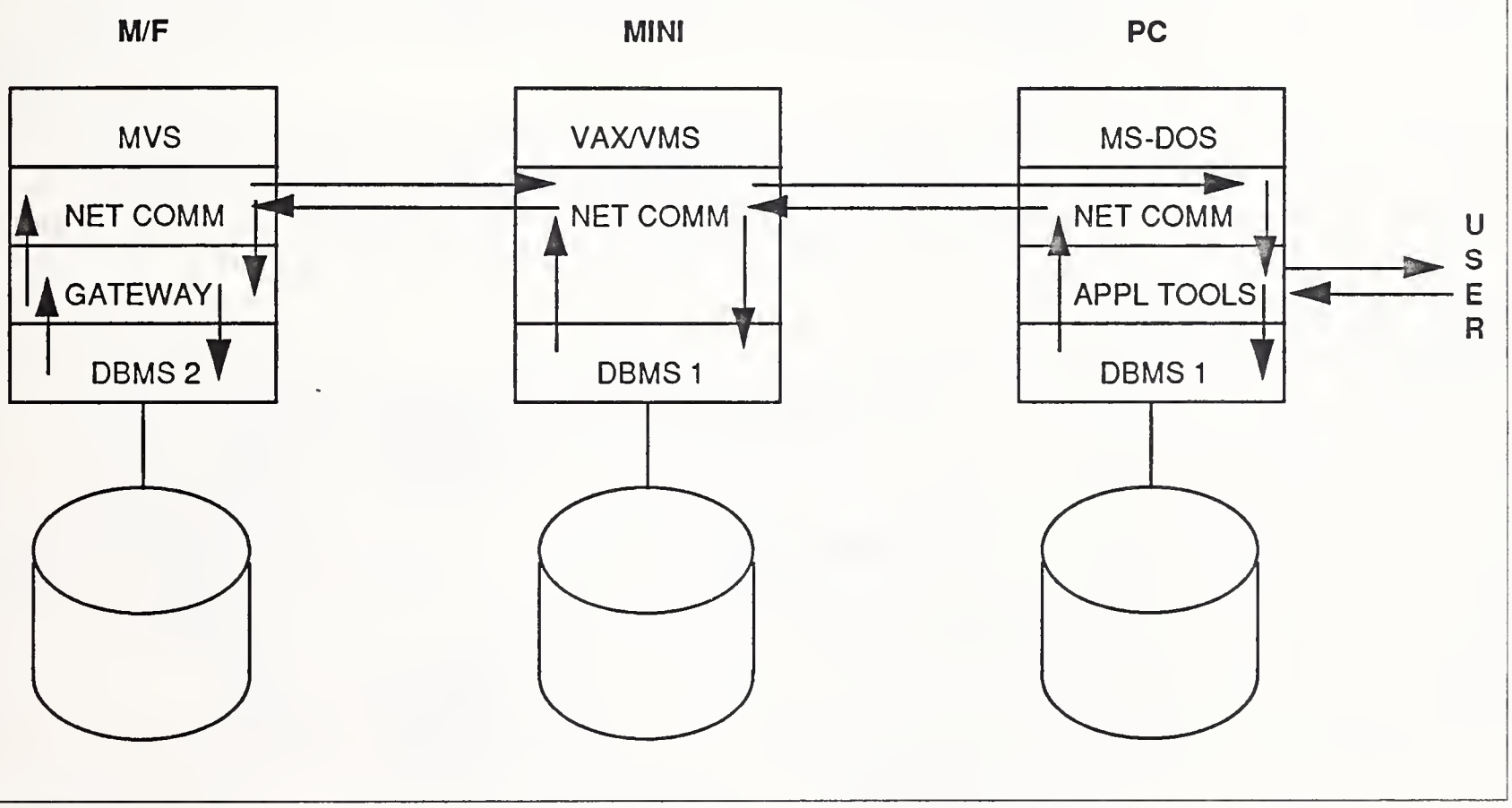


Exhibit V-2 provides another example of a DDBMS, and shows the complexity of a potential DDBMS environment. In this example, the data base includes relationships that cross platforms and multiple DBMSs. The DDBMS manages the relationships and interfaces with the communications capability to provide access to other DBMSs when necessary.

EXHIBIT V-2

Distributed DBMSs—A Data Communications Example



There are numerous implications of distributed data base processing. Key considerations include the following.

- Although it permits central administration of all data within an application, there are significant considerations of data redundancy and accuracy.

- Data base location will become transparent. Users of a data base will no longer need to know where data is located and will have access to a wide range of data located on a variety of platforms.

- Because of data redundancy and increased use, modeling will be much more important. Modeling and analysis will become increasingly dynamic as administrators seek to achieve optimum processing efficiency.

- The role of the administrator will become increasingly important to the application development process. Administrators will have greater knowledge of the business use of the data and will be able to influence application design.

2. Integrated Workstation Applications

With the advent of workstations and technologies such as DDBMS, the traditional approach to application design—placing the entire application on a single computer—is unnecessary. Instead, the application can be spread across a network, optimizing the capability of each tier of processing.

Exhibit V-3 shows a possible distribution of system functions between the workstation tier and the central processor that would lead to an integrated workstation application. Such applications will increase dependence on data base design and the data administrator.

EXHIBIT V-3

Integrated Workstation Applications Distribution of Functions

Workstation	Central Processor
User interface	Main file maintenance
Data entry and verification	Application network management
Secondary data management	Primary data management
	Primary system output
Current activity analysis and reporting	Weekly, monthly processing
Ad hoc analysis	Backup, security, etc.

B

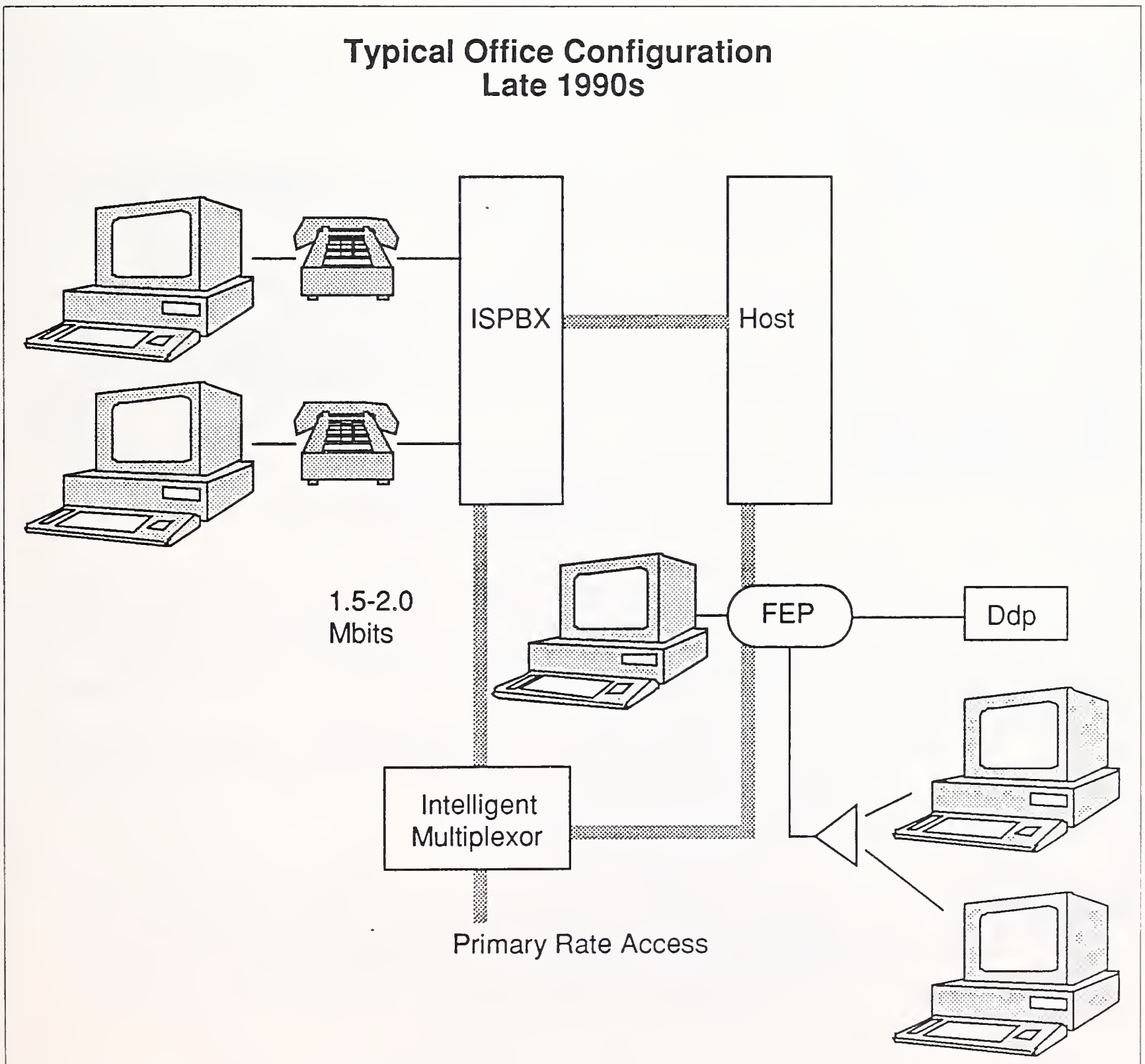
Impact of Network Growth

Changes in the telecommunications industry are having a dramatic effect on where and how users access systems and accomplish processing.

Among the changes, growth of ISDN will have a dramatic impact on the ability of users to access mainframes, minis and other micro-based systems.

As indicated in Exhibit V-4, the advent of highspeed (1.5-2.0 Mbit) connections between an organization's PBX, the host system and an intelligent multiplexor will have a major impact on where data bases are located and how users will be able to access systems.

EXHIBIT V-4



- With highspeed links, both locally through LANs, fiber cables and high-speed busses, and remotely through highspeed ISDN connections, users will be able to access data with equal ease at a local site, at the central site or at another remote site.

- With this expanded capability, defining the optimal location of a data base (considering use and frequency of access) through data modeling and application design will place increased importance on the data administration function.

- Adding to the flexibility of the end user will be the advent of integrated LAN systems that will further permit the sharing of data across departments, applications and users.

C

Role of the Data Dictionary

Current research indicates that the role of the data dictionary may be beginning to change.

While previous research indicated that use was highly fragmented, there are indications that greater acceptance of relational DBMSs will foster increased use of data dictionaries.

Previous research cited an example of a user that expended great effort to implement a data dictionary, and scrapped the work at the first budget cut. Current research indicates that a greater number of organizations are looking more seriously at use of the dictionary as a way to define corporatewide standards.

- Dictionaries being provided with relational systems are more understandable and are reported to be of greater value, even though data administrators are having to learn multiple systems.

- With dictionaries that are easier to use, users are beginning to define their own standards, forcing information systems staff to place greater emphasis on development or lose users to their own development processes. The result could be a wide range of conflicting standards.

As applications become more highly integrated (a definitive trend) and distributed systems begin to evolve, data dictionaries—even in an elemental state—will be a necessity.

D

Data Administration and CASE

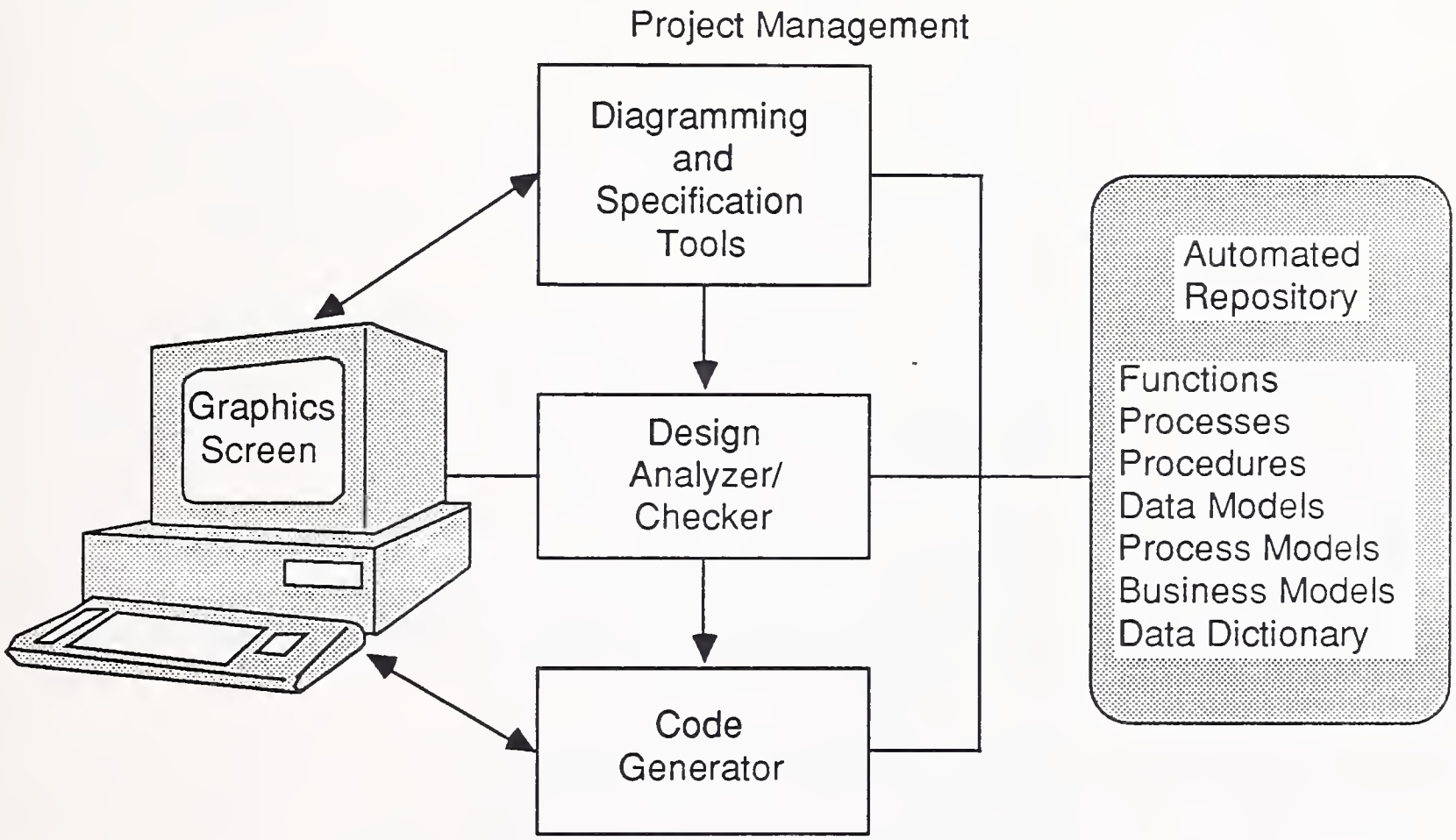
Computer-assisted systems engineering (CASE) technology has significant meaning for the process of data administration. The technology brings new data administration tools and will support the eventual reengineering of major applications for movement to new technology.

1. The Repository

Exhibit V-5 shows a complete CASE system. Of the key elements, the automated repository is perhaps the most important part of the system, certainly from the standpoint of the data administrator.

EXHIBIT V-5

Complete CASE System—A Schematic



While still in the early stages of development, the automated repository might well be viewed as the heart of the system. The repository is a data dictionary, a recorder of data and logic relationships, and is a source of previously used code. The repository has two important roles:

- For a single application, the repository serves as the source for all relationships and logic and provides a single source of historical information about a particular application.

- For a multiple application system, the repository serves as a single source of information about all data relationships and ensures consistency of data elements across applications.

The repository will be the tool that builds a solid interactive interface between systems development and data management. This technology is expected to develop quickly over the next several years.

2. Re-Engineering Tools

From the numerous comments made during INPUT's research, there is considerable evidence to suggest that if converting applications from existing DBMSs was a simple process, movement to new, relational systems would progress quickly. Likewise, if data dictionaries were simple and straightforward, more organizations would be using them.

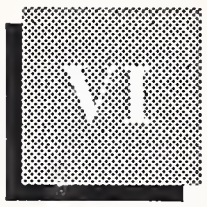
However, conversion is not simple, and older versions of data dictionaries are not straightforward. These situations have contributed to the lack of use of data dictionaries and the multiplicity of data base systems in many organizations. This situation is expected to change in the next few years, stimulated by the availability of re-engineering tools.

Recognizing that there is a vast number of existing applications, that many are poorly documented, and that creating new, complex, integrated systems would require major commitments, some CASE vendors are developing tools that will provide for the re-engineering of applications to meet new processing requirements.

The new tools are designed for use by the data manager and the DBA. Their purpose is to provide methods of understanding data structures and the relationships of existing applications so as to understand how they can fit into an expanded relational DBMS environment.

While development of these re-engineering tools is still in the early stages, information service organizations indicate that the availability of such tools could provide the needed impetus for progressing with use of data dictionaries and meeting the needs of the future.

Data Management: Future Environment



Data Management: Future Environment

To adequately portray the role of data management in the future, it is important to understand the evolution of organizational data management.

- Prior to the arrival of computers, and through the first fifteen years or more of their use, the controller of a business organization's information was the accounting organization.

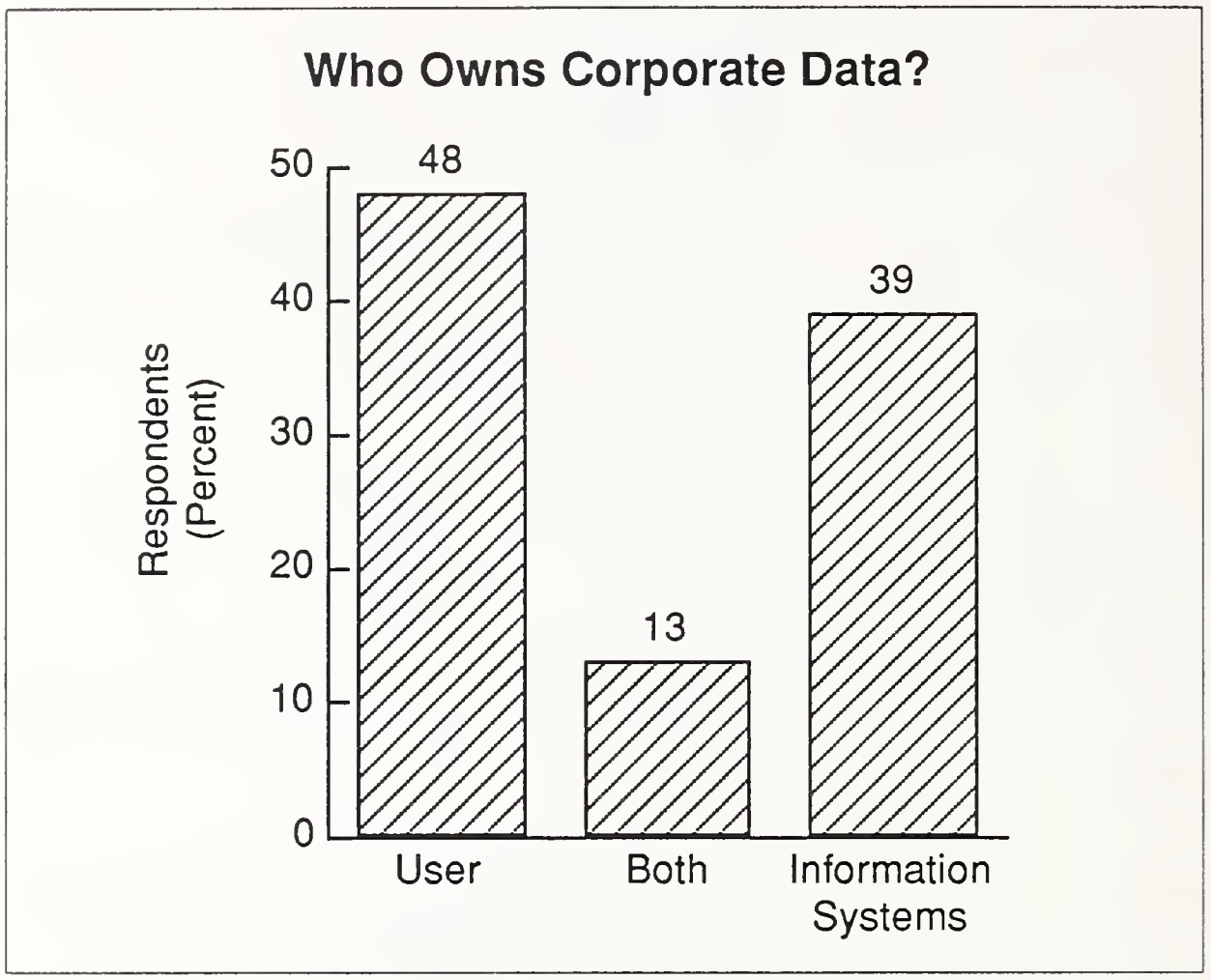
- For operational purposes outside of accounting, data contained in systems began to take on an ownership of its own. It became common for an application to serve more than one department and for operational and accounting data to be intertwined. Soon, operational data became more important than accounting data, and today's operational data is the focus of most information systems activities.

- With the advent of data base management technology, it became possible to integrate data from many applications. Once DBMS technology was in common use, the data it stored became the computer's data, owned by information systems.

- The counteraction to this apparent ownership by information systems was the tendency in the 1980s to distribute data to minicomputers and then to PCs. Very quickly, in some companies, small duplicate applications were created to permit a local department to own its data. Significant effort was expended to develop methods to transfer data files between the central site and operating departments and locations.

- Lower priced minis and desktop processing have altered the situation dramatically. Today, information service managers indicate that over 60% of data is owned by the user or jointly with information systems (Exhibit VI-1). Slightly less than 40% of information service executives indicate that they own corporate data.

EXHIBIT VI-1



- Much of this distribution of the computing process has been driven by improved user-oriented technology, greatly increasing the ease of use. Another driving factor has been the push by senior management to flatten organizations (reduce the number of organizational layers), reduce central staff, and place increased decision making in the hands of line management. For managers to operate their business units, they have taken an active role in planning and deployment of information technology.

- As with past experiences with swings between centralized and decentralized processing, distribution has resulted, in some organizations, in chaos in the information network. Because of this, there has been increased emphasis on integrating the information network while continuing to support operational demands.

To meet the challenge, data architecture and data management processes are necessary. Data is needed at many levels of the organization, but the information system is the keeper of the keys and is the only company function that can fulfill the need. Fulfilling the need is therefore a primary role of information systems.

A

Data Management's Role in the Organization

1. Characteristics of Today's Data Manager

Today's data manager generally does a good job, but he/she is an administrator, not a manager. Grounded in technology, today's administrator is generally detail-oriented, interested in "getting the job done", has strong orientation to the application, and has had little opportunity to understand the data needs of the numerous operating business units.

Exhibit VI-2 portrays the key characteristics of todays data base administrator. Note is quickly made that the administration function is geared to maintaining the "corporate" data base.

EXHIBIT VI-2

Characteristics of Today's Data Base Administrator

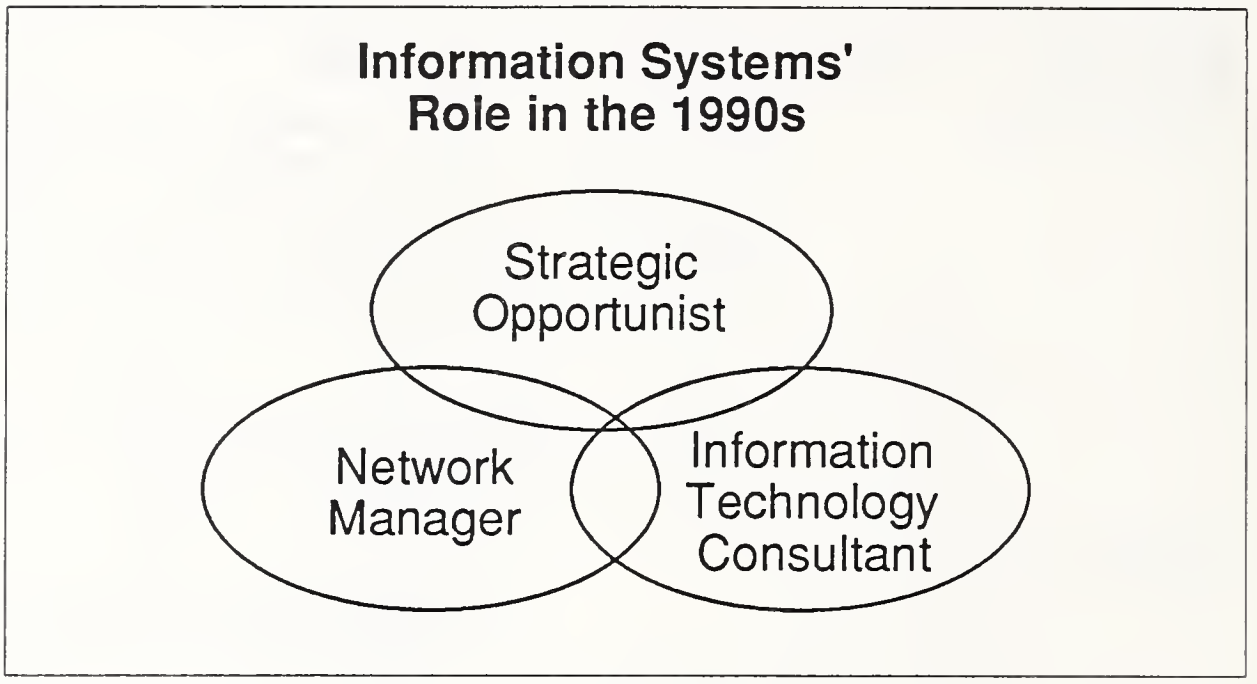
- Technologist—understands the DBMS in use
- Detailed orientation—performs a quality control function
- Service orientation—wants to get the job done
- Application orientation—focuses on the specific application and the data bases involved
- Corporate data focus—sees responsibility primarily tied to the corporate data bases and applications

2. Information Systems' Role in the 1990s

The role of the data manager (administrator) will change in the 1990s because the role of information systems will change.

INPUT believes that the role of the central information systems organization in the 1990s will be comprised of three basic elements, as shown in Exhibit VI-3.

EXHIBIT VI-3



- Information systems will assume the role of a network manager, ensuring the ability to access and transfer data between groups and locations, but not necessarily having control of the operating sites.

- Information systems will become the organization's chief technology (application) consultant. They will have extensive interface with operating units, to show how to obtain information needed for conduct of the business unit.

- As a strategic opportunist, the information systems group will be the point of contact for senior management to ascertain how the organization can better utilize information to strategic advantage.

Within this framework, information systems will have less responsibility for actual development work that directly affects each operating unit. Each unit will have greater ability to access and manipulate data necessary to manage its business.

3. Data Management in the 1990s

As a result of the shift in the role of information systems, the data management process will also change.

In the 1990s, INPUT believes that the data management process will be three-tiered, as indicated in Exhibit VI-4.

EXHIBIT VI-4

Data Management Process for the 1990s

Tier	Description	Functions
1	Infrastructure	Information architecture Data network modeling Technology selection
2	Definition	Definition standards Definition review Coding schemes
3	Execution	Data element definition Design execution DBMS installation and support

- Tier one will be the primary purview of information systems. User management will need to understand how the overall architecture supports the business structure and strategy, but it will be information system's responsibility to define and develop policies related to corporatewide architecture and data structure.

- At tier two, there will be shared responsibility. Information systems will define and review standards and structure universal coding schemes. However, this will be done in conjunction with users, who will need to ensure that their needs and requirements are represented.

- At tier three, users will become more autonomous. They will increasingly implement their own DBMSs and structure their own query and analysis programs. At this level, the corporate information service department will act primarily in a consulting role.

4. Characteristics of Tomorrow's Data Manager

As a result of the change in the role of information systems and the process of data management, the data manager must also change.

As indicated in Exhibit VI-5, the role of the administrator needs to change to become that of a manager. The manager needs to have a broader understanding of the business, needs to understand the broader value of information, and needs to be more of a consultant in dealing with users.

EXHIBIT VI-5

Characteristics Of Tomorrow's Data Manager

- Generalist—understands the value of information technology and can market its value
- Broad picture—has top-down view of the information infrastructure of the organization
- Consulting orientation—advises the development and data administration groups on specific development projects
- Business data orientation—views the data and information flow across the entire organization
- Systems integration focus—focuses on data sharing by groups of applications in support of a strategic business operation

Compared side-by-side (Exhibit VI-6), the data manager and the data administrator are distinctly different individuals. While the administrator focuses on technological detail, the manager will focus on the provision of information to support business operations.

EXHIBIT VI-6

Data Manager Versus Data Administrator—A Comparison

Data Manager	Data Administrator
Generalist	Technologist
Broad picture	Detailed orientation
Consulting orientation	Service orientation
Business information orientation	Application orientation
Systems integration focus	Corporate data focus

B

Roadblocks to Success

The transition of data management from a detailed, technical, administrative orientation to a general, strategic, corporate orientation will not happen quickly.

Due to a number of roadblocks summarized in Exhibit VI-7, the transition will not be made until well into the 1990s. However, the transition must be made, and organizations that fail to make it will be at a disadvantage in the strategic use of technology. They could find themselves at a competitive disadvantage.

EXHIBIT VI-7

Future Data Management Role Inhibiting Factors

- Who owns the data
- Skills of today's data administrator
- Status of the information network
- Status of DBMS technology
- Understanding need
- End-user training

1. Who Owns the Data?

The issue of who owns corporate data has been smoldering for years. It underlies many information technology decisions that affect users.

Following years of having unresponsive central information service organizations, and fueled by PC and new relational systems, many users have adopted the "I'd rather do it myself" attitude. The result has been a patchwork of systems that are frequently uncoordinated, and of end-user development that is uncontrolled.

While somewhat overstated, the fact remains that information systems will have to reorient and build a framework under which ownership is apparent. Ownership will generally be granted to the most appropriate end user. Infrastructure and development processes will need to support easy sharing with others.

2. Skills of Today's Data Administrator

With a foundation in technology and the development of large, centralized applications, many of today's administrators are not prepared to meet the requirements of data management of the 1990s.

In addition, few organizations have granted the data administration function the same level of responsibility of other groups in information systems. This situation has had several results:

- The data administration function has continued to develop along technical lines. With continued focus on large central systems, the user is frequently given little consideration, let alone involvement, in the systems design process.

- Data administration is viewed as a technical necessity rather than as a contributor to corporate assets.

- There is little evidence that data administration is being marketed internally as a valuable corporate management function.

- There is little evidence that the end user is being prepared to share in the responsibility and benefits of the data management process.

3. Status of the Information Network

DBMSs at central sites, minicomputers in departments and divisions, and PCs for users have contributed to the availability of vast amounts of information throughout the organization. Systems are increasingly connected, but are far from integrated. The availability of information has increased dramatically during the 1980s, but users' needs are far from satisfied.

Information technology is increasingly a foundation for the conduct of business, and users are quickly becoming decision makers on many, if not most, major decisions. However, much work remains before applications become integrated and the network is open to all users to obtain information when and where needed.

4. Status of DBMS Technology

Relational technology is increasingly viewed as a truly new and better answer for many computing requirements. However, information systems are just beginning to move into RDBMSs, and there is considerable work to train users and IS staff. Existing systems carry ten to fifteen years of investment in traditional DBMS technology. Most data managers simply see major technical tasks ahead and have not been able to address the broader aspects of the data management process.

5. Understanding the Need

In an era of decentralized management and limited corporate staffs, the need for a corporatewide data management process is difficult for senior corporate management to understand.

Corporatewide data management is an abstract concept. The level of effort necessary to sell the concept will be equal to, and perhaps greater than, the effort that was needed to sell the concepts of early DBMSs. But the effort must be made if information is to be used as an asset.

6. Training the End User

Emergence of 4GLs, the PC and the corporate information center has resulted in considerably increased interaction between information systems and the end user over the past few years.

However, the success of the interaction can be questioned. While users have been assisted with technical solutions, the extent to which users have learned the value of corporate information and the need for data management is questionable. Whether as an oversight or as a loss of focus, information systems have not trained users in the value of corporatewide information technology.

The training required ranges from analysis and design, to data base concepts, to security. Until information systems begin training in these areas, other aspects of the central information system's role, including data management, will be met with some resistance.

C

Planning for the Future

In its 1988 *Information Systems Planning Report*, INPUT identified a number of priorities for 1989 and beyond. In the area of data management, there were two that were key:

- For 1989—Audit the data management function.
- Beyond 1989—Refocus data management to a companywide orientation.

To date, there appears to have been only limited effort to audit data management. While many organizations may be able to begin refocusing without an audit, there is a risk of mis-directing planning efforts into areas where there is little need.

In all, INPUT believes that the steps previously identified to meet the information system needs of the 1990s bear repeating.

1. Audit Data Administration

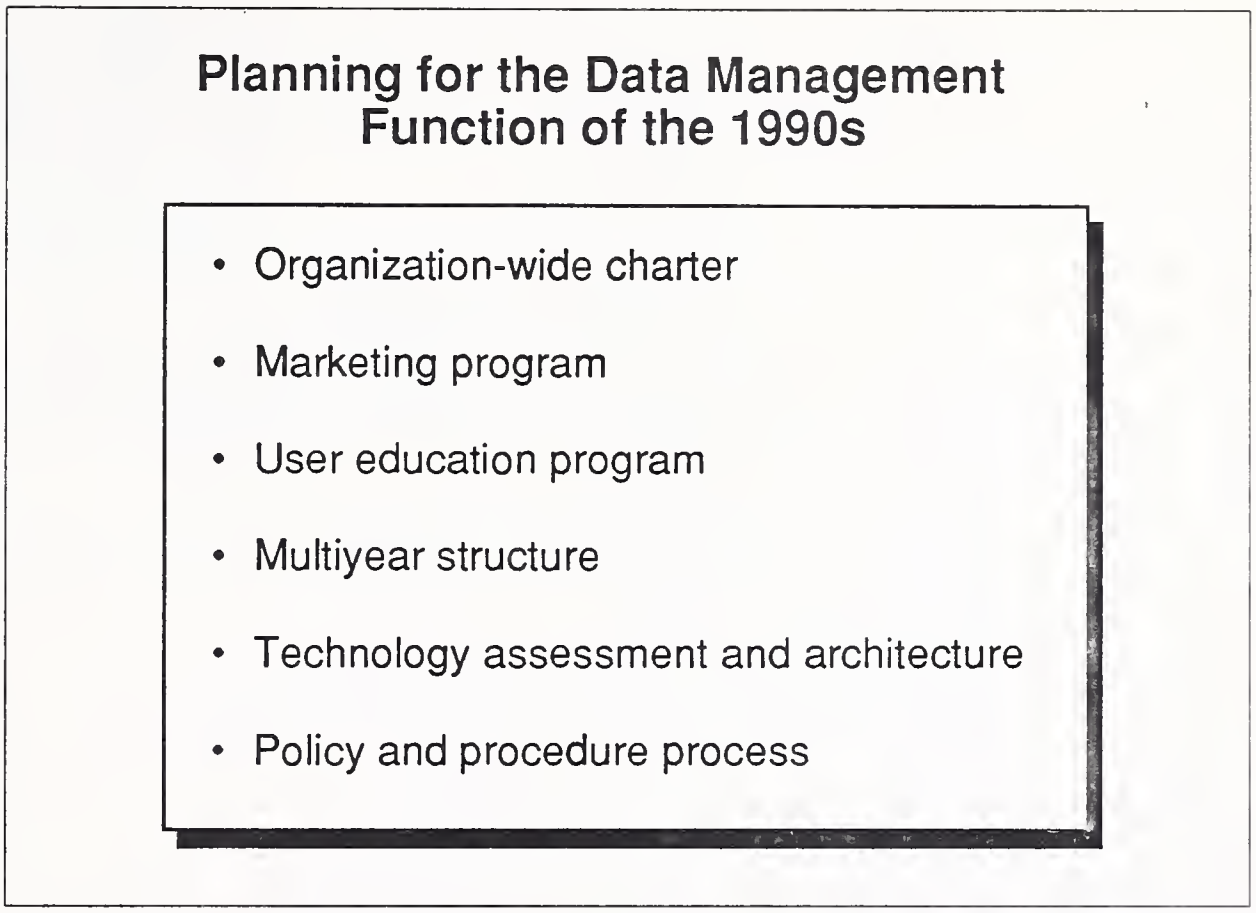
Many data administration functions are ten years old, or older, and could benefit from an audit to determine whether they will be able to meet the needs of the 1990s.

- The audit should result in a thorough understanding of the relationship of the data administration function and the changing user and technology environment.
- A byproduct of the audit should be a set of action plans to ensure that the administrator has the tools necessary to become a manager.

2. Planning for Change

Following an audit, building an action plan is an obvious, but necessary step. While the plan will vary to meet an organization's unique needs, it should include at least the elements identified in Exhibit VI-8.

EXHIBIT VI-8



- Organization Charter—An organization charter is frequently the most problematic aspect of a plan. Developing a charter is difficult, requiring recognition of responsibilities that will need to be given up and others that will need to be assumed. Because of the difficulty, it is frequently ignored. But gaining consensus between information systems and users about the role of data management needs to be the first step.

- Marketing Program—For individuals oriented to accomplishing technical tasks, marketing is frequently not valued. However, a plan that will gain acceptance of users and senior management is necessary to the success of data management.

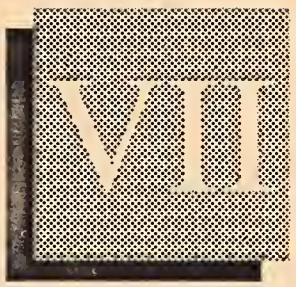
- User Education Program—With users better able to understand DBMS (relational) technology, training emphasis needs to shift from one of process to one of concept. Users need to understand, and accept, the concepts of corporatewide information and the value of accessibility by all corporate users.

- Multiyear Structure—Making the transition from technologist to generalist will not happen overnight. Plans need to reflect the long-term nature of the transition from data administration to data management.

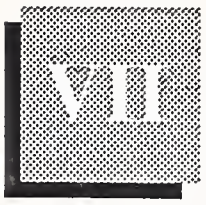
- Technology Assessment and Architecture—There needs to be a thorough understanding of the organization's technology and architecture. Only with a comprehensive understanding of the existing environment can strategies be developed to make changes for the future.

- Policy and Procedure Process—Since tasks related to definition and execution will be quickly decentralized to the end user, the data management process must begin to focus on corporatewide policies and procedures.

Making the change from data administration to data management must be initiated by information systems. Neither senior management nor the user will be the facilitator of the change. Information systems must make the decision that data management is a valuable corporate service, and assume the responsibility to provide it.



Conclusions and Recommendations



Conclusions and Recommendations

As indicated in the introduction, a primary purpose of this study was to consider development of DBMSs and data management "one year later". INPUT's interest was in assessing the extent to which there had been movement in accomplishing the transition from the historical data administration function to data management.

Many of the conclusions and recommendations that follow are the same as, or closely aligned with last year's. Modifications have been made where appropriate to reflect changes that were noted during the current research.

A Conclusions

Information systems are poised to change. In a number of respects, changes are happening more rapidly than information systems are able to accommodate. Unless information systems begin to make changes in the management of corporate data, systems will become increasingly fragmented and information will not be usable as a corporate competitive asset.

- Relational technology is fast becoming the underlying technology. Only investment made in older technology stands in the way of a complete changeover.

- With RDBMSs available on all three tiers of the information network, applications are being developed throughout the organization. In many cases, information systems are being left out of the loop.

- RDBMSs will provide opportunities for control and coordination of the data management process as well as for data accessibility that have not been available in the past. Opportunities include:

- SQL as a standard language
- Multitier portability
- Increased understanding, and consequently use, by the end user
- Integration of data dictionaries

- Other technologies, such as CASE, are having a significant stimulating effect on the development of new systems.

- Since relational technology is still fairly new, there will be continuous improvements over the next several years. The changes will provide evolutionary (and some revolutionary) growth, necessitating expansion of the data administration function.

  - Distributed DBMSs will have a revolutionary effect on the way systems are designed and managed.

  - The ability to make changes to older systems will result in the overhaul of integration systems previously considered too costly.

- The priority for the next few years will be to interconnect the information network while maintaining its distributed structure. New DBMS technology and a proactive data management process are essential to achieve this priority.

B

Recommendations

Like the conclusions, few changes in the recommendations are warranted. The fundamental recommendation remains that information systems must realize that the traditional data administration process is no longer adequate to serve and help control the information network. It is essential that a broader view be assumed and a top-down management process evolve over the next several years.

1. Data Management Objectives for 1990

The following objectives are recommended as actions to be accomplished during the coming year. The purpose is to establish a base from which to grow.

- Audit the data administration process

- Define a data management charter

- Since PC users learn best by doing, select an SQL DBMS as a standard for PC use

- Develop internal marketing plans that increase emphasis on the data dictionary. Use of data dictionaries will be a necessity in the DDBMS environment

- Introduce a generalist into the data administration function. Select an individual who has basic technological skills, but is oriented to consulting and to understanding corporate business information requirements.

- Audit the use of minicomputer- and PC-based DBMSs to ensure that they are in line with existing corporate information policies or are consistent with corporate direction.

2. Data Management Objectives Beyond 1990

The following objectives are recommended as data management priorities beyond 1990. They assume that a comprehensive data management process is an objective.

- Broaden the responsibility of data management to all levels of the information network.

- Launch a user training program in all aspects of data base use.

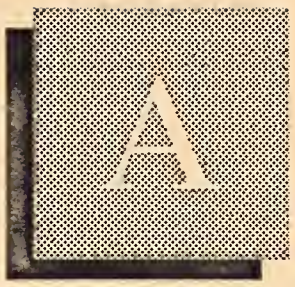
- Develop a data architecture, and gain acceptance from user management.

- Market the data management process at all levels of the organization.

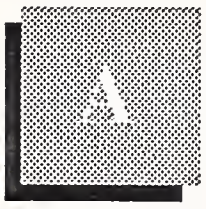
- Begin to segment the data management process for the eventual shift of some tasks to unit and end-user information systems groups. Shift the responsibility as soon as the units or users are prepared to accept it.

- Explore the impact of CASE on the data management and administrative process.

- Support experimentation with distributed DBMSs. This extended capability of relational DBMSs will eventually provide the control mechanism required to permit a true balance between data management control and freedom of data use by the end user.



Appendix: Data Administration Questionnaire



Appendix: Data Administration Questionnaire

INTRODUCTION

Hello, my name is _____. I'm calling for INPUT. Last year we talked with you or an associate about your data administration function and the changes it is undergoing.

At this time, I would like to take a few minutes to follow up on our conversation and explore some aspects of how the use of data bases and the data administration function may have changed since last year.

Is this a convenient time or would you prefer that we set an appointment for another time? It should take only 20 minutes and INPUT will be pleased to share a summary of the results with you. Your responses to the questions will be kept confidential and, as I mentioned, INPUT will be pleased to send you a complimentary copy of the results.

Individual Completing the Questionnaire

Name _____

Title _____

Organization _____

Address _____

_____

Telephone _____

First, I would like to confirm a few characteristics of your data administration environment. This information was obtained from you last year.

1. We understand that you are currently using the following DBMSs. Are these correct? **(Note - Question 1a asks where the DBMS is installed).**

DBMS	Mainframe	Mini	PC
_____	_____	____	____
_____	_____	____	____
_____	_____	____	____
_____	_____	____	____

1a. For each DBMS identified in 1, please indicate whether it is implemented on a mainframe, mini, or PC.

2. Are there any others in Use? Yes ____ No ____

   Which ones? _____

3. What is the number of employees in the data administration staff?
   1989 _____ 1990 Plan _____

4. Is there more than one data administration organization?
   Yes _____ No _____ **(If No skip to 6)**

5. How many separate data administration organizations are there within your firm? _____

6. Which of the following categories of data is the corporate data administration function responsible for? **READ THE LIST**

Central Data	Y___	N___
Division Data	Y___	N___
Departmental Data	Y___	N___
Distributed Systems	Y___	N___
Minicomputer Data Bases	Y___	N___
PC Data bases	Y___	N___
Non-Data base Systems	Y___	N___
Other (_____)	Y___	N___

7. How have data administration responsibilities changed in the past year?

   _____

   _____

   _____

8. How is the performance of data administration measured in your company?

_____

_____

_____

9. What are the top three issues facing your data administration function?

1. _____
2. _____
3. _____

10. What are the top three objectives for your data management function over the next two years?

1. _____
2. _____
3. _____

11. On a scale of one (poor) to five (excellent), how would you rate the effectiveness of the corporate data administration function today?

1 _____ Poor
2 _____
3 _____ Average
4 _____
5 _____ Excellent

} **Go to 13**

12. **(If answer to 11 was three or lower).** What plans exist to improve the effectiveness of the corporate data administration function?

_____

_____

_____

13. On a scale of one (poor) to five (excellent), how would you rate the quality of your use of data dictionary capabilities?

1 _____ Poor
2 _____
3 _____ Average
4 _____
5 _____ Excellent

} **Go to 15**

14. **(If answer to 13 was three or lower)**. What plans exist to improve the effectiveness of your data dictionary capabilities?

_____

_____

_____

15. Who sets definitions for data in your organization?
User _____ Information Systems _____ Both _____

16. Who owns the data in your organization?
User _____ Information Systems _____ Both _____

17. Please describe what data ownership means?

_____

_____

_____

18. Please describe the policies on data administration in your organization?

_____

_____

_____

19. Do these policies apply to...

Corporate IS Department	Yes___	No___
User Computing	Yes___	No___
Departmental Computing	Yes___	No___
Operating Division	Yes___	No___
Personal Computing	Yes___	No___

20. On a scale of one (low) to five (high), please rate the level of data definition policies:

1 _____ Low
2 _____
3 _____ Average
4 _____
5 _____ High

Next, we would like to learn what the primary applications are for mainframes, minis, and personal computers.

21. What are the primary DBMS applications on your mainframe(s)?

_____

_____

_____

22. What are the primary DBMS applications on your minicomputers?

_____

_____

_____

23. What are the primary DBMS applications on your microcomputers/PCs?

_____

_____

_____

24. Are there any new DBMSs under consideration? For example DB2, Oracle, Adabas, etc.
Yes _____ No _____

25. Which ones and for what applications?

DBMS	Applications
_____	_____
_____	_____

The next set of questions addresses your use of relational data base management systems.

26. Has your organization developed any relational DBMS-based applications?
Yes _____ No _____ **(If No, skip to 37)**

27. When did RDBMS use begin? _____

28. Is the use primarily:

Mainframe _____ Mini _____ PC _____

29. How many applications have...

Been developed? _____ Are in process? _____

30. Please identify three key criteria that have been used to decide to use an RDBMS versus a conventional DBMS:

1. _____

2. _____

3. _____

31. What has been the impact on data administration regarding:

1. Training

_____

_____

_____

2. Data Modeling

_____

_____

_____

3. Data Administration/Data Dictionary

_____

_____

_____

4. Other

_____

_____

_____

32. Please rate the productivity of data administration with RDBMSs compared to conventional DBMSs:

Worse _____ Same _____ Better _____

33. How much better or worse?

10% ___ 25% ___ 50% ___ 75% ___ Other ___

34. Please rate the productivity of application development with RDBMSs compared to conventional DBMSs:

Worse _____ Same _____ Better _____

35. How much better or worse?

10% ___ 25% ___ 50% ___ 75% ___ Other ___

36. If using mainframe- and minicomputer-based RDBMSs, please compare the activities/productivity/etc.:

_____

_____

_____

End User Use of RDBMSs

37. Are end users developing systems with RDBMSs?

Yes _____ No _____ **(If no, skip to 42)**

38. Is the corporate data administration group coordinating this use?

Yes _____ **(If yes, skip to 40)** No _____

39. Who is coordinating these activities? _____

_____

40. How would you rate the coordination?

Less than adequate _____
Adequate _____
More than adequate _____

41. What type of training is being provided to end users?

_____

_____

_____

Personal Computer/Workstation DBMSs

42. What type of workstation-based DBMS activity is there?

_____

_____

_____

43. What are some examples of workstation-based DBMS use?

1. _____

2. _____

3. _____

Distributed Data Base Management

(Interviewer Note: A distributed data base is a data base which is replicated, in whole or in part, on one or more systems located remotely from a central processing facility.)

44. Is a distributed DBMS...

1. In use in your organization? Yes ___ No ___

2. Under consideration? Yes ___ No ___

3. Under investigation? Yes ___ No ___

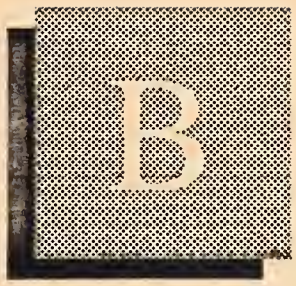
45. What distributed data base is in use?

_____

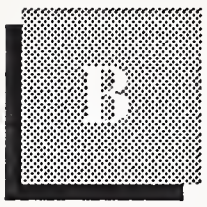
46. Who is coordinating your research into use of distributed DBMSs?

Corporate group _____ End users _____ Other ____

That's it! I want to thank you for your help. Let met double check you adddress in order to send you an Executive Overview of the report.



Appendix: Definitions



Appendix: Definitions

- *Data Administration* - The more traditional definition of the data control process that addresses the technical definition and control tasks required by DBMS technology, in particular the traditional data base management systems.

- *Data Base Administrator (DBA)* - Both the task and the specific job assignment of the individual who establishes the detailed definition logic required in programming a DBMS application.

- *Distributed Data Base* - A data base system which is replicated, in whole or in part, on one or more systems located remotely from a central processing facility.

- *Relational Data Base Management System (RDBMS)* - Those data base management systems that are based on the SQL language standard and that provide the flexible, table-like, view of data relationships.

- *Traditional/Conventional Data Base Management Systems (DBMSs)* - Those traditional data base management systems that are hierarchical or use the Codasyl standard for specifying data relationships.

Report Quality Evaluation

To our clients:
To ensure that the highest standards of report quality are maintained, INPUT would appreciate your assessment of this report. Please take a moment to provide your evaluation of the usefulness and quality of this study. When complete, simply fold, staple, and drop in the mail. Postage has been pre-paid by INPUT if mailed in the U.S..

Thank You.

1. Report title: ***Data Base Systems Development*** (UMDA)

2. Please indicate your reason for reading this report:
   - ☐ Required reading
   - ☐ Area of high interest
   - ☐ Area of general interest
   - ☐ New product development
   - ☐ Business/market planning
   - ☐ Product planning
   - ☐ Future purchase decision
   - ☐ Systems planning
   - ☐ Other _____

3. Please indicate extent report used and overall usefulness:

	Extent		Usefulness (1=Low, 5=High)				
	Read	Skimmed	1	2	3	4	5
Executive Overview	☐	☐	☐	☐	☐	☐	☐
Complete Report	☐	☐	☐	☐	☐	☐	☐
Part of Report (_____ %)	☐	☐	☐	☐	☐	☐	☐

4. How Useful was:

Data presented	☐	☐	☐	☐	☐
Analyses	☐	☐	☐	☐	☐
Recommendations	☐	☐	☐	☐	☐

5. How useful was the report in these areas:

Alert you to new opportunities or approaches	☐	☐	☐	☐	☐
Cover new areas not covered elsewhere	☐	☐	☐	☐	☐
Confirm existing ideas	☐	☐	☐	☐	☐
Meet Expectations	☐	☐	☐	☐	☐
Other _____	☐	☐	☐	☐	☐

6. Which topics in the report were the most useful? Why?_____

_____

7. In what ways could the report have been improved? _____

_____

8. Other comments or suggestions: _____

_____

Name _____ Title _____

Department _____

Company _____

Address _____

City _____ State _____ ZIP _____

Telephone _____ Date Completed _____

Thank you for your time and cooperation.

M&S 633/01 12/89

INPUT

FOLD HERE

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL
First Class Permit No. 982 Mountain View, CA

POSTAGE WILL BE PAID BY ADDRESSEE

Attention: Marketing Department

INPUT
1280 Villa Street
Mountain View, CA 94041-9912

FOLD HERE