U.S. Department
of Commerce

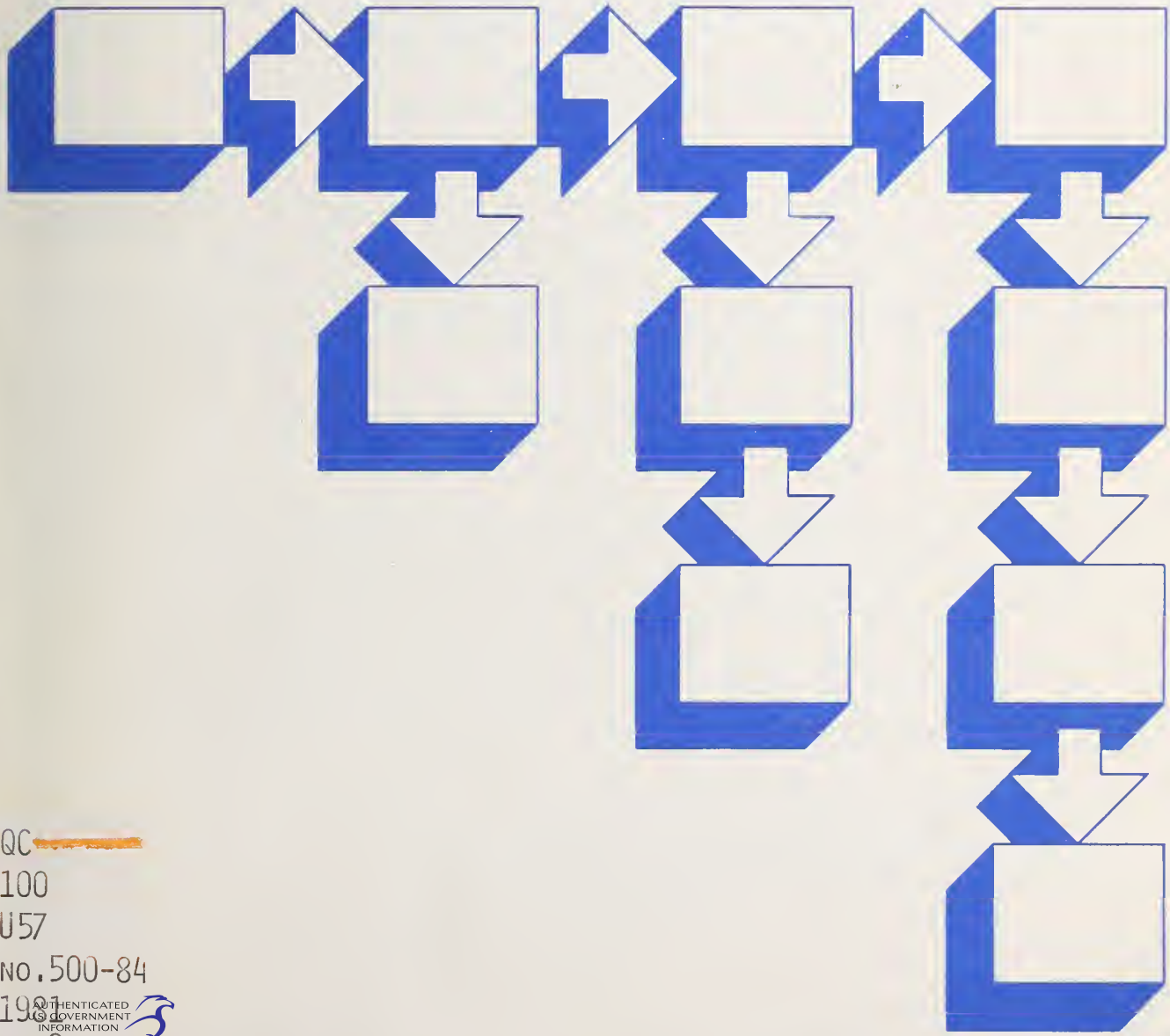
National Bureau
of Standards

# Computer Science and Technology



NBS Special Publication 500-84

## Costs and Benefits of Database Management: Federal Experience



# NATIONAL BUREAU OF STANDARDS

The National Bureau of Standards[1] was established by an act of Congress on March 3, 1901. The Bureau's overall goal is to strengthen and advance the Nation's science and technology and facilitate their effective application for public benefit. To this end, the Bureau conducts research and provides: (1) a basis for the Nation's physical measurement system, (2) scientific and technological services for industry and government, (3) a technical basis for equity in trade, and (4) technical services to promote public safety. The Bureau's technical work is performed by the National Measurement Laboratory, the National Engineering Laboratory, and the Institute for Computer Sciences and Technology.

**THE NATIONAL MEASUREMENT LABORATORY** provides the national system of physical and chemical and materials measurement; coordinates the system with measurement systems of other nations and furnishes essential services leading to accurate and uniform physical and chemical measurement throughout the Nation's scientific community, industry, and commerce; conducts materials research leading to improved methods of measurement, standards, and data on the properties of materials needed by industry, commerce, educational institutions, and Government; provides advisory and research services to other Government agencies; develops, produces, and distributes Standard Reference Materials; and provides calibration services. The Laboratory consists of the following centers:

Absolute Physical Quantities[2] — Radiation Research — Thermodynamics and Molecular Science — Analytical Chemistry — Materials Science.

**THE NATIONAL ENGINEERING LABORATORY** provides technology and technical services to the public and private sectors to address national needs and to solve national problems; conducts research in engineering and applied science in support of these efforts; builds and maintains competence in the necessary disciplines required to carry out this research and technical service; develops engineering data and measurement capabilities; provides engineering measurement traceability services; develops test methods and proposes engineering standards and code changes; develops and proposes new engineering practices; and develops and improves mechanisms to transfer results of its research to the ultimate user. The Laboratory consists of the following centers:

Applied Mathematics — Electronics and Electrical Engineering[2] — Mechanical Engineering and Process Technology[2] — Building Technology — Fire Research — Consumer Product Technology — Field Methods.

**THE INSTITUTE FOR COMPUTER SCIENCES AND TECHNOLOGY** conducts research and provides scientific and technical services to aid Federal agencies in the selection, acquisition, application, and use of computer technology to improve effectiveness and economy in Government operations in accordance with Public Law 89-306 (40 U.S.C. 759), relevant Executive Orders, and other directives; carries out this mission by managing the Federal Information Processing Standards Program, developing Federal ADP standards guidelines, and managing Federal participation in ADP voluntary standardization activities; provides scientific and technological advisory services and assistance to Federal agencies; and provides the technical foundation for computer-related policies of the Federal Government. The Institute consists of the following centers:

Programming Science and Technology — Computer Systems Engineering.

[1]Headquarters and Laboratories at Gaithersburg, MD, unless otherwise noted; mailing address Washington, DC 20234.
[2]Some divisions within the center are located at Boulder, CO 80303.

# Computer Science and Technology

NBS Special Publication 500-84

# Costs and Benefits of Database Management: Federal Experience

Jesse M. Draper

Center for Programming Science and Technology
Institute for Computer Sciences and Technology
National Bureau of Standards
Washington, DC 20234

**U.S. DEPARTMENT OF COMMERCE**
**Malcolm Baldrige, Secretary**

**National Bureau of Standards**
Ernest Ambler, Director

Issued November 1981

## Reports on Computer Science and Technology

The National Bureau of Standards has a special responsibility within the Federal Government for computer science and technology activities. The programs of the NBS Institute for Computer Sciences and Technology are designed to provide ADP standards, guidelines, and technical advisory services to improve the effectiveness of computer utilization in the Federal sector, and to perform appropriate research and development efforts as foundation for such activities and programs. This publication series will report these NBS efforts to the Federal computer community as well as to interested specialists in the academic and private sectors. Those wishing to receive notices of publications in this series should complete and return the form at the end of this publication.

Library of Congress Catalog Card Number: 81-600152

National Bureau of Standards Special Publication 500-84
Nat. Bur. Stand. (U.S.), Spec. Publ. 500-84, 106 pages (Nov. 1981)
CODEN: XNBSAV

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 1981

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402
Price $5.00
(Add 25 percent for other than U.S. mailing)

TABLE OF CONTENTS


Page

1. INTRODUCTION ..... 1

2. FEDERAL DATA PROCESSING AND DBMS'S ..... 2

2.1 ADP Expenditures ..... 3

2.1.1 Personnel ..... 3
2.1.2 Hardware ..... 5
2.1.3 Outside Services ..... 6

2.2 Software Development ..... 6

2.3 Potential Impact of DBMS's ..... 8

2.4 DBMS Software Availability ..... 8

2.5 Federal Inventory of DBMS Software ..... 10

2.6 Database on Timesharing Services ..... 11

2.7 Potential for DBMS Growth in Government ..... 12

2.7.1 Computing Power ..... 14
2.7.2 On-line Secondary Storage ..... 14

3. COST AND BENEFIT FACTORS ..... 15

3.1 Benefits ..... 16

3.1.1 Centralized Data Management ..... 16
3.1.2 Reduced Redundancy ..... 16
3.1.3 Data Independence ..... 16
3.1.4 External User Views ..... 17
3.1.5 Logical and Physical Storage Structures ... 17
3.1.6 On-line Processing and Ad-hoc Queries ..... 18
3.1.7 Access and Integrity Control ..... 18

3.2 Costs ..... 18

3.2.1 Vulnerability and User Conflicts ..... 18
3.2.2 Performance ..... 19
3.2.3 Personnel Skills ..... 19
3.2.4 Overhead and Initial Costs ..... 19

3.3 Summary ..... 20

4. INTERVIEWS WITH FEDERAL AGENCIES ..... 20

4.1 Applications ..... 21

4.2 Requirements and Cost/Benefit Analyses ..... 21

4.3 DBMS Selection ..... 23

4.4 Benefits ..... 24

4.4.1 Centralized Data Management ..... 25
4.4.2 Data Independence ..... 26
4.4.3 High-Level Query Facilities ..... 26
4.4.4 High-Level User Languages ..... 27

4.5 Costs ..... 28

4.5.1 Initial Costs ..... 28
4.5.2 Resource Requirements ..... 30
4.5.3 Increased Need for Technical Skills ..... 30
4.5.4 Portability ..... 31

4.6 Future Directions ..... 31

5. CONCLUSION ..... 32

6. ACKNOWLEDGMENTS ..... 33

REFERENCES ..... 34

APPENDIX A: QUESTIONNAIRE ..... 37

APPENDIX B: A TABULATION OF RESPONSES TO THE QUESTIONNAIRE ..... 44

APPENDIX C: INTERVIEW REPORTS ..... 50

APPENDIX D: DBMS COST/BENEFIT PARAMETERS ..... 92

# Costs and Benefits of Database Management: Federal Experience

Jesse M. Draper

The Federal Government has a large investment in a wide variety of database management systems (DBMS's) and in diverse applications using those systems. Data managers from eight Federal agencies report that they are pleased with the power, flexibility, and cost effectiveness of DBMS's, whether they use their system as sophisticated access software or make it the focal point of a large, integrated data processing system. While DBMS's may save an agency money in the long run, their major immediate benefits are almost always increased functionality and productivity, not reduced cost. Centralized data and high-level user languages reduce the time for developing and maintaining computer programs, and the latter enable nonprogrammers to use computers effectively.

The amount of cost/benefit analysis an agency needs before deciding to buy a DBMS increases with the complexity of the application. The experiences of the interviewed agencies, together with a structured list of cost/benefit parameters, should help Federal managers in understanding the potential value of DBMS technology and in defining their requirements for data management.

Key words: application development; computer software; cost/benefit analysis; database management system; data management; program maintenance; requirements studies; software procurement.

## 1. INTRODUCTION

This report discusses the costs and benefits of database management systems (DBMS's) as they are used by executive agencies of the Federal Government. For the purposes of the report, a DBMS is any general-purpose, application-independent software package used in association with on-line mass-storage to facilitate the entry, storage, processing, and retrieval of logically structured data. To provide the Federal community with information about DBMS's,

NBS has gathered material about four related topics. Several sources, including the Office of Management and Budget and the General Services Administration, have provided useful data for Chapter 2, which estimates the size of the Federal data processing effort and the number of DBMS's already installed in Federal ADP shops. These data suggest that DBMS's can provide the Government with substantial benefits and savings during the next decade. Textbooks and articles supplied the theoretical costs and benefits of database management that constitute Chapter 3. To determine whether the Government has actually gained these benefits and incurred these costs, NBS interviewed data managers from eight Federal agencies, and Chapter 4 analyzes and discusses the results of the interviews. More detailed information about the interviews appears in the first three Appendixes. Chapter 5 briefly states the conclusions of the report. Finally, Appendix D covers the fourth topic of interest to Federal users of DBMS's. It presents some parameters that Federal agencies may want to consider in selecting a DBMS. The list of parameters and their descriptions are preliminary results of an NBS contractor who is developing a cost/benefit decision model for data management.

Because this report addresses the practical experiences of Federal agencies, it is necessary in a number of places to mention vendors and commercial products. The inclusion or omission of a particular company or product does not imply either endorsement or criticism by NBS.

## 2. FEDERAL DATA PROCESSING AND DBMS'S

Every year the Federal Government spends a large amount of money to maintain its existing hardware and software inventory, to buy new equipment, programs, and services, and to pay for ongoing ADP operations. Public expenditures for data processing differ considerably from those in the private sector, and the differences directly affect the prospects for DBMS usage in the Government. This chapter characterizes those expenditures, identifies important differences between Federal and private data processing, and discusses both the current Federal inventory of DBMS's and the potential for future growth in the inventory.

## 2.1 ADP Expenditures

In FY 1981 executive agencies of the Federal Government will spend approximately $6.0 billion on electronic data processing and related activities. This figure represents slightly less than 1 percent of the total Federal budget and about 10 percent of all data processing expenditures in the United States [OMB80]. Actual expenditures have been increasing at an annual rate of 13 percent since 1973, representing a real growth in constant dollars of slightly more than 4 percent per year [OMB 80]. As Robert V. Head has pointed out, the Federal ADP budget is much larger than the ADP budget of even the largest corporations. "Even the individual budgets of relatively small agencies, such as HUD at about the $25 million level, outrank most Fortune 500 companies" [HEAD81b].

Besides being much larger than corresponding budgets in private industry, the Federal ADP budget differs from them in the way it is divided among activities. Figure 1 shows percentages for various categories of expenditures in 1980 for both the Federal Government and the private sector. The most significant difference between Federal and private sector spending patterns is that the percentage of Federal expenditures for outside services is more than triple the percentage of expenditures in the private sector. In comparison with private companies, Federal agencies thus have a smaller percentage of their budgets to spend for hardware and personnel. Based on OMB projections through 1985, these differences will continue to grow as the Federal workforce stabilizes or shrinks and expenditures for commercial systems analysis and programming services increase [OMB80]. In each of the first three categories, the differences between public and private spending percentages correspond to differences in ADP constraints and practices. Each of the following sections addresses such differences in one of the given categories.

2.1.1 Personnel. The Federal data processing workforce has grown only slightly since 1973. According to OMB projections, it will continue to remain stable at about 120 thousand through 1985 [OMB80]. Published accounts of current policy suggest that the workforce may even decline. In 1973 computing personnel accounted for 54 percent of the total data processing budget, whereas by 1981 the percentage will drop to about 38 percent [OMB80]. The Federal trend sharply contrasts with the large increase in the number of ADP personnel nationwide, which, according to the Bureau of Labor Statistics, is growing at about 8 percent per year [BLS79]. The OMB attributes this difference to the Federal Government's relying heavily on the private sector for computing services, including systems analysis and design,

| CATEGORY | FEDERAL $6.0 billion | PRIVATE $66.6 billion |
|---|---|---|
| Personnel | 38.3% | 46.9% |
| Hardware | 22.2% | 33.7% |
| Outside Services | 30.4% | 9.6% |
| Supplies and Overhead | 9.1% | 9.8% |
| TOTAL | 100.0% | 100.0% |

Source: [OMB80] and [IDC81a]

Figure 1. Distribution of ADP Costs 1981

| CATEGORY | STAFF-YEARS | PERCENTAGE |
|---|---|---|
| Systems Analysis and Design | 16,856 | 13 |
| Programming | 18,999 | 15 |
| Equipment Operation and Maintenance | 34,024 | 27 |
| Key Punch | 16,770 | 13 |
| Services and Support | 39,193 | 32 |
| Total | 125,842 | 100 |

Source: [GSA79b]

Figure 2. ADP Staff-years by Category (FY1978)

software development, hardware maintenance, and facilities management [OMB80].

Using GSA data on annual work-year equivalents in various data processing categories, Figure 2 shows the breakdown of the Federal DP workforce in FY78. The category for services and support includes management support, planning, selection and procurement of hardware and software, and all data entry other than key punch. Because GSA no longer collects data on staff activities, the only available update is the total number of ADP work-years, which was 119,835 in FY79 and 121,327 in FY80 [GSA80b].

Between 1977 and 1979 the total number of Federal computer programmers and systems analysts declined slightly. During this time there was a 5 percent drop in Federal computer specialists compared with a 43 percent increase nationwide [BLS79]. Part of this trend can be accounted for by the increase in expenditures for computer services: computer programmers and systems analysts account for more than 36 percent of the total Federal dollars spent on commercial services between 1978 and 1981 [OMB80].

If demands for computer-related services continue to grow as OMB has projected, the probable decline of the Federal ADP workforce means that in the future agenices will spend more for outside services. Federal managers face personnel ceilings, hiring restrictions, and salary regulations that do not apply to private industry, and they may lose their best staff to private firms who actually develop the applications that the Government needs.

2.1.2 Hardware. From 1977 through 1979 the Federal Government spent an average of $553 million per year to acquire computing equipment. By the end of 1979 the Government's total investment in computer hardware amounted to almost $5.4 billion distributed among more than 10,000 computing systems [GSA80a]. A configuration of ADP equipment including one or more CPU's, a computer system may range from small minicomputer systems used for special purposes to large general data processing systems with huge mass memory. In recent years the greater part of the 11 percent annual average increase in the Federal computer inventory has come in the Special Management classification, perhaps because minicomputer technology has made inexpensive computing equipment available to organizational units that could neither afford nor support large mainframes in the General Management Classification.

The percentage of total Federal ADP expenditures that go for new hardware (22.2%) is significantly lower than the corresponding percentage in private industry (33.7%). As a consequence, the inventory of Federal computers is relatively old. Robert V. Head, writing in Government Executive, reported that "the average age of all computer systems in the Federal Government is over 7 years" [HEAD81a]. Because computer technology advances so rapidly, many Government agencies rely on computers that are now obsolete. These agencies cannot take advantage of newer software designed to run on modern hardware. Though the Government faces some of the most challenging of modern computing problems, it cannot always tackle those problems with the best technology available to private industry.

2.1.3 Outside Services. The outside services category includes line charges for timesharing, direct expenses for software purchase, and expenditures for all outside consulting, development, or maintenance support. The Office of Management and Budget considers direct purchase of software packages as a separate line item under capital investment, but expenditures under this category are currently only about 1 percent of the total ADP budget. Figure 3 categorizes Federal expenditures for commercial services other than direct purchase of software packages. Executive branch expenditures for commercial services have increased from $380 million or 16.5 percent of the ADP budget in 1973 to an estimated $1.76 billion or 28 percent of the budget in 1981 [OMB80]. Over the same period the percentage of private industry budgets that has gone for outside services has remained relatively stable at 8-10 percent [IDC73, IDC81a]. Since the Federal workforce is projected to decrease, the percentage of Federal ADP budgets spent on commercial services will probably grow through 1985.

## 2.2 Software Development

Figure 2 shows that 28 percent or nearly 34 thousand persons in the Federal data processing workforce are systems analysts or computer programmers. Figure 1 shows that the Federal Government spends 38.3 percent of $6.0 billion, or about $2.3 billion, on personnel. If the average salary of systems analysts and programmers is equal to or higher than the average salary of all ADP personnel, then 28 percent of $2.3 billion--or $644 million--represents a minimum Federal personnel cost for in-house programming and systems analysis. Combined with the $606 million the Government spends on commercial programming and analysis (see Figure 3), this figure gives a total of at least $1250 million. Hardware expenditures almost certainly represent

| CATEGORY | EXPENDITURES (in $M) | PERCENTAGE |
|---|---|---|
| Hardware Maintenance | 307 | 17 |
| Systems Analysis and Programming | 606 | 35 |
| Timesharing Services | 348 | 20 |
| Facilities Management | 285 | 16 |
| Management Studies | 216 | 12 |
| Total | 1762 | 100 |

Source: [OMB80]

Figure 3. Projected Commercial Services Expenditures (FY 1981)

less than 20 percent of the total, leaving about $1 billion per year for development and maintenance of application and system software.

## 2.3 Potential Impact of DBMS's

Because of the way the Government spends its ADP budget, database management systems could become particularly important to Federal agencies. The size of the Government's expenditures for software development and maintenance emphasizes the need for innovative technology to cut costs and increase productivity. In many cases DBMS's can boost the productivity of in-house personnel and help hold down the rate of increase in contracted services. In a database environment Federal computer specialists can maintain and develop more of their own application programs. In cases where agencies procure outside help to develop applications on their own computers, a database environment should reduce maintenance costs. Initial development costs for specialized applications may run higher than similar efforts in a traditional environment, but the DBMS will allow more generalized development and should reduce costs for subsequent applications that merely require minor modifications of an existing program or system. And while people usually think of a DBMS in relation to large databases supported by powerful mainframes, many of the Government's smaller, special-purpose computer applications could also benefit from a common database managed by specialized software.

## 2.4 DBMS Software Availability

The availability of DBMS packages to solve Federal ADP problems depends in part on hardware environments, and the Federal mix of hardware by manufacturer is considerably different from that in the private sector. In the private market IBM is by far the dominant hardware vendor (see Figure 4). In 1978, the last year for which International Data Corporation published such data, IBM recorded 53 percent of all general purpose and minicomputer sales by dollar value, with no other manufacturer posting even 10 percent.

In the Federal market the mixture of equipment is considerably different. IBM is still the dominant vendor, but CDC, Univac, and Honeywell are significant competitors (see Figure 5). Figures 4 and 5 are not directly comparable: the former is based on total 1978 shipments, while the latter is based on the total value of installed computers. Nevertheless, IBM hardware is clearly much less

| VENDOR | DOLLAR VALUE (in $M) | MARKET SHARE |
|---|---|---|
| IBM | 9605 | 53% |
| DEC | 1610 | 9% |
| Honeywell | 1020 | 6% |
| Univac | 988 | 6% |
| Burroughs | 750 | 4% |
| CDC | 340 | 2% |
| Other | 3637 | 20% |
| Total | 17,950 | 100% |

Source: [IDC79]

Figure 4. Dollar Value of 1978 Shipments
General Purpose and Minicomputers

| VENDOR | DOLLAR VALUE (in $M) | MARKET SHARE |
|---|---|---|
| IBM | 544 | 28% |
| CDC | 374 | 19% |
| Univac | 219 | 11% |
| Honeywell | 203 | 11% |
| DEC | 145 | 8% |
| Burroughs | 45 | 2% |
| Other | 413 | 21% |
| Total | 1943 | 100% |

Source: [GSA80a]

Figure 5. Federal Installed Base of CPU's
(By Dollar Value)

prevalent in the Federal Government than in private industry. As a consequence, Federal ADP managers who want to procure a DBMS generally have a smaller selection than do their counterparts in the private sector. According to Datapro Software, an industry buyer's guide, 8 of 18 major database management systems run only on IBM equipment, and another 7 run on only one vendor's hardware [DATA80]. In some cases the Federal manager may have to choose between buying the DBMS sold by the agency's hardware manufacturer or giving up altogether the potential advantages of a database environment.

## 2.5 Federal Inventory of DBMS Software

Determining how many DBMS packages the Federal Government now owns is not easy. No direct count of systems is available, and estimates from indirect sources vary from 491 [IDC81b] to 3720 [GA079] or even higher. Licensing arrangements between commercial vendors and the Government may be complex and varied: an "installation" can mean one thing to one vendor or department and something else to another. Finally, ownership of a DBMS does not guarantee a consistent level of usage, particularly in situations where the DBMS came "bundled" with the vendor's hardware.

In several cases commercial vendors have supplied NBS with estimates of DBMS installations for the Government as a whole. According to Computer Corporation of America, there are between 35 and 40 Federal installations of Model 204. IBM reports a similar number of IMS or DL/1 installations, and Cullinane says there are about 60-70 Federal shops with IDMS. INTEL reports between 110 and 120 installations of System 2000, Burroughs reports 20-25 Government installations of DMS-II on large mainframes, and Honeywell says that it has 6 IDS-I accounts and 12 DM-IV accounts. One of these DM-IV accounts covers installations at 35 separate sites. This list is far from complete, but it does indicate the extent of DBMS usage by Federal agencies.

For the purpose of determining the Government's actual investment in database management, qualitative factors are probably more important than simple counts of products. As subsequent chapters explain, the cost of a DBMS itself may be small compared to the cost of converting an entire ADP shop into a database environment in which users may reap the greatest benefits of the database approach to data processing. According to the GAO, total costs may run as much as $1-2 million or higher [GA079]. If figures like these apply even to 10 percent of Government DBMS installations, the size of the Federal investment in DBMS technology warrants scrutiny of DBMS costs and benefits.

## 2.6 Database on Timesharing Services

Not all Federal agencies that use a DBMS own or lease it. Some buy time on a system through commercial timesharing services, which came into existence in the mid 1960's and have since grown into a nearly $3 billion per year industry. The Federal Government has long been a major user of these services, with an estimated executive branch expenditure in FY 1981 of $348 million, or almost 10 percent of all ADP expenditures reaching the private sector [OMB80]. Federal use of timesharing services has increased by more than 14 percent annually since 1978 and will probably continue to be a significant portion of the Federal ADP budget.

| YEAR | EXPENDITURES (in $M) |
|---|---|
| FY 78 | 234 |
| FY 79 | 271 |
| FY 80 (EST) | 325 |
| FY 81 (EST) | 348 |

Source: [OMB80]

Figure 6. Commercial Timesharing Expenditures Federal Executive Branch

Timesharing services originally sold computer time only with access to several general-purpose programming languages, but most services have significantly expanded their software offerings to include accounting systems, statistical and scientific packages, bookkeeping, payroll, inventory, and generalized database management systems. In a recent Datapro 70 report more than 74 percent of the 125 timesharing services listed offered at least one database management system [DATA79]. According to the Datapro report approximately 30 percent of commercial timesharing customers use the service to maintain a database. This figure is consistent with an IDC survey of the ten largest

timesharing services (CDC, CSC, ADP, EDS, GE, SDC, McAuto, Tymshare, PRC, Bradford), all with over $100 million in revenue annually. The IDC report shows that 25 percent of all interactive use of timesharing services is for database inquiry [IDC79].

Several database software vendors offer their products through a large number of services. For example, System 2000 from Intel Corporation is available on over 30 different timesharing services. The largest timesharing services--like General Electric with DMS-III, NCSS with NOMAD, and Computer Sciences Corporation with MANAGE--have developed their own database management systems fully integrated with their other software offerings. Those that use IBM hardware and software will normally make available IMS and one or more other systems designed to run on IBM equipment. Figure 7 lists some of the major vendors of timesharing services, the type of hardware they employ, and the database management systems that they offer to users. Some of the services also maintain large databases that are of general interest to a wide variety of users. For example, Dialcom maintains Fabers, a Municipal Information Database originally developed for the Department of Agriculture.

Here again it is difficult to gauge the size of the Federal investment in DBMS's. INTEL says that the Government spends $15-20 million per year to buy time on System 2000. Since this figure includes only one DBMS, and since it does not include personnel costs, we may conclude that the Government's investment in timeshared databases is a significant addition to its expenditures on systems that agencies own or lease.

## 2.7 Potential for DBMS Growth in Government

Growth in Federal use of DBMS's depends upon several factors. The most important is the desire on the part of data processing managers to implement the database approach to data processing. Such a desire should come only as the result of a positive cost effectiveness study on the advantages of the database approach over a more traditional file approach. Chapter 3 discusses important considerations in determining this cost effectiveness. Physical factors are also important in the potential for growth in Federal usage of DBMS's. Federal agencies must have adequate computing power and secondary storage capability to support increased use of DBMS software, and DBMS packages that run on Federal hardware must be available in the marketplace. Industry trends point to substantially increased use of DBMS's in the coming years because both computing power and

| SERVICE NAME | MAJOR HARDWARE | DBMS PACKAGES |
|---|---|---|
| ADP Network Timesharing Services | DEC | DBMS-10 |
| Control Data Corp. Cybernet | CDC | System 2000<br>Total<br>IPF |
| Computer Sciences Corp. Infonet | UNIVAC<br>IBM | MANAGE<br>System 2000 |
| Dialcom | Prime<br>Honeywell | INFO |
| General Electric | Honeywell<br>IBM | DMS-III<br>System 2000<br>Mark IV<br>BASIS<br>Inquire |
| McDonnell Douglas | IBM | IMS<br>System 2000 |
| NCSS | IBM<br>Amdahl | NOMAD<br>Mark-IV<br>Ramis |
| TYMSHARE | Xerox<br>DEC<br>IBM | System 2000 |

Figure 7. DBMS Availability on Some Commercial Timesharing Services

secondary storage are becoming cheaper and more available and because more commercial DBMS packages are available in the marketplace. The following paragraphs address the impact of these trends in the Federal domain.

2.7.1 Computing Power. Although people tend to associate database management systems with powerful mainframes, one cannot eliminate even special-purpose minicomputers from potentially supporting a DBMS. Databases need not be particularly large before generalized access and centralized control are desirable. For example, a small mailing list database may contain names, company affiliations, membership in organizations, and addresses of each. With the use of a DBMS, numerous users may retrieve the exact list of names satisfying complex combinations of membership and affiliation. NBS research has shown that a complete state-of-the-art database management system can run on a microcomputer system having 256K bytes of main memory and 10 megabytes of disk storage with total system cost less than $30 thousand. Although exact data is not available, NBS estimates that a majority of existing Federal computer systems have these capabilities and thus could support a DBMS if such a package were available. In addition, most new CPU's purchased for use with a disk system will probably be capable of supporting a DBMS.

2.7.2 On-line Secondary Storage. On-line secondary storage units in the Federal Government have been increasing at approximately 15 percent per year and in FY 1979 numbered nearly 24 thousand, with a total dollar value greater than $657 million. Figure 8 shows the growth in both number and dollar value from FY 1975 to FY 1979.

| YEAR | # DEVICES | GROWTH | $-VALUE | GROWTH |
|---|---|---|---|---|
| FY 1975 | 13,479 | | $502 M | |
| FY 1976 | 15,676 | 16% | $512 M | 2% |
| FY 1977 | 18,302 | 17% | $592 M | 16% |
| FY 1978 | 20,899 | 14% | $601 M | 2% |
| FY 1979 | 23,687 | 13% | $657 M | 9% |

Source: [GSA79c]

Figure 8. On-line Secondary Storage Units

On-line storage units include conventional disks and drums as well as huge mechanically automated [illegible] capable of storing over a million megabytes of data. Since the latter use a conventional disk as an intermediate buffer, they are actually tertiary storage, and we do not consider them here. By far the predominant form of on-line secondary storage is conventional disks, which vary considerably in size, ranging from inexpensive floppy disks to large conventional drives in the 400-800 megabyte range costing approximately $60 thousand apiece. Estimates from industry show that typical large disk systems cost about $56 per megabyte to purchase [THEI78] and approximately $1.12 per megabyte per month to lease. Many industry experts expect these figures to drop dramatically in the next 5 to 10 years as vendors begin to market mass data storage units that take advantage of new technologies for fast retrieval.

Because of these falling prices and increased density per storage unit, total disk capacity in the United States is increasing dramatically. In 1978, IDC estimated that U.S. disk capacity was growing at 24 percent annually and would reach over 45 million megabytes by 1980 [IDC78]. Assuming that the Federal Government owns or leases about 9 percent of this total, NBS estimates that the Government's total disk capacity is about 4 million megabytes or an average of over 400 megabytes per computer system.

The increased availability of on-line secondary storage has produced tremendous growth in the development of on-line information systems. The Quantum Science Corporation estimates that in 1974 only 50 percent of installed information systems were on-line, whereas by 1980, the figure had grown to 74 percent. This trend demonstrates that the demand for on-line data management is increasing substantially. The versatility of comprehensive DBMS's will support this growth.

## 3. COST AND BENEFIT FACTORS

The literature on database management includes numerous taxonomies of DBMS costs and benefits [DATE77, MART75, SIRC78, GA079]. Most of these costs and benefits are theoretical and therefore general enough to apply to a number of different situations. In synthesizing information from textbooks and articles, the following discusssion identifies the major advantages and disadvantages to consider when deciding whether or not to implement the database approach.

## 3.1 Benefits

Virtually every benefit that a DBMS offers represents increased functionality or flexibility rather than reduced costs. Like many other innovators of major software systems, DBMS developers have in general wanted to enable users to accomplish things that they could not do or would not attempt in traditional environments. A sophisticated software system like a DBMS costs money and may require very highly trained ADP personnel, but the benefits it offers can provide substantial improvements in the operations of a data processing unit.

3.1.1 Centralized Data Management. The fundamental assumption of the database approach is the importance of data as a corporate asset. Any organization that wants to protect its investment in this resource should consider the benefits of centralized control of data through data management. Traditional approaches to file processing let individual programmers define, structure, and duplicate data for particular applications. The database approach requires centralized control of data so that different applications can share information.

3.1.2 Reduced Redundancy. Because applications share data, an organization using the database approach requires fewer files of redundant data and less effort to control data integrity, to maintain data consistency, and to enforce data standards. The organization can define validation procedures to control the insertion and update of database entities. A single update makes new data immediately available to all applications that share the data.

3.1.3 Data Independence. The phrase "data independence" refers to the separation of data from programs. This separation usually results from the definition of three separate views of the data [MART75]:

1. The physical representation,
2. The overall database logical representation, and
3. The logical representations for individual application programs.

A DBMS handles two interfaces: one between the logical and physical representations of the whole database, and a second between the logical representation of the database and the external user views of that representation. The physical representation and access mechanisms for stored data remain

invisible to the user and to individual application programs.

3.1.4 External User Views. A typical database management system allows the Database Administrator (DBA) to tailor a logical view of data necessary to each application program. For example, a user from the personnel department may regard the database as a collection of department records plus a collection of employee records, while a user in the purchasing department sees only department, supplier, and part records. Furthermore, these two users might have different restrictions on their use of the data they can access. One may be able only to retrieve data, while the other can retrieve, update, and delete particular records.

Increased flexibility to meet new user demands is the key benefit of external user views. Development and maintenance of application programs require less effort in a database environment than in a traditional setting, although application programmers may need more sophisticated knowledge of data structures and operations to use the database as effectively as they used conventional data files. External logical views increase a system's responsiveness to changes in user demands, reduce the time to develop applications, and insulate applications from changes in hardware and support software.

3.1.5 Logical and Physical Storage Structures. In addition to external user views, generalized database management systems often provide independent logical and physical storage structures for the entire database. A DBMS maintains generalized logical access paths that insulate record logic from data storage; consequently, users can peruse the database without knowing anything about its indices, pointers, chains, and other means of physically locating records.

Future DBMS's may offer what is now only a potential benefit of DBMS technology--the freedom to change hardware and operating systems without incurring large conversion costs for application programs. In the database environment applications depend only on logical structures provided by the DBMS, but most available DBMS packages run in only a limited range of environments; statistics in the previous chapter show that many require a single specific line of hardware. Users may have to wait for a DBMS that provides flexibility in the choice of computer equipment and support software, although they can already realize the benefits of flexible data structures.

3.1.6 On-line Processing and Ad-hoc Queries. While not necessarily a part of the database environment, on-line processing of ad hoc queries is one of the major benefits of some DBMS products. With some training in such DBMS features as query languages and report generators, users without programming experience can answer structured questions and prepare reports in minutes. In a traditional environment these operations would take hours, days, or even weeks. Ease of access to data increases the rate at which users can perform transactions, especially those that use a variety of structured data collections to produce special reports or one-time analyses [SIRC78].

3.1.7 Access and Integrity Control. With some database management systems the DBA can control logical access and check integrity for the entire database, specific files, relations or realms, and specific data elements or attributes. More advanced systems may be able to differentiate among records or tuples according to the contents of data fields: for instance, mid-level managers may be able to retrieve personnel records for all those under their supervision but not for their superiors. At the external level, some users may be able only to retrieve, some to retrieve and modify, some to create and delete data, and some to restructure and define databases. The DBMS may allow the DBA to define validation procedures to make sure that all storage transactions comply with defined data characteristics. In a traditional file-processing environment these capabilities are either impossible or very costly.

## 3.2 Costs

Though a database management system can increase the productivity of ADP applications and services, it may not suit the purposes of every data processing shop. Not every Federal agency has to do the things that DBMS's do best, and even those agencies that can benefit from a particular feature or two may find cheaper, less sophisticated software systems to meet their needs. Before purchasing a DBMS or implementing a database environment, the ADP manager should consider carefully the potential costs of converting from traditional file processing to database management.

3.2.1 Vulnerability and User Conflicts. Centralizing data and eliminating redundant files necessarily increases an organization's vulnerability to system crashes. Backup and recovery of a complex database are more difficult than similar processes for traditional files. Furthermore, unauthorized personnel can gain access to restricted data unless the DBA implements complex and costly procedures for

controlling access and preserving security. Different users within an organization may disagree about the best structures and controls for shared data, and resulting conflicts can have political repercussions [SIRC78].

3.2.2 Performance. To maintain external user views, and to provide indices and multiple access paths, a DBMS will often require significantly more computing capacity, main memory, and secondary storage than conventional files. Security and integrity controls add to the burden on machine resources. In fact, a DBMS can place so much strain on existing hardware and system software that the user may have to acquire more processing power just to support the DBMS [GA079].

3.2.3 Personnel Skills. While a DBMS may allow inexperienced personnel to retrieve data and generate reports, it can require greater sophistication for system support and application programming. In order to develop and maintain new application systems, programmers need to know data structures and operations for the data model as well as for the host language that communicates with the database. Even end users may have to develop skill in using complex Boolean languages. The agency will need to train many of its personnel in database methods, and it may have to hire additional computer specialists to handle the database management system itself.

3.2.4 Overhead and Initial Costs. Though each of these disadvantages can cause problems for an organization, the most significant disadvantage of a database environment is its cost. An ADP unit that fully implements the database approach incurs a significant start-up cost long before it begins to benefit from DBMS technology. Maintaining the DBMS itself and implementing the database administrator functions necessary to carry out the DBMS concept impose a substantial overhead cost. Furthermore, the cost of simply converting existing application programs so that they can use the database can be prohibitive. One recent report hypothesizes that implementing a database environment shifts costs from operation and maintenance toward the beginning of the system's life cycle [DEUT78, p. 63]. An agency could forego some of these early planning costs, but only at great risk. Federal ADP units that acquired and implemented database management systems without requirements studies and cost/benefit analyses sometimes wasted $1-2 million, while an agency that carefully planned its conversion lost almost nothing and successfully began database operations only 2 months behind schedule [GA079]. Such a contrast in cost effectiveness demonstrates the importance of studying users' requirements to see if database management is the best way to meet users' needs and accomplish agency missions.

## 3.3 Summary

The decision to convert from traditional file processing to database management should come only after careful weighing of the costs and benefits of a DBMS. Almost everything one gains by using a DBMS entails a corresponding cost. Centralizing data increases an organization's control over one of its major assets, but makes that asset more vulnerable to failure and misuse. Data independence insulates application programs from changes in the physical representation of data, and thereby removes the option of structuring data for efficient retrieval. A DBMS enables the user to derive more kinds of information from his data, but to do so it requires increased computing resources. Finally, a DBMS can increase system functionality and reduce maintenance costs, but its immediate effect is higher costs for overhead, acquisition, implementation, and conversion. No agency should commit itself to a database environment without making a detailed cost/benefit analysis of the advantages and disadvantages of the database approach to the agency's data processing problems.

# 4. INTERVIEWS WITH FEDERAL AGENCIES

In order to determine how well database management systems are serving the Federal Government, NBS interviewed data managers from eight Federal agencies. The participants contributed a significant amount of their valuable time to respond to our questions and to share their experience and advice with other Federal agencies. Each manager filled out a questionnaire and participated in a subsequent interview. In addition, two other managers filled out questionnaires to supply procurement information that the interviewee could not provide for his particular agency. Both the questionnaire and the interview focused on the agency's requirements for data management and the costs and benefits of DBMS's with respect to those requirements. Most of the questions from the questionnaire required qualitative answers like "little or none," "some," "much," and "very much." Since it was impractical to try to find a few agencies that represented the needs of the entire Federal community, NBS sought instead to talk with agencies using a variety of DBMS products to solve diverse application problems.

Drawing upon the results of the questionnaire and interviews, this chapter discusses the extent to which agencies study their requirements, the ways they select particular products, and the actual costs and benefits they have experienced in using DBMS's. Throughout the discussion we try to draw general conclusions while acknowledging both the important differences among Federal agencies and the limits of the sample. The next section briefly summarizes the kinds of applications that have led agencies to acquire and use database management systems. Readers who want more explicit details about the interview process or the responses of particular agencies may wish to consult Appendixes A, B, and C.

## 4.1 Applications

Two of the participating agencies use DBMS's to manage data for a number of administrative programs. In both cases the data structuring abilities of the DBMS have enabled the database administrators to model many-to-many relationships between record types. The results have been complex and centrally controlled administrative systems combining data from several organizational units. These two systems manipulate data through commands embedded in host languages like COBOL and PL/1. Neither requires high-level query facilities, and both rely primarily on batch processing of compiled programs.

Six other agencies use DBMS's in both batch and interactive modes to keep track of frequently updated records. One of the six monitors water resources across the nation, handling 50,000 updates and about 300 queries per week. Database management systems help another four agencies keep records of their own activities and respond quickly to changes in the industries they regulate. The last agency uses its DBMS to protect the Government against fraud in the direct distribution of money. Interactive processing of predefined queries enables the agency to fulfill its legislated mission with minimal danger of fraud.

## 4.2 Requirements and Cost/Benefit Analyses

Before acquiring a DBMS, most of the interviewed agencies tried to identify the requirements of a software package to handle some of the agency's data. All 10 respondents reported doing a requirements study, and 6 said they consulted users "much" or "very much" in the process. Such consultations were not always fruitful. In one case users did not understand computers well enough to define what software they needed, even though they thoroughly

understood the job that was being automated. At another agency the interviewee felt that users did not want to take the time to define application requirements; they wanted the ADP staff not only to provide data services, but also to help them analyze their own requirements. Problems like these, coupled with management pressure to provide needed services quickly, contributed to the general difficulty of defining requirements for computer software.

Nine of the respondents said they had done a cost/benefit analysis to determine if a DBMS was the best way to meet their requirements, but only three of the analyses were formal. A few managers based their decisions on their own experience with software tools and understanding of agency resources and needs. At least a couple thought that formal cost/benefit analysis satisfied procurement regulations without substantially increasing the likelihood of success with a purchased product. These managers wanted to have to justify their decisions only to their supervisors, not to procurement experts as well. In some cases database management systems were perhaps the only available products that provided the complex data structures or rapid access and update capabilities that these people considered essential. However they may have selected a DBMS product, the interviewees have been pleased with their systems and feel that they analyzed costs and benefits as much as was necessary. Nonetheless, it is impossible to tell whether a manager's satisfaction with a DBMS stems from its being the best product available or simply from its superiority to the previous system or method for performing the task.

Most of the interviewees had at least some trouble identifying costs and benefits of a DBMS, but nobody complained of "very much" difficulty. Trying to quantify intangible costs and benefits was harder, with 4 of the 10 respondents rating the difficulty as "much" or "very much." It is probably significant that the agency acknowledging the greatest difficulty in these two tasks was the one that performed the most extensive cost/benefit analysis. Agencies that found the process easy may have overlooked or underestimated some of the indirect or intangible costs and benefits, including those related to reorganization of staff, conversion of programs, productivity of staff, and flexibility of the DBMS. Four of the ten respondents said that a guideline on cost/benefit methodology would help them "much" or "very much" in future DBMS acquisitions, but only two thought that the cost/benefit analysis they had done had helped them "much" in their plans to select, acquire, and install the DBMS they have now.

Despite the difficulties in determining the cost effectiveness of a DBMS, some cost/benefit analysis is probably necessary in almost all cases. One of the interviewed agencies acquired its DBMS to provide an access method and some data independence for a particular application. Because the DBMS did not do everything it advertised, the systems programmer who performed an "intuitive" cost/benefit analysis ultimately had to write a special program to do the job. Such a problem might have occurred even after a thorough cost/benefit analysis, but it had a better chance of escaping detection in an informal study. Fortunately for the agency, generalized database management systems are flexible enough that subsequent staff members have been able to use the agency's DBMS to develop several other applications. The DBMS has increased the responsiveness and productivity of the staff in developing and maintaining application programs for users throughout the organization. Other agencies might not always be so lucky.

Although all agencies should probably analyze costs and benefits before acquiring a DBMS, the depth of the analysis does not have to be uniform. As the previous section showed, Federal agencies use DBMS's in different ways. Those agencies that want a sophisticated access method for a single application may have well-defined costs and benefits that make an analysis straightforward. Other Government organizations, which plan to centralize data across a number of applications and to control this data with a DBMS, risk more serious losses and therefore need correspondingly more extensive cost/benefit analyses to ensure that a DBMS is justified. Initial costs for developing a centralized data management system around a DBMS may run as high as $1-2 million, and development time can easily be 2 or 3 calendar years. When considering such a large commitment to a particular class of software products, an agency needs a full-scale analysis to identify and quantify all the potential costs and benefits of various ways to solve the given problem. In these cases the generalized capabilities of a DBMS are not likely to compensate for a mistake in procurement or development.

## 4.3 DBMS Selection

For the purposes of discussion we have separated DBMS selection from requirements studies and cost/benefit analyses. Theoretically, an agency would first define its requirements, then perform a cost/benefit analysis to determine if a DBMS was the best method to meet those requirements, and only then evaluate the features of commercially available products. In practice these

distinctions are less clear. Because DBMS products vary considerably, the features of particular systems may play a role in cost/benefit analysis. For example, a high-level query facility might help both to justify a DBMS and to eliminate some products from consideration. Several of the managers we interviewed had particular products already in mind when they compared the costs and benefits of a commercial DBMS to those of other software packages.

Although 8 of the 10 respondents evaluated in detail the features of commerical DBMS products, only 3 of the interviewed agencies actually issued a Request for Proposals (RFP) as part of the procurement process. Two of these did not have a particular DBMS in mind, and both developed a benchmark and used it to evaluate potential systems. The other agencies compared systems -- in some cases formally and extensively -- before justifying a sole-source procurement of the product they wanted. Though the sample of agencies was too small to draw many conclusions about the way Federal agencies select DBMS's, the basic procedure seems to involve some kind of requirements study and an effort to match the features of particular products against those requirements. This kind of comparison is minimal; anything less would probably not be in an agency's best interest. To take full advantage of the benefits offered by a generalized database management system, an organization must make sure that the product provides essential features. In one case a benchmark would probably have kept the agency from buying a product that did not do everything it advertised and hence could not do the job for which it was acquired.

## 4.4 Benefits

Virtually every major benefit of a DBMS represents increased functionality or productivity rather than reduced cost. Which benefits an agency actually receives depends in part on both the particular product and the way the organization uses it. Systems that offer complex network data structures enable agencies to reduce data redundancy and to control data access and integrity. Federal agencies that use such systems for centralized data management are likely to find that their application programs provide more timely and accurate information while being less sensitive to changes in physical storage structures. Other DBMS's, which feature rapid access and high-level query facilities or user languages, shorten the time for developing applications and enable nonprogrammers to use computers effectively. Some products and applications may offer agencies both kinds of benefits, although it is unlikely that a single agency will receive all of the benefits

described in the rest of this section.

4.4.1 Centralized Data Management. Combining redundant data files and regarding data as an organizational asset are two prime theoretical advantages of a DBMS. In practice such wholesale integration can be politically difficult, but even a lesser degree of centralization can provide important benefits. Two of the interviewed agencies have managed to overcome their users' reluctance to give up control of data. Both agencies listed reduced data redundancy and improved control of data as major advantages of using a DBMS. One stressed the improved timeliness and accuracy of data that result from "point-of-transaction data capture." As soon as one user adds or updates data, the new values are immediately available to all other programs having access to the updated records. Such timeliness is as important in a batch environment as it is in on-line processing, and it is feasible primarily because there are not multiple copies of the same data. The same agency that cited point-of-transaction data capture also praised the security, recovery, and backup procedures provided by the DBMS. Such procedures provide tight control of agency data while minimizing the danger of putting all organizational data in a single place.

Both of the agencies reporting large-scale integration of data files have DBMS's that support complex data structures and that emphasize record-at-a-time processing by application programs written in languages like COBOL and PL/1. Other kinds of DBMS's also enable an organization to centralize data, but only two of the other interviewed agencies have actually used their DBMS's to integrate files from more than one application area. The others acquired their DBMS's as sophisticated access methods for particular applications and continue to use them in that way. Two interviewees had high praise for relational DBMS's, and one explicitly stated that a relational system with the ability to define different "views" of the data would probably eliminate many of the users' conflicts over control of shared data. If he is right, relational DBMS's may significantly increase the number of Federal agencies that make a DBMS the focal point of their data processing.

Combining data files from several applications is not the only kind of centralization that a DBMS encourages. Three of the interviewed agencies use their systems to handle data from field offices across the nation. Geographic centralization enables the agencies to coordinate regional activities from a single national office. Though other kinds of software packages might provide similar capabilities, the high-level query and update facilities provided by the DBMS's make geographic centralization

practical and effective. The managers we interviewed at each of these three agencies recognized the benefits of integrating data across applications, but also acknowledged the practical and technical difficulties of further centralization.

4.4.2 Data Independence. Besides reducing redundancy and centralizing control of data, DBMS's separate logical data structures from physical storage structures and hence isolate application programs from changes in data storage. Seven of the ten respondents to the questionnaire reported that data independence reduced by "much" or "very much" the effort needed to design and program new applications. Four people named easier application development as a major advantage of a DBMS, though part of their enthusiasm probably comes from powerful user languages that take advantage of the distinction between logical and physical data structures. Another respondent praised the flexibility that results from data independence.

According to our sample, data independence has a significant impact on program maintenance and modification as well as design and development. Six of the respondents reported "much" or "very much" easier maintenance with a DBMS than with conventional files. Four people named easier modification of application software as a major benefit of a DBMS. However, the application manager for one of the large, centralized administrative systems dissented, stating that the "maintenance effort for daily systems which are undergoing regular modification is about the same" for systems with or without a DBMS. The most likely explanation for this difference in experience is that the dissenter probably has stable physical storage structures for application data. The advantage of data independence is isolating programs from changes in the physical storage of data. If physical storage structures are static, the programmer who maintains an application system is unlikely to experience many benefits from data independence.

4.4.3 High-Level Query Facilities. A number of commercial DBMS's provide high-level query facilities designed to provide quick responses to on-line queries about the database. One of the agencies with centralized administrative systems expressed no interest in queries, but seven of the other nine respondents thought that the ability to ask ad hoc queries had significantly improved ("much" or "very much") the timeliness and accuracy of information available to managers. For another agency the on-line query facility is absolutely critical to the major application of the DBMS, but the queries are all predefined rather than ad hoc. Without the ability to process these predefined queries, the agency would not be able to accomplish its

legislated mission without risking fraudulent claims and significant losses to the Government.

Whereas ad hoc query facilities have provided Government managers with useful information, they have not been so successful in reducing the need for custom programs. Only four of the respondents said that they had experienced "much" reduction in this area. As one person put it, the query facility "answers questions but does not produce extensive reports." The responses to this question suggest that the need for custom programs is very different from the need for answers to particular questions. Managers who did not have access to a DBMS with a query language may have simply done without the information that they now obtain through the query facility. If so, this is another case where the benefit of a DBMS is increased functionality rather than reduced cost.

4.4.4 High-Level User Languages. The questionnaire did not address the issue of languages, but five of the respondents stressed the importance of high-level user languages provided by the DBMS vendor specifically for use with a database. These languages are more functional than query facilities, but not so complex as conventional programming languages. One manager said that he could produce an 1800 page report with only four instructions in the language of his report writer. Another asserted that the user language provided by his DBMS provided a 90 percent savings in code and a 75 percent savings in effort relative to comparable COBOL programs. Three of the interviewees called these languages major benefits of a DBMS because they enable nonprogrammers to retrieve information and produce reports that were previously available only through trained programmers. One manager pointed out that user languages help users to understand their data requirements and to describe their problems when they need advice from the ADP staff. For managers who have difficulty hiring programmers or who want independence from ADP shops, such languages are especially useful. A good user language can reduce the time for developing applications and increase both the productivity and the responsiveness of agency staff. Moreover, several DBMS's have active users' groups that provide specific advice for user languages as well as general support for the DBMS. Four of the interviewees specifically mentioned their users' groups as major benefits of their DBMS's.

## 4.5 Costs

The primary costs of a database management system come from three sources. At the beginning of the DBMS life cycle Federal agencies have faced one-time costs of procurement, changes in hardware and software environments, reorganization of the ADP staff, and conversion of existing application files and programs. Throughout the life cycle they have experienced recurring costs for both computer resources and skilled technical personnel. Like the benefits, the costs of a DBMS vary widely and depend in part on both the particular product and the way the agency uses it.

4.5.1 Initial Costs. The initial costs of implementing a database management system directly depend on the role the DBMS will play in an agency's data processing. One agency that performed an extensive cost/benefit analysis and developed a benchmark application to evaluate systems incurred costs of about $400,000 for procurement alone, and another $250,000 for conversion of existing files and programs. Another agency, which is making its DBMS the center of a complex administrative application, has already incurred costs of about $1.5 million for requirements studies, procurement, vendor support, and development of application programs using the DBMS. Large initial costs (more than $500,000) usually include some development costs as well as procurement, reorganization, conversion, training, and new hardware and support software. Costs for smaller systems are correspondingly less. Six other agencies estimated their total initial costs as "some," which could range from $100,000 to $200,000 or more. Even in these cases, then, initial costs of a DBMS are significant.

Though requirements studies and cost/benefit analyses may constitute a substantial investment in a DBMS, their costs can be small relative to the risks they avoid. In 1979 the General Accounting Office cited cases where poor planning had cost agencies $1-2 million [GA079]. Balanced against such losses, even $400,000 for procurement costs does not seem unreasonably high. At a minimum, an agency must expect a cursory comparison of available products to take at least 2 months of effort by a skilled member of the ADP staff. More detailed analyses and feature evaluations may take several months of effort by one or more ADP professionals. Making sure that a particular DBMS is the right product to buy can cost more than the purchase of the system itself.

Most of the interviewed agencies spent about what they had expected on procuring a DBMS. Their surprise costs came from new support software, new hardware, and new personnel needed to make the system work. For discussions of hardware and personnel costs, see later sections on recurring costs for computer resources and personnel. The only kind of new support software specifically mentioned was a teleprocessing (TP) monitor. Since Federal agencies sometimes use DBMS's for centralized processing of data from across the country, TP monitors can be essential parts of a large-scale data communications and processing system. In other cases the availability of a DBMS at a central agency site may make divisions of the agency want access to database software through some kind of telecommunications. Finally, using DBMS's with host programming languages sometimes entails writing support programs not supplied by the DBMS vendor. For example, one agency had to modify the vendor's PL/1 programs to support local recovery from disk instead of tape. The manager we interviewed considered the problems with PL/1 support a major cost and disadvantage of the DBMS.

Two categories of initial costs, reorganization and conversion, have actually accounted for considerably less expense than database theory would have predicted. In both cases the situation of Federal agencies is usually quite different from the "typical" situation addressed by textbooks. Reorganization generally costs Federal agencies less than agency managers might expect simply because less reorganization occurs. For example, political difficulties kept one manager from integrating files across applications; consequently, he did not need to reorganize the ADP staff to accommodate a new kind of data processing. The two managers whose agencies have thoroughly centralized their data report that reorganization has cost them "very much." One of them has had to reorganize his staff three times in the past 18 months to deal with changing application needs. Other agencies, which centralized data only in a particular application area, did not incur such costs. Six of the respondents said they had experienced little or no costs for reorganization, and another estimated the costs as "some."

Similarly, conversion costs for Federal agencies have been relatively low because the agencies have not been converting conventional applications to database applications. Federal agencies have often acquired DBMS's to solve new problems or to develop applications for systems that have not been automated before. Moreover, agencies frequently acquire new hardware at the same time and hence do not face conversion from one kind of equipment to another. Only 2 of the 10 respondents said that conversion had cost them "much"; 3 reported "some" costs, and 5 said that they had spent little or nothing on conversion of

existing files and programs.

4.5.2 Resource Requirements. Measuring the impact of a DBMS on system performance is not always easy. Respondents to the questionnaire reported somewhat more overall degradation of system performance than they had expected, but the only consensus was that a DBMS has little impact on the number of simultaneous users of the system. One singled out the disk storage requirements, another mentioned conflict between batch and interactive processing, and a third cited database loading as particular problems associated with a DBMS. For the most part, interviewed agencies did not consider degradation of system performance a major cost of a DBMS, but in a number of cases the DBMS's are running on hardware specifically acquired for them.

This dedication of a computer to a single software package represents a significant recurring cost of a DBMS. One agency ran its database applications on a central facility until yearly costs for time on the central processing unit (CPU) amounted to $250,000. At that time the agency acquired a minicomputer and its own copy of the DBMS, which now uses 80-90 percent of total CPU time on the new computer. The DBMS showed its need for resources in CPU charges on the mainframe and in slow response time on the minicomputer. Some other agencies use most of the resources of a large mainframe to run the DBMS; their annual costs for computer resources may be even higher than the $250,000 mentioned above. One agency with a large administrative application using a DBMS reported production costs for 1980 of almost $400,000, and another $150,000 for query, reporting, backup, and audit trail. Since this system was not previously automated, no data are available for a comparable system without a DBMS.

4.5.3 Increased Need for Technical Skills. The second important recurring cost of a DBMS results from the complexity of database software. To use a DBMS effectively, application programmers often have to know a good deal about the structure of data in the database. For some DBMS's, such knowledge requires a programmer to understand tree structures or more complex network structures. One interviewee thinks that many ADP staff members lack the conceptual ability to handle anything more difficult than the "mechanics" of routine operation. Designing databases is even more complex than using them. Five of the ten respondents said that they had experienced "much" or "very much" difficulty in designing databases for multiple users. While the difficulties were occasionally political, most of them involved technical issues like the modeling of many-to-many relationships between record types.

To meet the demand for skilled technical personnel, Federal agencies have had to spend time and money recruiting new personnel, training the current staff, or hiring contract support. Four of the respondents to the questionnaire stated that their DBMS had increased the need for technical support and advice by "much" or "very much," and five experienced "much" or "very much" difficulty in recruiting or training DBMS staff. In addition to the scarcity of programmers with appropriate DBMS experience, Federal agencies have had to contend both with the resistance of some traditional programmers and with discrepancies in the salaries of programmers in Government and industry. One interviewee told us that she knew of an employee who had left the agency to become an assistant database administrator after only six months of experience as a systems programmer and another year in database applications. This kind of demand for DBMS expertise often makes it difficult for agencies to hire and keep skilled technical personnel. While some vendors provide excellent technical support, only two of them provided "much" training to the agencies we interviewed.

4.5.4 Portability. The final important cost of a DBMS is its lack of portability. Only a few commercially available systems run on more than one kind of equipment, and agencies that want to change hardware may find that they are unable to do so without also changing their DBMS. One of the managers we interviewed decided to continue with the same DBMS in order to avoid conversion costs for existing applications. As a result his hardware procurement was limited to a major vendor and some smaller companies that sell "plug compatible" computers. Any agency that uses a DBMS extensively must recognize that its investment in application programs represents a commitment to a particular DBMS and, in many cases, to a particular hardware vendor.

## 4.6 Future Directions

While only four of the respondents to the questionnaire expect to procure another DBMS in the next 10 years, all of them expect database management systems to become increasingly important in the future. Several plan to expand their usage of DBMS's, and two are planning some kind of distributed database system. Despite the costs and complexity of a DBMS, all of them consider their current systems good investments. A couple of managers whose agencies are currently using their DBMS's primarily as access methods for particular applications hope to increase the number of application areas that make use of the generalized data management capabilities of a DBMS. If the attitudes of the interviewees is any key to the plans of ADP

managers throughout the Government, database management systems will assume an increasingly larger role in Federal data processing.

## 5. CONCLUSION

While theories about data management and cost/benefit analysis offer Federal ADP managers sound advice about the advantages and disadvantages of a DBMS, the substance of this report is the actual experience of the Government with DBMS products. Though the sample was small and the agencies varied, the interviews suggest several important conclusions about DBMS usage in the Federal Government.

1. The Federal Government already has a large investment in a wide variety of DBMS products and in diverse application programs using those products. Agency practices range from using a DBMS as a sophisticated access method for a single application area to making it the focal point of large, integrated data processing systems.

2. Regardless of how they use their DBMS's, ADP managers in Federal agencies seem pleased with the power, flexibility, and cost effectiveness of database management systems.

3. The major benefits of a DBMS are increased functionality and productivity, not reduced cost.

4. The amount of cost/benefit analysis needed to ensure success with a DBMS increases with the complexity of the system's intended use. Agencies that want a sophisticated access method for a single well-defined application may need little more than a comparison of available products. Agencies that expect to centralize data from several applications take greater risks and need correspondingly more analysis of requirements, costs, and benefits before committing themselves to database management in general or to a particular DBMS product. With centralized systems, large initial costs are likely to precede benefits by at least 1 or 2 years.

5. Many special applications legislated by Congress virtually require that an agency obtain a DBMS. The application manager may need software for frequent updates and queries or for quick response to agency management and to the public at large.

6. High-level user languages, including simple query facilities, often increase productivity by reducing the time for developing applications and enabling nonprogrammers to use computers effectively.

7. While database management systems can be cost effective in a number of different uses, any agency that plans to develop substantial database applications should recognize that those applications will probably not be easily transportable to another DBMS. Because DBMS packages often run on only a small number of computer systems, a commitment to one of these packages may restrict the agency's ability to change hardware environments.

These conclusions should be useful both to Federal managers considering a DBMS and to policymakers responsible for Government-wide ADP practices.

## 6. ACKNOWLEDGMENTS

Three groups of people contributed valuable material to this report. Managers from eight Federal agencies took their time to fill out questionnaires, to participate in interviews, and to give their advice to other agency managers. Vendor representatives supplied information about the number of Government installations of their products. Finally, several members of the Data Management and Programming Languages Division of NBS participated in interviews, provided important information, or read and criticized early versions of this report. They are J.V. Upperman, E.G. Adami, L.J. Gallagher, W.T. Hardgrave, D.R. Deutsch, and D.W. Fife.

# REFERENCES


[BLS79] Bureau of Labor Statistics, Department of Labor, Unpublished data showing Federal and U.S. ADP employment based on Current Population Surveys taken 1975-1979. Data exclude military and overseas personnel. ADP classifications are based on 1979 Census definitions.

[DATA79] Datapro Research Corporation, "All about Timesharing and Remote Computing Services," Datapro 70: The EDP Buyer's Bible, Vol. 3, September 1979.

[DATA80] Datapro Research Corporation, "Data Base Management Systems," Datapro Directory of Software.

[DATE77] Date, C. J., An Introduction to Database Systems, Addison-Wesley Publishing Company, Inc., Reading, 1977, pp. 6-9, 373-417. Date describes and compares the structures, operations, and access languages of three common data models.

[DEUT78] Deutsch, D., E. Fong, and J. Collica, "Cost Considerations for Generalized Database Management Systems," Generalized Data Management Systems and Scientific Information, a report of the specialist study on computer software, pp. 50-70. This paper identifies classes of costs for the entire life cycle of a database management system.

[GAO79] General Accounting Office, Data Base Management Systems -- Without Careful Planning There Can Be Problems, Washington, D.C., June 29, 1979.

[GSA79a] General Services Administration, Automatic Data Processing Equipment Inventory in the U.S. Government, Washington D.C., Automated Data and Telecommunications Service, 1979. Publication provides an overview of the use of ADP equipment by Federal agencies. Excluded from the document is comprehensive information regarding specialized digital computers.

[GSA79b] General Services Administration, Automatic Data Processing Activities Summary, Washington D.C., Automated Data and Telecommunication Service, 1979. A summary of ADP activities in the Federal Government as of September 30, 1978.

[GSA79c] General Services Administration, Unpublished data showing the number and dollar value of computers and storage devices in the Federal Government for FY77-FY79.

[GSA80a] General Services Administration, Automatic Data Processing Equipment Inventory in the U.S. Government, Washington D.C., Automated Data and Telecommunications Service, 1980. Publication provides an overview of the use of ADP equipment by Federal agencies. Excluded from the document is comprehensive information regarding specialized digital computers.

[GSA80b] General Services Administration, Automatic Data Processing Activities Summary, Washington D.C., Automated Data and Telecommunication Service, 1980. A summary of ADP activities in the Fedrral Government as of September 30, 1979.

[HEAD81a] Robert V. Head, "Federal ADP Systems: Atrophy in the Sinews of Government," Government Executive, February, 1981.

[HEAD81b] Robert V. Head, "The Complex Nature of Federal Data Processing," Government Executive, March, 1981.

[IDC73] International Data Corporation, "Briefing Session: The Computer Industry," Newtonville, MA, February 1973.

[IDC78] International Data Corporation, informal phone call in late October, 1978.

[IDC79] International Data Corporation, "Review and Forecast", EDP Industry Report, June 29, 1979. Report contains forecasts and trends in the U.S. ADP industry.

[IDC81a] International Data Corporation, "1981 Computer Industry Briefing Session," Waltham, MA, March, 1981.

[IDC81b] International Data Corporation, "Data Base Management Systems," contract report delivered to NBS in March, 1981.

[MART75] Martin, James, Computer Data-Base Organization, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1975, pp. 19-42. Martin explains in depth how database management systems use logical and physical organization to achieve data independence.

[OMB80] Office of Management and Budget, Federal Data Processing Resources, January 25, 1980. Summarizes resources for general purpose data processing activities of executive agences and reports trends in spending patterns for these resources.

[SIRC78] Sircar, Sumit, "The Importance of Data Base Management Systems," Business Quarterly, Vol. 43, No. 2, Summer 1978, pp. 45-49. Sircar explains how experiences in traditional operations have led to the need for "data independence." He also discusses the advantages, costs, and availability of DBMS's.

[THEI78] Theis, Douglas J., "An Overview of Memory Technologies," Datamation, Vol. 24, No. 1, January 1978, pp. 113-131. This article summarizes the costs and capabilities of several kinds of storage devices.

# APPENDIX A

## QUESTIONNAIRE

In order to determine how well database management systems are serving the Federal Government, NBS interviewed employees from eight Federal agencies. In each case we tried to select an ADP manager with broad responsibilities for data management and application development. Prior to each interview we sent the manager a three-part questionnaire about the agency's experiences with DBMS's. Part 1 asked for expectations and experiences with respect to six topics that seemed particularly important: centralization of data, data independence, ad hoc queries, performance degradation, personnel, and start-up costs. A response of 0 meant "don't know" or "not applicable," a response of 1 meant "little or none," and 2,3, and 4 meant "some," "much," and "very much," respectively. Part 2 used the same scale for questions about experiences (not expectations) in defining requirements and analyzing costs and benefits. Part 3 asked for information on selection and usage that did not fit either of the other formats. In three cases the person whom we interviewed had not been involved in procuring the DBMS, which serves a larger organizational unit. In each case we sent a copy of the questionnaire to someone who had participated in procurement, and two of the recipients responded. The rest of Appendix A consists of the questionnaire itself.

| QUESTION | EXPECTATIONS | | | | | EXPERIENCE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | don't know | little or none | some | much | very much | don't know | little or none | some | much | very much |
| | 0 | 1 | 2 | 3 | 4 | 0 | 1 | 2 | 3 | 4 |
| 1. Many advocates of database management systems stress the advantages of centralizing data and combining redundant data files, ultimately to form an organization-wide database. | | | | | | | | | | |
| 1.1. How much centralization of data did you expect (and experience) from using a DBMS? | 0 | 1 | 2 | 3 | 4 | 0 | 1 | 2 | 3 | 4 |
| 1.2. How much difficulty did you expect (and experience) in designing databases for multiple users? | 0 | 1 | 2 | 3 | 4 | 0 | 1 | 2 | 3 | 4 |
| 2. A second potential advantage of a DBMS is the ability to develop application data structures without concern for physical storage structures. The following questions concern this kind of data independence. | | | | | | | | | | |
| 2.1. To what extent did you expect (and experience) a reduction in the effort needed for designing and programming new applications as a result of application independence from storage structures? | 0 | 1 | 2 | 3 | 4 | 0 | 1 | 2 | 3 | 4 |
| 2.2. To what extent did you expect (and experience) easier program maintenance and modification because of the independence of logical data structures? | 0 | 1 | 2 | 3 | 4 | 0 | 1 | 2 | 3 | 4 |
| 3. One of the major advantages of some DBMS's is a high-level query language interface that enables users to retrieve answers to ad hoc queries and even to add, update, and delete data from the database. The following questions apply to on-line query processing: | | | | | | | | | | |
| 3.1. To what extent did you expect (and experience) improved timeliness and accuracy of information available to managers from the use of an ad hoc query facility? | 0 | 1 | 2 | 3 | 4 | 0 | 1 | 2 | 3 | 4 |
| 3.2. To what extent did you expect (and experience) that an ad hoc query facility would reduce the requirement for developing custom programs? | 0 | 1 | 2 | 3 | 4 | 0 | 1 | 2 | 3 | 4 |

| | | |
|---|---|---|
| 4. One potential disadvantage of a generalized DBMS is its degradation of overall computer system performance. The following questions concern this effect: | | |
| 4.1. How much degradation of system performance did you expect (and experience) from a DBMS? | 0 1 2 3 4 | 0 1 2 3 4 |
| 4.1.1. Of response time? | 0 1 2 3 4 | 0 1 2 3 4 |
| 4.1.2. Of throughput? | 0 1 2 3 4 | 0 1 2 3 4 |
| 4.1.3. Of number of simultaneous users? | 0 1 2 3 4 | 0 1 2 3 4 |
| 4.1.4. Of other performance features (identify)? | 0 1 2 3 4 | 0 1 2 3 4 |
| 5. Another potential disadvantage of a generalized DBMS is the need for skilled personnel to run it properly. The following questions concern this need: | | |
| 5.1. How much difficulty did you expect (and experience) in recruiting or training DBMS staff? | 0 1 2 3 4 | 0 1 2 3 4 |
| 5.2. What premium did you expect (and experience) to pay for ADP personnel with DBMS expertise? | 0 1 2 3 4 | 0 1 2 3 4 |
| 5.3. To what extent did you expect (and experience) increased needs for DBMS technical support and advice? | 0 1 2 3 4 | 0 1 2 3 4 |
| 5.4. How much training did you expect (and experience) from the DBMS vendor? | 0 1 2 3 4 | 0 1 2 3 4 |
| 6. In addition to its purchase price, a DBMS may require large start-up costs. What expenses did you expect (and experience) in each of the following categories? | | |
| 6.1. DBMS procurement (including requirements studies, cost/benefit studies, specification, selection, and software acquisition)? | 0 1 2 3 4 | 0 1 2 3 4 |
| 6.2. DBMS installation? | 0 1 2 3 4 | 0 1 2 3 4 |
| 6.3. New support software? | 0 1 2 3 4 | 0 1 2 3 4 |
| 6.4. New hardware? | 0 1 2 3 4 | 0 1 2 3 4 |
| 6.5. New personnel? | 0 1 2 3 4 | 0 1 2 3 4 |

| QUESTION | EXPECTATIONS | | | | | EXPERIENCE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | don't know | little or none | some | much | very much | don't know | little or none | some | much | very much |
| | 0 | 1 | 2 | 3 | 4 | 0 | 1 | 2 | 3 | 4 |
| 6.6. Training? | 0 | 1 | 2 | 3 | 4 | 0 | 1 | 2 | 3 | 4 |
| 6.7. Reorganization (including the establishment of a Database Administrator)? | 0 | 1 | 2 | 3 | 4 | 0 | 1 | 2 | 3 | 4 |
| 6.8. Conversion of existing data files and programs? | 0 | 1 | 2 | 3 | 4 | 0 | 1 | 2 | 3 | 4 |
| 6.9. Other (identify)? | 0 | 1 | 2 | 3 | 4 | 0 | 1 | 2 | 3 | 4 |
| 6.10. TOTAL? | 0 | 1 | 2 | 3 | 4 | 0 | 1 | 2 | 3 | 4 |

DBMS Costs and Benefits

A Questionnaire for Selected Federal ADP Units

Part 2

| QUESTION | EXPERIENCE | | | | |
|---|---|---|---|---|---|
| | don't know | little or none | some | much | very much |
| | 0 | 1 | 2 | 3 | 4 |
| **1. Cost/Benefit Analysis** | | | | | |
| 1.1. How much cost/benefit analysis did you do before you acquired a DBMS? | 0 | 1 | 2 | 3 | 4 |
| 1.2. How much cost/benefit analysis do you think is necessary? | 0 | 1 | 2 | 3 | 4 |
| 1.3. How much difficulty did you experience in identifying costs and benefits? | 0 | 1 | 2 | 3 | 4 |
| 1.4. How much difficulty did you experience in quantifying intangible costs and benefits? | 0 | 1 | 2 | 3 | 4 |
| 1.5. How much would a guideline on cost/benefit methodology help you in future DBMS acquisitions? | 0 | 1 | 2 | 3 | 4 |
| 1.6. How much help was your cost/benefit analysis in your planning for DBMS selection, acquisition, and installation? | 0 | 1 | 2 | 3 | 4 |
| **2. Requirements** | | | | | |
| 2.1. To what extent did you consult users to determine their application requirements before you acquired a DBMS? | 0 | 1 | 2 | 3 | 4 |
| 2.2. How much difficulty did you experience in defining application requirements? | 0 | 1 | 2 | 3 | 4 |

Part 3

1. Deciding to Acquire a DBMS

1.1. Which of the following did you do before you decided to acquire a DBMS?

Requirements Study

Cost/Benefit Analysis

Other (please identify)

1.2. Besides a DBMS, what solutions to your ADP problems did you consider?

1.3. What did you consider the major advantages and disadvantages of a DBMS (please list the most important first)?

ADVANTAGES DISADVANTAGES

1.4. What do you now consider the major advantages and disadvantages of a DBMS (please list the most important first)?

ADVANTAGES DISADVANTAGES

2. Selection and Specification

2.1. Which of the following did you do before selecting a DBMS?

Detailed Feature Evaluation

Develop a Benchmark

Other (please identify)

2.2. How many DBMS products did you consider?

2.3. How many vendors bid on your RFP for a DBMS?

2.4. How many DBMS products satisfied your mandatory requirements?

2.5. Did you want a particular DBMS product? If so, did you get it?

3. DBMS Usage

3.1. Does your DBMS satisfy your users' requirements?

3.2. Which of the following factors limit your usage of your DBMS?

Poor performance?

Complicated languages, commands, or structures?

Lack of database skills?

Poor vendor support?

Limited DBMS capabilities?

DBMS features not necessary to applications?

3.3. What features of your DBMS have disappointed you most?

3.4. What features of your DBMS have pleased you most?

3.5. Will you be converting to another DBMS within the next 3 years? 5 years? 10 years?

4. What hardware configuration supports your DBMS?

4.1. CPU Manufacturer and Model?

4.2. Size of Main Memory?

4.3. Size of Secondary Memory?

# APPENDIX B

## A TABULATION OF RESPONSES TO THE QUESTIONNAIRE

This Appendix reproduces the questionnaire with some minor changes to show overall results. For every question, entries in each column indicate the number of people who gave the response shown in the column heading. Since 10 people completed the questionnaire, there are a total of 10 answers to each question except those labeled "other." Questions left blank counted as a "don't know" response. To highlight the differences between expectations and experience in Part 1 of the questionnaire, we have added a column labeled "DIFFERENCE" that did not appear in the actual questionnaire. In calculating the differences for this column, we subtracted each agency's EXPECTATIONS from its EXPERIENCE to get a number between -3 and +3, inclusive. We decided that a response of "don't know" for either expectations or experience gave a difference of 0, no matter what the other response was. The entry in each column under DIFFERENCE indicates the number of respondents who reported the difference between experience and expectations shown in the column heading.

We have also tabulated responses to Part 2 of the questionnaire, which dealt only with experiences of agencies, not with their expectations. We have not tried to write a general summary of all the narrative responses to Part 3; however, a narrative account of each response appears in Appendix C as part of the appropriate agency's interview report.

DBMS Costs and Benefits

A Tabulation of Responsee to the Queetionnaire

| QUESTION | EXPECTATIONS | | | | | EXPERIENCE | | | | | DIFFERENCE (EXPERIENCE - EXPECTATIONS) | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | don't know | little or none | some | much | very much | don't know | little or none | some | much | very much | | | | | | | |
| | 0 | 1 | 2 | 3 | 4 | 0 | 1 | 2 | 3 | 4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 |
| 1. Many advocates of database management systems stress the advantagee of centralizing data and combining redundant data files, ultimately to form an organization-wide database. | | | | | | | | | | | | | | | | | |
| 1.1. How much centralization of data did you expect (and experience) from using a DBMS? | 1 | 2 | 3 | 2 | 2 | 0 | 1 | 3 | 3 | 3 | 0 | 0 | 0 | 9 | 0 | 0 | 1 |
| 1.2. How much difficulty did you expect (and experience) in designing databases for multiple users? | 1 | 1 | 5 | 2 | 1 | 0 | 4 | 2 | 2 | 2 | 0 | 1 | 4 | 2 | 1 | 2 | 0 |
| 2. A second potential advantage of a DBMS is the ability to develop application data structures without concern for phyeical storage etructures. The following questions concern this kind of data independence. | | | | | | | | | | | | | | | | | |
| 2.1. To what extent did you expect (and experience) a reduction in the effort needed for designing and programming new applications ae a result of application independence from storage structuree? | 1 | 0 | 3 | 4 | 2 | 0 | 1 | 2 | 6 | 1 | 0 | 0 | 4 | 3 | 2 | 1 | 0 |
| 2.2. To what extent did you expect (and experience) easier program maintenance and modification because of the independence of logical data structures? | 0 | 0 | 5 | 4 | 1 | 0 | 1 | 3 | 4 | 2 | 0 | 0 | 3 | 4 | 2 | 1 | 0 |

A Tabulation of Responses to the Questionnaire -- Continued

| QUESTION | EXPECTATIONS | | | | | EXPERIENCE | | | | | DIFFERENCE (EXPERIENCE - EXPECTATIONS) | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | don't know | little or none | some | much | very much | don't know | little or none | some | much | very much | | | | | | | |
| | 0 | 1 | 2 | 3 | 4 | 0 | 1 | 2 | 3 | 4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 |
| 3. One of the major advantages of some DBMS's is a high-level query language interface that enables users to retrieve answers to ad hoc queries and even to add, update, and delete data from the database. The following questions apply to on-line query processing: | | | | | | | | | | | | | | | | | |
| 3.1. To what extent did you expect (and experience) improved timeliness and accuracy of information available to managers from the use of an ad hoc query facility? | 2 | 1 | 0 | 5 | 2 | 1 | 1 | 1 | 5 | 2 | 0 | 0 | 0 | 10 | 0 | 0 | 0 |
| 3.2. To what extent did you expect (and experience) that an ad hoc query facility would reduce the requirement for developing custom programs? | 3 | 1 | 1 | 5 | 0 | 2 | 0 | 4 | 4 | 0 | 0 | 0 | 2 | 7 | 1 | 0 | 0 |
| 4. One potential disadvantage of a generalized DBMS is its degradation of overall computer system performance. The following questions concern this effect: | | | | | | | | | | | | | | | | | |
| 4.1. How much degradation of system performance did you expect (and experience) from a DBMS? | 1 | 5 | 4 | 0 | 0 | 1 | 3 | 4 | 2 | 0 | 0 | 0 | 0 | 7 | 2 | 1 | 0 |
| 4.1.1. Of response time? | 1 | 5 | 4 | 0 | 0 | 0 | 6 | 2 | 2 | 0 | 0 | 0 | 1 | 7 | 1 | 1 | 0 |
| 4.1.2. Of throughput? | 0 | 5 | 5 | 0 | 0 | 1 | 5 | 2 | 2 | 0 | 0 | 0 | 2 | 5 | 2 | 1 | 0 |
| 4.1.3. Of number of simultaneous users? | 0 | 6 | 3 | 1 | 0 | 0 | 7 | 1 | 2 | 0 | 0 | 0 | 1 | 8 | 1 | 0 | 0 |

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4.1.4. Of other performance features (identify)?<br><br>USGS reported a large difference for disk storage requirements. | 2 | 4 | 1 | 1 | 0 | 2 | 2 | 1 | 2 | 1 | 0 | 0 | 0 | 5 | 2 | 0 | 1 |
| 5. Another potential disadvantage of a generalized DBMS is the need for skilled personnel to run it properly. The following questions concern this need: | | | | | | | | | | | | | | | | | |
| 5.1. How much difficulty did you expect (and experience) in recruiting or training DBMS staff? | 0 | 2 | 7 | 0 | 1 | 0 | 2 | 3 | 3 | 2 | 0 | 1 | 2 | 1 | 3 | 3 | 0 |
| 5.2. What premium did you expect (and experience) to pay for ADP personnel with DBMS expertise? | 2 | 3 | 2 | 3 | 0 | 4 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 10 | 0 | 0 | 0 |
| 5.3. To what extent did you expect (and experience) increased needs for DBMS technical support and advice? | 0 | 1 | 5 | 4 | 0 | 0 | 0 | 6 | 1 | 3 | 0 | 0 | 2 | 4 | 3 | 0 | 1 |
| 5.4. How much training did you expect (and experience) from the DBMS vendor? | 0 | 0 | 8 | 2 | 0 | 0 | 2 | 6 | 2 | 0 | 0 | 0 | 3 | 6 | 1 | 0 | 0 |
| 6. In addition to its purchase price, a DBMS may require large start-up costs. What expenses did you expect (and experience) in each of the following categories? | | | | | | | | | | | | | | | | | |
| 6.1. DBMS procurement (including requirements studies, cost/benefit studies, specification, selection, and software acquisition)? | 0 | 3 | 6 | 0 | 1 | 0 | 3 | 6 | 0 | 1 | 0 | 0 | 0 | 10 | 0 | 0 | 0 |
| 6.2. DBMS installation? | 1 | 2 | 7 | 0 | 0 | 1 | 4 | 4 | 1 | 0 | 0 | 0 | 2 | 7 | 1 | 0 | 0 |
| 6.3. New support software? | 1 | 2 | 6 | 1 | 0 | 1 | 1 | 4 | 3 | 1 | 0 | 0 | 1 | 5 | 2 | 2 | 0 |
| 6.4. New hardware? | 2 | 4 | 3 | 1 | 0 | 2 | 2 | 4 | 2 | 0 | 0 | 0 | 0 | 8 | 1 | 1 | 0 |

NOTE: Entries in each column indicate the number of people who gave the response shown in the column heading.

A Tabulation of Responses to the Questionnaire -- Continued

| QUESTION | EXPECTATIONS | | | | | EXPERIENCE | | | | | DIFFERENCE (EXPERIENCE - EXPECTATIONS) | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | don't know | little or none | some | much | very much | don't know | little or none | some | much | very much | | | | | | | |
| | 0 | 1 | 2 | 3 | 4 | 0 | 1 | 2 | 3 | 4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 |
| 6.5. New personnel? | 1 | 5 | 4 | 0 | 0 | 1 | 4 | 2 | 3 | 0 | 0 | 0 | 0 | 7 | 2 | 1 | 0 |
| 6.6. Training? | 1 | 0 | 7 | 1 | 1 | 1 | 0 | 6 | 1 | 2 | 0 | 1 | 1 | 5 | 1 | 2 | 0 |
| 6.7. Reorganization (including the establishment of a Database Administrator)? | 1 | 3 | 2 | 3 | 1 | 1 | 6 | 1 | 0 | 2 | 0 | 2 | 1 | 6 | 1 | 0 | 0 |
| 6.8. Conversion of existing data files and programs? | 0 | 4 | 5 | 1 | 0 | 0 | 5 | 3 | 2 | 0 | 0 | 1 | 0 | 7 | 2 | 0 | 0 |
| 6.9. Other (identify)? For further details about the large differences, see the individual questionnaires for NAVSEA (AEGIS) and the Civil Aeronautics Board. | 0 | 3 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 0 | 0 | 0 | 0 | 2 | 0 | 2 |
| 6.10. TOTAL? | 3 | 1 | 6 | 0 | 0 | 3 | 0 | 6 | 1 | 0 | 0 | 0 | 0 | 8 | 2 | 0 | 0 |

NOTE: Entries in each column indicate the number of people who gave the response shown in the column heading.

DBMS Coete and Benefite

A Tabulation of Resbonees to the Questionnaire

Part 2

| QUESTION | EXPERIENCE | | | | |
|---|---|---|---|---|---|
| | don't know | little or none | eome | much | very much |
| | 0 | 1 | 2 | 3 | 4 |
| 1. Cost/Benefit Analysis | | | | | |
| 1.1. How much coet/benefit analysis did you do before you acquired a DBMS? | 1 | 1 | 6 | 0 | 2 |
| 1.2. How much cost/benefit analysis do you think is necessary? | 1 | 2 | 5 | 1 | 1 |
| 1.3. How much difficulty did you experience in identifying costs and benefite? | 1 | 3 | 4 | 2 | 0 |
| 1.4. How much difficulty did you experience in quantifying intangible costs and benefite? | 1 | 3 | 2 | 2 | 2 |
| 1.5. How much would a guideline on cost/benefit methodology help you in future DBMS acquisitions? | 2 | 2 | 2 | 3 | 1 |
| 1.6. How much help was your coet/benefit analysis in your planning for DBMS eelection, acquisition, and installation? | 3 | 3 | 2 | 2 | 0 |
| 2. Requirements | | | | | |
| 2.1. To what extent did you consult usere to determine their application requirements before you acquired a DBMS? | 0 | 2 | 2 | 3 | 3 |
| 2.2. How much difficulty did you experience in defining application requirements? | 1 | 1 | 3 | 4 | 1 |

NOTE: Entries in each column indicate the number of people who gave the responee shown in the column heading.

# APPENDIX C

## INTERVIEW REPORTS

Appendix C consists of responses by particular agencies. For each agency there is a table of answers to the questionnaire and a narrative summary of the interview. When the interview included discussions of specific answers to the questionnaire, we have summarized each discussion in the table of responses rather than in the general narrative. On the other hand, answers to fill-in-the-blank questions are subsumed within the interview narrative. For one of the two cases where we received a second questionnaire, the responses appear separately with no identification of the agency. For the other, the second person's answers to Parts 1 and 2 appear in a separate table from those provided by the interviewee. Answers to Part 3 are included in the interview narrative.

In accordance with NBS policy, the following interview reports do not name particular commercial products. Instead, we have used phrases like "Product A" and "the vendor of Product B." We have included the criticisms, compliments, and comparisons of these products in order to identify specific features that have plagued or pleased Federal ADP managers. In every case the opinions about DBMS products, applications, and problems are only the opinions of an individual. With two exceptions, the agencies we interviewed have agreed to let us identify them, but this identification does not constitute official endorsement of the opinion expressed by the manager. Neither NBS nor the agency for whom a manager works endorses any explicit or implicit descriptions or evaluations of specific products.

INTERVIEW REPORT

BUREAU OF RADIOLOGICAL HEALTH

The Bureau of Radiological Health (BRH) monitors compliance with Federal standards for radiological emissions from commercial products, conducts research to determine the biological effects of radiation, and conducts educational programs on radiation safety. The manager we interviewed first came to BRH in 1973, when the information system was maintained on 80 column punch cards. At that time BRH used FORTRAN programs to generate some reports from this data. Some of the current applications did not exist then, others were inefficiently automated, and still others were not automated at all. BRH's parent agency already owned a DBMS that the interviewee considered ideal for the bureau's needs. He had to have quick solutions to problems, and he preferred to deal with the nonprogrammers on his staff rather than recruit specialized programmers. He also thought his division needed the impact of a dramatic, visible change in its methods of data processing.

The interviewee felt that Product A was flexible and simple to learn and use; he also recognized that a good understanding of the database would be helpful in designing applications and databases. For example, when he first began using Product A, it did not allow users to redefine keys conveniently. The database designer had to be certain that the initial choice of key fields was right; to make a change meant recreating the database. Centralization of data was not originally a motivation for using a DBMS, but it has turned out to be an advantage.

Using Product A at the parent agency's computing center, BRH's data management staff developed one large application consisting of seven smaller intertwined applications. This application tracks administrative records and performs technical analysis of compliance data. All of the interactive programs and some of the batch programs are written in Product A's user language. PL/1 programs edit data, reformat records, perform mathematical calculations, and produce final reports. Eventually even these programs will be rewritten in the user language.

As BRH developed more programs for Product A, the application became very expensive to operate on the parent agency's computer. The bureau spent about $250,000 for computer time during the last full year of leasing, after which management decided to develop or buy a new system. BRH undertook an informal requirements and cost/benefit study, which covered five commercial DBMS products. The bureau

decided to stay with Product A because it provided a flexible data manipulation mechanism and report generator, and because doing so would avoid training and conversion costs. BRH's copy of Product A was the second within the department and had a reduced purchase price; the interviewee did not know how the cost of their maintenance contract compares with that of the parent agency. The procurement of the computer system (hardware) was an open solicitation; however, the new system had to support Product A, which runs on only one vendor's equipment or compatible systems. Only that vendor submitted a bid, and BRH procured the vendor's minicomputer.

With the architecture of the new computer, Product A showed its resource needs in slow response time rather than costs for CPU time. The DBMS used 80-90% of the resources of the new computer. Because the computer offered only one megabyte of main memory, it spent a lot of time in paging of user areas. BRH has since upgraded to an enhanced version of the minicomputer, which provides four megabytes of main memory and reportedly doubles CPU responsiveness. The agency has 1.1 billion bytes of disk space. While the CPU is very reliable, BRH has had some problems with disks. Since the upgrade, CPU time has improved by a factor of two; moreover, better I/O has improved real time responsiveness (wall clock time) by a factor of 10 or more.

BRH's usage of Product A remains high. Ten employees write programs in the user language, and another 20 use canned programs for various tasks. From June of 1980 through May of 1981 BRH averaged 1100 sessions (logins), 565,000 terminal I/O's, and 385,000 disk I/O's per month. The bureau adds about 800-900 records per week and updates about 200 records per week (5-6 fields per record). BRH enters most of its data in batch mode using a data loading language. Though the user language is not so efficient in its use of CPU time, it enables nonprogrammers to perform all these functions. The language's query capabilities have not eliminated the need for individual programs, but such programs are easier to develop in the user language than they would be in a conventional programming language.

In general, Product A is easy to learn and use, although some employees learn only the basics and do not use the more complex and sophisticated features of the system. The DBMS vendor provides user training and is best in the user language; for other techniques the national and local users' groups for the DBMS are very helpful. Each of them meets twice a year. Attendance at the national meetings averages about 100, of whom slightly more than half represent Federal agencies.

The DBMS vendor compensates for some lack of initial systems training by providing excellent support. Though not quite so individualized as it was when the user community was smaller, this support still enables an owner of the system to call the vendor, ask a question or describe a problem, and get a rapid response. Some requests for modifications may also receive a response in a number of weeks. The vendor has been very helpful in upgrading its product and in occasionally making specific modifications tailored to BRH's needs.

## ADVANTAGES of a DBMS

1. Rapid "program" development without the need to wait for programmers or specialized personnel
2. Simplicity in learning how to use the system, make requests, and develop reports
3. Ease of data interaction (multi-files and multi-formats) without tight data structure or formatting requirements.

## DISADVANTAGES of a DBMS

1. Vulnerability to loss of data integrity due to system crashes or TP line drops (Product A flags errors that result from system crashes, but it also flags some false alarms.)
2. Lack of simple and quick file reorganization facility.
3. Need for good programming techniques to avoid "locking" one user out of data that another user may be processing.

Responses to the Questionnaire

Bureau of Radiological Health

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| Part One | | |
| 1. Centralization | | |
| 1.1 Extent | little or none | very much |
| 1.2 Difficulty | little or none | little |
| 2. Data Independence | | |
| 2.1 Reduced Effort in Application Development | very much | much |
| 2.2 Easier Maintenance and Modification | very much | very |
| 3. High-level Queries | | |
| 3.1 Improved Timeliness and Accuracy | much | much |
| 3.2 Reduced Need for Custom Programs | little or none | some |
| 4. Degradation of System Performance | | |
| 4.1 Overall | little or none | much |
| 4.1.1 Response Time | little or none | much |
| 4.1.2 Throughput | little or none | some |
| 4.1.3 Number of Simultaneous Users | some | some |
| 4.1.4 Other<br>Lock out of record updating. | little or none | some |
| 5. Personnel | | |
| 5.1 Difficulty in Recruiting or Training | little or none | some |
| 5.2 Premium for DBMS Expertise | don't know | don't know |
| 5.3 Increased Need for Technical Support | some | some |
| 5.4 Amount of Vendor Training | some | little or none |
| 6. Start-up Costs | | |
| 6.1 DBMS procurement | little or none | little or none |
| 6.2 DBMS installation | little or none | little or none |
| 6.3 New support software | little or none | some |
| 6.4 New hardware | much | much |

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| 6.5 New personnel | little or none | little or none |
| 6.6 Training | some | some |
| 6.7 Reorganization | little or none | little or none |
| 6.8 Conversion of existing data files and Programs | little or none | little or none |
| 6.9 Other<br>High cost of manuals. | little or none | some |
| 6.10 TOTAL | some | some |

Responses to the Questionnaire

Bureau of Radiological Health

| QUESTION | EXPERIENCE |
|---|---|
| Part Two<br>1. Cost/Benefit Analysis | |
| 1.1 Amount Done | some |
| 1.2 Amount Necessary | much |
| 1.3 Difficulty in Identifying Costs and Benefits | some |
| 1.4 Difficulty in Quantifying Intangible Costs and Benefits | much |
| 1.5 Need of a Guidsline | much |
| 1.6 Amount of Help Provided by Cost/Benefit Analysis in DBMS Procurement | little or none |
| 2. Requirements | |
| 2.1 Extent of Consultation with Users Before Procurement | little or none |
| 2.2 Difficulty in Defining Application Requirements | much |

## INTERVIEW REPORT

## CIVIL AERONAUTICS BOARD

The Civil Aeronautics Board is now operating under "sunset" legislation, and the person actively involved in database management acquired Product B to overcome the shortness of staff by reducing the time for developing applications. Of the seven systems whose vendors responded to the RFP, only three satisfied CAB's mandatory requirements. The agency eliminated another because it seemed to require standards and a full-size application staff, including a database administrator (DBA). To compare systems, the manager contacted each of the three vendors and set a requirement that the vendor first install the DBMS and then let the CAB try the system for 2 weeks with the agency's own users, files, and site. Only the vendor of Product B responded; the others dropped out. In the interviewee's estimation, Product B has fully satisfied the original requirements.

The CAB did not have to reorganize its data processing staff to accommodate a DBMS. Few of the agency's applications share common data; in fact, the person we interviewed said that centralized data seems better suited to a manufacturing environment than to governmental monitoring of private industry. The agency did experience conflicts between conventional programmers used to batch processing and new end users who wanted immediate results from on-line processes. The manager believes that a system designed to operate in batch mode cannot really handle on-line demands. He also said that a successful DBMS environment requires a strong DBA who can persuade or force programmers to consider central concerns while doing their work. Converting to a DBMS implies change, and the threat of change can evoke resistance from current ADP staff. The interviewee knew of several organizations in which database administration proved impossible: the DBA gave in to individual demands and finally had to resign when things did not work out.

Product B's primary advantage for CAB is its user language. Whereas procedural languages do not tie in well with a DBMS (e.g., COBOL's End of Data statement), Product B's user language ties in directly to the data dictionary. The CAB develops almost everything in the user language, which enables end users to deal with problems more effectively because they know what questions to ask when they run into problems.

The interviewee said that he preferred Product B to a similar DBMS for a number of reasons. The other system is powerful but not controlled -- e.g., a user can divide a name by 2 unless a DBA writes an application program to prevent such things. Product B's compression techniques are simple but effective. The system saves about 50 percent merely by eliminating the leading 0's of numerics and the trailing blanks of character strings.

## ADVANTAGES of a DBMS

1. Reduction of effort
2. Better responsiveness to user requirements
3. Relieving programmers from trivial tasks
4. Placing responsibility for data in user's territory
5. Quicker development cycle

## DISADVANTAGES of a DBMS

1. Dealing with Federal procurement regulations
2. Conflict between batch and on-line processing
3. Retraining and recruiting staff

Responses to the Questionnaire

Civil Aeronautics Board

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| Part One | | |
| 1. Centralization | | |
| 1.1 Extent | some | some |
| Few of the agency's applications share common data. It's too much trouble to eliminate duplicate files. | | |
| 1.2 Difficulty | some | little or none |
| 2. Data Independence | | |
| 2.1 Reduced Effort in Application Development | some | very much |
| The DBMS's user language provides a 90% saving in code and a 75% saving in effort compared to COBOL. | | |
| 2.2 Easier Maintenance and Modification | some | very much |
| 3. High-level Queries | | |
| 3.1 Improved Timeliness and Accuracy | much | much |
| 3.2 Reduced Need for Custom Programs | much | much |
| 4. Degradation of System Performance | | |
| 4.1 Overall | some | some |
| 4.1.1 Response Time | some | some |
| 4.1.2 Throughput | some | little or none |
| 4.1.3 Number of Simultaneous Users | little or none | little or none |
| 4.1.4 Other | much | much |
| Batch contention with on-line. | | |
| 5. Personnel | | |
| 5.1 Difficulty in Recruiting or Training | little or none | much |
| The answer is for recruiting. Training presented no problems. | | |
| 5.2 Premium for DBMS Expertise | little or none | little or none |
| 5.3 Increased Need for Technical Support | some | some |
| 5.4 Amount of Vendor Training | some | some |
| 6. Start-up Costs | | |
| 6.1 DBMS procurement | some | some |
| The start-up costs for the DBMS included $100,000 for the purchase price and about eight months of the interviewee's time. | | |

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| 6.2 DBMS installation | some | little or none |
| 6.3 New support software | some | some |
| 6.4 New hardware | some | some |
| 6.5 New personnel | little or none | little or none |
| 6.6 Training | some | some |
| 6.7 Reorganization | much | little or none |
| 6.8 Conversion of existing data files and Programs | much | little or none |
| 6.9 Other<br>Resistance from traditional programmers. | little or none | very much |
| 6.10 TOTAL | some | some |

Responses to the Questionnaire

Civil Aeronautics Board

| QUESTION | EXPERIENCE |
|---|---|
| Part Two | |
| 1. Cost/Benefit Analysis | |
| 1.1 Amount Done<br>The manager spent about three months of his time dealing with Governmental procurement regulations. | very much |
| 1.2 Amount Necessary | some |
| 1.3 Difficulty in Identifying Costs and Benefits | some |
| 1.4 Difficulty in Quantifying Intangible Costs and Benefits | some |
| 1.5 Need of a Guideline | little or none |
| 1.6 Amount of Help Provided by Cost/Benefit Analysis in DBMS Procurement | much |
| 2. Requirements | |
| 2.1 Extent of Consultation with Users Before Procurement | very much |
| 2.2 Difficulty in Defining Application Requirements | little or none |

INTERVIEW REPORT

NAVAL SEA SYSTEMS COMMAND (PMS-400)
(AEGIS)

The AEGIS Shipbuilding Project Office of the Naval Sea Systems Command (NAVSEA) is currently developing a large-scale integrated database for use with Product C. AEGIS has had a management information system (MIS) since 1969. In 1976 managers decided to get another MIS for the AEGIS project, which in 1977 became a Shipbuilding Project for certain destroyers and nuclear cruisers.

The project's old PL/1 system with flat files did not have the flexibility or portability that AEGIS needed. Project managers contracted with Syscon Corporation of America to help them define their requirements. They investigated several possibilities, including in-house development and at least five commercial DBMS's. They wanted more flexible storage and relationships to ensure adaptability for new applications.

In selecting a DBMS, the AEGIS Project managers faced two particularly important constraints. First of all, it was quickly clear that the administrative system involved many-to-many relationships and therefore required the flexibility of network data structures. Secondly, because they were buying computer time from the Federal Data Center at the National Institutes of Health, they needed a package that could run on a particular vendor's equipment. In early 1978 NIH turned down their request to use that vendor's DBMS, and AEGIS arranged instead to use Product C, which another part of the Navy had acquired to use on NIH's equipment. To work with Product C, AEGIS also used the same vendor's report writer and paid $12,000 for the vendor's data dictionary. Vendor representatives said they would support recovery from disk (tape is too slow at NIH), and AEGIS agreed to pay the vendor the regular annual maintenance fee of less than $10,000 to cover updates, consultation, and other services for all three products.

Since the beginning of database design, the Product C project has employed 5-10 people, including management analysts and technical staff. In trying to recruit new employees, the project has found that there are only a small number of people with training in Product C, and an even smaller number skilled in both Product C and PL/1. The project soon gave up on trying to get people with both skills, and has sought instead to hire PL/1 programmers interested in learning a DBMS. In particular, the interviewee has had trouble hiring a DBA with experience; he opted to hire a

person with the potential to learn.

AEGIS has received a lot of good advice through the Product C User's Group. At least half of the members of the local users' group represent Federal agencies, and the chairman of the national users' group has enormous experience in the impact of a DBMS on an organization.

The new system that AEGIS is developing is cheaper than the old system in updates, but more expensive in complex queries that are common for this application. Because the old system was extraordinarily redundant, update costs were expected to increase rapidly as files grew. To update two data elements would have cost $10 on the old system compared to 75 cents on the new one.

In general, the manager of the Product C project is satisfied with its progress, even though AEGIS has not yet realized many benefits from an initial investment of about $1.5 million and at least 3 calendar years. Development is nearing completion, and before long he hopes to have a fully functional, extremely flexible system for processing the data of the entire AEGIS project. While it was hard to convince people throughout the project that centralizing data was a good idea, and while the ADP staff have often underestimated their needs, the manager is convinced that the end product will justify its initial costs.

ADVANTAGES of a DBMS

1. Lack of data redundancy
2. Quality control of data
3. Easier modification of application software

DISADVANTAGES of a DBMS

1. Dependence on vendor support
2. Need for more highly skilled technical personnel
3. Scarcity of highly skilled technical personnel

Responses to the Questionnaire

Naval Sea Systems Command (PMS-400)

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| Part One | | |
| 1. Centralization | | |
| 1.1 Extent | much | much |
| 1.2 Difficulty | eome | very much |
| In designing databases for multiple users, employees had to wait for the policies and procedures of the new organization. They had to guess data needs and invest money and time up front. The good guessee have worked out; others have made it neceeary to redesign parts of the database. Several members of the etaff spent eight calendar monthe and a total of fifteen months of individual effort logically analyzing data elements from different projects. At the end of the period they had a couple of hierarchies linked in several places and some isolated data elements with many-to-many relationshipe. The manager of the DBMS project thinks that too few people were involved in the database design. While the software is straightforward for a talented person, the logic ie complex and difficult. | | |
| 2. Data Independence | | |
| 2.1 Reduced Effort in Application Development | much | eome |
| The DBMS does not yet support full logical data independence; it requiree programmers to know a lot about the database. Even with these difficultiee, etaff members produced 200 PL/1 programe in 1979 and another 350 in 1980. They have also modified eome DBMS software to read diek inetead of tape. | | |
| 2.2 Easier Maintenance and Modification | much | some |
| 3. High-level Queries | | |
| 3.1 Improved Timelinese and Accuracy | don't know | don't know |
| 3.2 Reduced Need for Custom Programs | don't know | don't know |
| 4. Degradation of System Performance | | |
| 4.1 Overall | little or none | little or none |
| 4.1.1 Response Time | little or none | little or none |
| 4.1.2 Throughput | little or none | little or none |
| 4.1.3 Number of Simultaneous Users | little or none | little or none |
| 4.1.4 Other | don't know | don't know |
| 5. Pereonnel | | |
| 5.1 Difficulty in Recruiting or Training | eome | very much |
| Only a emall number of people are available who know both PL/1 and the agency's DBMS. | | |
| 5.2 Premium for DBMS Expertise | much | don't know |

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| 5.3 Increased Need for Technical Support | some | much |
| 5.4 Amount of Vendor Training | much | much |
| 6. Start-up Costs | | |
| 6.1 DBMS procurement | some | some |
| 6.2 DBMS installation | some | some |
| 6.3 New support softwars | little or none | much |
| 6.4 New hardware | little or none | little or none |
| 6.5 New personnel | soms | much |
| 6.6 Training | some | some |
| 6.7 Reorganization | much | very much |
| 6.8 Conversion of existing data filss and Programs | some | much |
| 6.9 Other<br>Vendor support of PL/1 language | little or none | very much |
| 6.10 TOTAL | don't know | don't know |

Responses to the Questionnaire

Naval Sea Systems Command (PMS-400)

| QUESTION | EXPERIENCE |
|---|---|
| Part Two | |
| 1. Cost/Benefit Analysis | |
| 1.1 Amount Done | some |
| 1.2 Amount Necessary | some |
| 1.3 Difficulty in Identifying Costs and Benefits | little or none |
| 1.4 Difficulty in Quantifying Intangible Costs and Bensfits | very much |
| 1.5 Need of a Guideline | some |
| 1.6 Amount of Help Provided by Cost/Benefit Analysis in DBMS Procurement | don't know |
| 2. Requirements | |
| 2.1 Extent of Consultation with Users Before Procursment | some |
| 2.2 Difficulty in Defining Application Requirements | much |

INTERVIEW REPORT

NATIONAL INSTITUTES OF HEALTH

The central facility for data processing at the National Institutes of Health (NIH) is the Computer Center in the NIH Division of Computer Research and Technology (DCRT). This central facility serves over 8,000 users who have a wide variety of scientific, statistical, and administrative computing tasks. It includes large mainframes linked by communication lines and systems software. In the late 1960's and early 1970's, the DCRT Computer Center Branch (CCB) developed a multi-processor configuration for its computers that features a single batch job queue for all machines and shared disk files accessible to all machines. In addition, each machine ran one or more interactive systems to user terminals over telephone lines. There are currently over 600 ports to serve more than 2000 user terminals.

About 1972 CCB began looking at database management systems (DBMS's). It chose Product D as one that would run under this overall utility design and showed evidence of sufficient capability and vendor support. It tested and explored the DBMS by designing and developing a database application for information about its several thousands of users, students in its training courses, and the hundreds of terminals rented by users. An NIH contractor chose this same DBMS in 1976 to begin development of an automated NIH materiel management system for purchasing, inventory, and related functions. Several Institutes at NIH have subsequently used the DBMS for other database applications.

Product D has its own facilities for telecommunications support. Its major uses are all on-line database applications with data entry from CRT or typewriter terminals. The DBMS itself is supported by a small group in the CCB systems team covering software and telecommunication and by a small group in the CCB Program Support Section that monitors operation and is responsible for all DBMS file maintenance operations. Programming groups outside of CCB are responsible for developing specific database applications using the DBMS.

The major application of Product D is an NIH administrative database system under continuing development, segment by segment, over the last 4 years by the Data Management Branch (DMB) of DCRT, in conjunction with the NIH offices of Administrative and Financial Management. This project followed the initial contractor's efforts on the materiel management system. It is designed to handle on-line

entry of tens of thousands of line transactions from more than 100 decentralized terminals, some locally multiplexed. Key to the success of this application is an experienced, central Database Administrator (DBA) who reviews all programs and approves the database strategy used by all programmers. Product D supports interactive queries, but another commercial package generates reports more efficiently.

## ADVANTAGES of a DBMS

1. Phased project implementation in an organized manner
2. Application independence from file structure or organization
3. Fully synchronized data
4. Point of transaction data capture
5. Immediate availability of captured data
6. Formatting service for data entry at the page level
7. Full integration of all data
8. Elimination of multiple data capture of redundant data
9. Minimum of redundancy in programming
10. Excellent security
11. Excellent recovery and backup procedures already implemented

## DISADVANTAGES of a DBMS

1. Need to learn new DBMS software
2. Overhead of the DBMS
3. Training people to use the on-line DBMS systems and concern about the impact of a multi-user environment
4. Longer lead time in developing applications
5. Potential for becoming locked in to the DBMS
6. More complex maintenance of programs and systems
7. Limitation caused by need for support personnel to make DBMS available for longer hours on-line
8. Demand for new application support once others see impact of a database environment
9. Total reliance on the DBA function for integrating new applications and difficulty in providing adequate backup for this function

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| Part One | | |
| 1. Cantralization | | |
| 1.1 Extent | vsry much | very much |
| 1.2 Difficulty | soms | little or none |
| 2. Data Independence | | |
| 2.1 Reduced Effort in Application Developmsnt<br>The agency needs a more technically orientsd group to dsvslop applications around data. | much | some |
| 2.2 Easier Maintenanca and Modification<br>The maintenanca effort for daily systems which ars undergoing regular modification is about ths sams for database and non-databass systems. Savings are on ths databass side, not on the side of program maintenanca. | soms | littls or none |
| 3. High-level Queries | | |
| 3.1 Improved Timeliness and Accuracy<br>Timelinsss and accuracy result not from the DBMS query facility, but from centralized data with on-line updates. | much | much |
| 3.2 Reduced Need for Custom Programs<br>The agency has a separate file management system to handle overnight queries, most of which are batch and sequential. This product has virtually eliminated custom programming of requssts. | much | much |
| 4. Degradation of System Psrformance | | |
| 4.1 Overall<br>The DBMS does not degrade the total performance of the computer because of ths overall design and total job stream balancing across several CPU's. Other telecommunication systems operate on CPU's separate from the one used by the DBMS, with responss tims independent of the DBMS and telecommunications load. | little or none | little or none |
| 4.1.1 Response Tims<br>The average responss time is about fivs ssconds; ths worst is sevsn seconds. | little or none | littls or none |
| 4.1.2 Throughput<br>The DBMS handles 15-20 thousand transactions psr day. | little or none | little or none |
| 4.1.3 Number of Simultansous Users | little or none | little or none |
| 4.1.4 Other | little or none | little or none |

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| 5. Personnel | | |
| 5.1 Difficulty in Recruiting or Training<br><br>The Government's personnel process, which took about six months, caused the agency a lot of troubls in finding and hiring a DBA. For good programmers, the transition to DBMS applications is relatively smooth. | vsry much | soms |
| 5.2 Premium for DBMS Expertise | soms | some |
| 5.3 Increased Need for Technical Support | much | soms |
| 5.4 Amount of Vsndor Training | soms | littls or none |
| 6. Start-up Costs | | |
| 6.1 DBMS procurement<br><br>Start-up costs for procurement and installation of the DBMS itsslf were part of the overhead of the large NIH Central Computer Utility. They were of the order of one system- programmer year spread over two years. The major manpower costs resultsd from the implsmentation of continuing application support. | soms | some |
| 6.2 DBMS installation | some | some |
| 6.3 New support software | soms | some |
| 6.4 New hardware<br><br>The hardware (CPU) to support a DBMS already existed as part of ths multi-machine design of the computer center. The decision was made to use one machine for a DBMS rather than to base another facility on telecommunications. The major nsw hardwars costs wers for tsrminals and local multiplexers. | don't know | don't know |
| 6.5 New personnel<br><br>Very few new personnel were hired. About a dozen peopls (operations, systems support, systems team) have been reallocated to support the DBMS in the Computer Center itself. For the NIH administrative database system, the DBA was hired specifically for DBMS expertise and in ths end about eight applications programmsrs will have been trained and converted to using the DBMS for the total materiel and financial management systems. | littls or none | littls or none |
| 6.6 Training<br><br>Becauss of what the managers had hsard, they sxpectsd training to be very expensive, sspecially for a complex DBMS. Thsir estimates turned out high. | vsry much | some |
| 6.7 Reorganization<br><br>The application staff has reorganizsd thres times in sighteen months, and the manager expects ons mors. Rsorganization has cost some time, not a lot of money. | vsry much | very much |
| 6.8 Conversion of existing data files and Programs | soms | some |
| 6.9 Other | | |
| 6.10 TOTAL<br><br>Total application development costs for ths administrativs databass system amounted to about 1.3 million dollars over four years. This figure covered machins time, personnsl, and new terminals. | | |

Responses to the Questionnaire

National Institutee of Health

| QUESTION | EXPERIENCE |
|---|---|
| Part Two | |
| 1. Cost/Benefit Analysie | |
| 1.1 Amount Done<br>NIH did not do a full cost/benefit analysie for the adminietrative database eystem becauee the coste (personnel and dollare) were fixed and the problems were not primarily in reducing coete, but in finding a way to get acceptable computer support for the overall materiel and financial management functions at NIH. | eome |
| 1.2 Amount Necessary | little or none |
| 1.3 Difficulty in Identifying Coets and Benefite | some |
| 1.4 Difficulty in Quantifying Intangible Costs and Benefite | little or none |
| 1.5 Need of a Guideline | don't know |
| 1.6 Amount of Help Provided by Cost/Benefit Analysis in DBMS Procurement | little or none |
| 2. Requiremente | |
| 2.1 Extent of Coneultation with Ueers Before Procurement<br>NIH analyeed ueere' requirements for each application to eee whether the exieting DBMS ie appropriate. To date only a few centralized applicatione have been appropriate for the DBMS. Thie number will grow in the adminietrative area as other applicatione tie onto the central NIH adminietrative databaee eyetem. | eome |
| 2.2 Difficulty in Defining Application Requiremente<br>Relative to application development. | eome |

INTERVIEW REPORT

AGENCY 5

A few years ago Agency 5 acquired Product E for direct multiple access to files and ease of changing data files. The procurement involved one systems programmer, who considered another DBMS as well as Product E to run on the agency's existing equipment. The decision was intuitive rather than formal. Once Product E was installed on the existing equipment, it proved too slow to implement the target application.

After coming to Agency 5 two years ago, the manager we interviewed began to use Product E as a generalized DBMS rather than as a file access method. She had gone through an extensive evaluation of DBMS's for another Governmental agency, and from that experience she concluded that another DBMS and Product E were the most satisfactory DBMS's available for her agency's hardware. As a result, she was pleased with the intuitive choice of Product E.

To develop support for centralizing data, the interviewee continued to operate as the agency had until continuing discrepancies in data built a case for centralized data management. She then explained to management that integrating files could solve the problem. Together, she and another Agency 5 branch chief now oversee 14 or 15 database application systems, several of which share data. Their largest file, which is inverted on two fields, contains 2,000,000 records, each with about 106 characters. Another application is inverted on 58 fields, which is the maximum for any current Agency 5 file using Product E.

The interviewee identified two major problems in data processing at Agency 5. The first is the inability of end users to define their requirements ahead of time. When they want information, they make last minute requests that she could not fill without a DBMS. End users tend not to acknowledge overall responsibility for control of their data and other data processing needs. In many ways her ADP staff act as management consultants, leading users in analyzing and articulating requirements and then helping them to satisfy these requirements. Each office has a user representative who acts as liaison between the office and the ADP staff.

The ADP staff at Agency 5 are always under pressure to get something up, provide quick information, and then modify programs as needed. To prepare for last minute requests, they concentrate on designing the logical structures of the

database so that it is flexible enough to answer unforeseen questions. Logical database design is the second major problem the manager faces in her daily work. Doing it right means coordinating a number of activities that are not strictly data processing. She likes the DBMS because it insulates her from potential problems with users; she can often meet even their toughest demands.

This particular manager does not think cost/benefit analysis is necessary to justify a DBMS, whose benefits are obvious. She does think that agencies need a method to determine which DBMS to buy. She very strongly favors an "evolutionary" approach in implementing a DBMS. Evolution enables the ADP manager both to avoid the large start-up costs of the database approach and to demonstrate the benefits of a DBMS before trying to persuade managers and users that centralizing data is useful and cost effective. She had little trouble integrating applications at Agency 5 since she had previous experience in integration. She estimates that integration of files saves 30 percent of the time needed for developing new applications.

Training database personnel is a problem in the Federal Government. Since in Washington the Government cannot compete with the pay scales of private industry, most of its beginning programmers are upwardly mobile employees who come from other fields and do not have degrees or backgrounds in computer science. While these people often develop considerable skill in the mechanics of data processing, they sometimes do not understand data administration. It is very hard to hire anyone at a GS-12 level, and it is almost impossible to hire people with database experience. One former programmer, who had only 6 months experience as a systems programmer and another year in database applications, left Agency 5 to become the assistant DBA for another organization.

Agency 5 programmers do most of their reports with a report writer since the agency does not have Product E's user language. However, the vendor recommends that users of Product E avoid "coupling" files because of the overhead involved. As a result, all programs that require data from multiple files are written in COBOL.

Product E has been very reliable at Agency 5. It has not gone down in the 2 years that the interviewee has been with the agency, and it has not lost any data or indexes. It makes file design easy, and it enables the DBA to integrate files after the fact. While the DBMS provides security at the data element level, it does not provide security by value, which is also necessary to satisfy Agency 5's security needs. When the agency needs technical advice and

support, the ADP staff relies more on the Product E users' group than on the vendor.

Within the next 3 years the ADP staff will have to give users the ability to manipulate data. The manager is not sure how she will do it unless Agency 5 acquires a relational DBMS interface. The interviewee definitely thinks that relational DBMS's will soon take over the market, and she strives to keep current files in Third Normal Form or even Fourth Normal Form.

ADVANTAGES of a DBMS

1. Ability to relate files conveniently
2. Reduced maintenance of application software due to file changes

DISADVANTAGES of a DBMS

1. Higher complexity requires higher conceptual ability in staff.
2. Overhead on computer system

Responses to the Questionnaire

Agency 5

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| Part One | | |
| 1. Centralization | | |
| 1.1 Extent<br>The DBMS was acquired as an access method; the database and database administration were afterthoughts. | don't know | much |
| 1.2 Difficulty<br>The interviewee had previous experience integrating files for another agency; otherwise, the 'some' would have been 'very much.' | don't know | some |
| 2. Data Independence | | |
| 2.1 Reduced Effort in Application Development | don't know | little or none |
| 2.2 Easier Maintenance and Modification | much | some |
| 3. High-level Queries | | |
| 3.1 Improved Timeliness and Accuracy | don't know | some |
| 3.2 Reduced Need for Custom Programs | don't know | much |
| 4. Degradation of System Performance | | |
| 4.1 Overall | little or none | some |
| 4.1.1 Response Time | little or none | little or none |
| 4.1.2 Throughput (of batch runs) | little or none | much |
| 4.1.3 Number of Simultaneous Users | little or none | little or none |
| 4.1.4 Other | | |
| 5. Personnel | | |
| 5.1 Difficulty in Recruiting or Training | some | very much |
| 5.2 Premium for DBMS Expertise | little or none | don't know |
| 5.3 Increased Need for Technical Support | little or none | very much |
| 5.4 Amount of Vendor Training | some | some |
| 6. Start-up Costs | | |
| 6.1 DBMS procurement | little or none | little or none |
| 6.2 DBMS installation | little or none | little or none |

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| 6.3 New support software | some | some |
| 6.4 New hardware | little or none | much |
| 6.5 New personnel | little or none | little or none |
| 6.6 Training | some | very much |
| 6.7 Reorganization | some | little or none |
| 6.8 Convereion of existing data filee and Programe | little or none | little or none |
| 6.9 Other | | |
| 6.10 TOTAL | little or none | eome |

Responees to the Questionnaire

Agency 5

| QUESTION | EXPERIENCE |
|---|---|
| Part Two | |
| 1. Cost/Benefit Analysis | |
| 1.1 Amount Done | little or none |
| 1.2 Amount Necessary | little or none |
| 1.3 Difficulty in Identifying Coste and Benefits | don't know |
| 1.4 Difficulty in Quantifying Intangible Coste and Benefite | don't know |
| 1.5 Need of a Guideline | little or none |
| 1.6 Amount of Help Provided by Cost/Benefit Analyeie in DBMS Procurement | don't know |
| 2. Requirements | |
| 2.1 Extent of Consultation with Ueere Before Procurement<br>Ueers don't know enough about their long term requiremente to define them. | little or none |
| 2.2 Difficulty in Defining Application Requirements | much |

INTERVIEW REPORT

SOCIAL SECURITY ADMINISTRATION

The Social Security Administration acquired Product F in 1972 to administer its Supplemental Security Income (SSI) program. Passed by Congress in 1972 and scheduled to go into effect in 1974, the Social Security Amendments Act authorized SSA to provide emergency cash advances to needy people who met special qualifications. In order to distribute this supplemental income while protecting the Government from fraud, SSA needed an on-line, interactive database connected to all field offices. SSA employees had to be able to enter an applicant's name and qualifications, determine immediately if the person qualified for a cash advance, and check to make sure that SSA had not made any previous payments.

When the person in charge of data management for the SSI program surveyed commercial database management systems, only three were available to run on the agency's mainframes: Product F, and two others. He selected Product F for quick, interactive access on selected keys, not for centralized management of shared data. SSA did not have complex queries on inverted files, nor did it need a good ad hoc query facility. The agency's main requirement was fast interactive response to predefined queries. For most of these queries Product F needs only two seeks: one to a master file and a second to a related file. One of the other DBMS's did not provide the features SSA needed, and the other one was new enough that its vendor support seemed uncertain. To complete the SSI system, Social Security also procured 4500 cathode ray tubes (CRT's) for its 1500 field offices.

Social Security is still using Product F, although the agency now has larger and more powerful mainframes from the same vendor to process both SSI and Retirement and Survivor's Insurance (RSI). The RSI database contains 33 billion bytes, which constitute 35 million records; the SSI database contains 7 billion bytes and 5 million records. With databases of this size, space is a consideration: Social Security could not afford to have records with fixed fields. Because of the factors affecting elegibility, records in the SSI database range from 110 characters to 15,000 characters.

During the next 3-5 years Social Security is planning to acquire another DBMS as part of its effort to update the entire Social Security Administration Data Acquisition and Response System (SSADARS). In four stages SSA will replace the currently saturated host computers (limited

competition), replace the DBMS and communications monitor, rewrite SSADARS in a high-level, structured language like Pascal or Ada (70% of SSADARS is still assembly language), and replace the CRT's now in the field offices with new microprocessors. This changeover should result in a reduction of telecommunications cost by permitting personnel at the field offices to maintain records there and ship them to the National office at periodic intervals. The future system design will probably be a distributed network with local databases.

ADVANTAGES of a DBMS

1. Potential for eliminating redundant software
2. Potential for reducing technical staff
3. Potential for facilitating software maintenance

DISADVANTAGES of a DBMS

none

Responses to the Questionnaire

Social Security Administration

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| Part One | | |
| 1. Centralization | | |
| 1.1 Extent<br>SSA's two large databases have redundant data elements, but political conflicts made it difficult to centralize data. | some | some |
| 1.2 Difficulty<br>Political rather than technical difficulty. | very much | much |
| 2. Data Independence | | |
| 2.1 Reduced Effort in Application Development | much | much |
| 2.2 Easier Maintenance and Modification<br>SSA put an interface on top of its DBMS to provide a subschema for COBOL programs. | some | some |
| 3. High-level Queries | | |
| 3.1 Improved Timeliness and Accuracy<br>Predefined rather than ad hoc queries provide timely and accurate data. | little or none | little or none |
| 3.2 Reduced Need for Custom Programs | don't know | don't know |
| 4. Degradation of System Performance | | |
| 4.1 Overall<br>Dedicated system | little or none | little or none |
| 4.1.1 Response Time | little or none | little or none |
| 4.1.2 Throughput | little or none | little or none |
| 4.1.3 Number of Simultaneous Users | little or none | little or none |
| 4.1.4 Other | little or none | little or none |
| 5. Personnel | | |
| 5.1 Difficulty in Recruiting or Training<br>The interviewee had trouble convincing management that SSA needed staff familiar with database management as well as application programming. | some | much |
| 5.2 Premium for DBMS Expertise<br>Higher grades for DBMS skills. | much | much |
| 5.3 Increased Need for Technical Support | some | some |
| 5.4 Amount of Vendor Training | much | some |

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| 6. Start-up Costs | | |
| 6.1 DBMS procurement<br>$40,000 purchase price plus two months of the interviewee'e time. | some | eome |
| 6.2 DBMS installation | eome | little or none |
| 6.3 New support software | much | much |
| 6.4 New hardwars | little or none | some |
| 6.5 New personnel | some | eome |
| 6.6 Training<br>Four staff psople have taken two one-week oourses each. Ths first dovered application development; the second, database administration. | much | some |
| 6.7 Reorganization | much | little or none |
| 6.8 Conversion of existing data files and Programs | little or none | little or none |
| 6.9 Other | | |
| 6.10 TOTAL | some | some |

Responses to the Questionnaire

Social Security Administration

| QUESTION | EXPERIENCE |
|---|---|
| Part Two | |
| 1. Cost/Benefit Analysis | |
| 1.1 Amount Done | eome |
| 1.2 Amount Necessary | some |
| 1.3 Difficulty in Identifying Costs and Benefits | much |
| 1.4 Diffioulty in Quantifying Intangible Costs and Bsnefits | much |
| 1.5 Need of a Guidslins | much |
| 1.6 Amount of Help Provided by Cost/Benefit Analysis in DBMS Procuremsnt | much |
| 2. Requirements | |
| 2.1 Extent of Consultation with Users Before Procurement | much |
| 2.2 Difficulty in Defining Application Requiremente<br>Users are nontechnical. Though familiar with the exieting eystem, they had difficulties understanding how a DBMS would help them. Theee ueers would probably be better now at defining their DBMS requirements. | very much |

INTERVIEW REPORT

FOOD SAFETY AND QUALITY SERVICE
DEPARTMENT OF AGRICULTURE

The Department of Agriculture acquired Product G in 1975 for a departmental ADP facility, not for a specific application. After performing a requirements study, USDA evaluated the features of about 25 commercial software packages. Only Product G satisfied the mandatory requirements, and the Department acquired it through a sole-source procurement. They wanted software for direct access to data, ad hoc access, decreased programming, and faster turnaround. They also expected increased cost and complexity with a DBMS. All of these expectations have proved correct.

After several years of using Produt G, USDA reports general satisfaction with the system. On the negative side, complicated languages, commands, and structures plus a lack of database skills have limited usage of the DBMS. Personnel have been disappointed with the report writer and have experienced complications with multi-user environments. While Product G does not satisfy all users' requirements, it has nonetheless satisfied the original ones, and its query capabilities have continued to please departmental personnel.

Within the Department of Agriculture, the Food Safety and Quality Service (FSQS) uses Product G extensively. One application manager maintains six databases on the system. The Food Management Communication System consists of data published in the Meat and Poultry Inspection Directory and used to print mailing labels that go on USDA regulations and circulars. This central database of about 16 megabytes (covering about 10,000 establishments) saves thousands of dollars on return mail alone. Five regional centers have 2741 executive typewriter terminals that communicate with this database by telecommunications. The central office controls access to the data so that one region cannot affect the data of another region.

FSQS also maintains a Chemical Compounds and Packaging database that keeps track of approved chemicals near food producers. At Beltsville, staff members update this 16 megabyte database daily; during a month the changes (revisions, additions, and deletions) number about 1000. Each update generates a letter, and every month FSQS generates an 1800 page printout of the database. Using a report writer, the application manager needed only four instructions to produce this report, and he estimates that the DBMS has saved 25 square feet of files and 9 months of labor.

Before acquiring Product G, FSQS handled most of the information in these two databases with batch sequential processing of COBOL programs. Some of the data in the Chemical Compounds and Packaging database was not even automated. Although FSQS uses the DBMS for several applications, it has not centralized data. There is still a lot of duplication; for example, some identification data that is already in the Food Management Communication System is not available to batch sequential programs.

One drawback of the DBMS is that users need to know something about the structure of the database. Nonetheless, FSQS has not had any difficulty training people to use Product G. The agency has sent a couple of people to advanced training courses, but for the most part managers rely on their staffs to provide on-the-job training of new employees. Good manuals help, and the turnover in the agency is low enough to eliminate the need for very much training. Besides, most of the uses of the DBMS are repetitive: inserting, deleting, and modifying records in the same way.

FSQS has used Product G for modeling. Changes in the budget have made it necessary to realign areas, circuits, and regions. The DBMS helped the agency predict the effects of prospective changes on current operations. Modeling these effects worked, but it was difficult to do.

ADVANTAGES of a DBMS (FSQS)

1. Computer letter writing capability
2. Regional headquarters maintenance
3. No more need for customized programs
4. Little or no maintenance of application programs
5. Faster implementation
6. Access and update by most knowledgeable personnel
7. File always current
8. Documentation supplied by vendor
9. Uniform and standard approach -- training carries over to other applications
10. Extra computer processing cost more than offset by savings in personnel cost.

DISADVANTAGES of a DBMS (FSQS)

1. Storage costs
2. Telecommunications costs

## ADVANTAGES of a DBMS (USDA)

1. Direct access to data
2. Ad hoc access
3. Decreased programming
4. Faster turnaround

## DISADVANTAGES of a DBMS (USDA)

1. Cost
2. Complexity

Responses to the Questionnairs

Food Safety and Quality Service

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| Part One | | |
| 1. Centralization | | |
| 1.1 Extent<br>Geographic centralization, not centralization of files from different applications. FSQS still has a lot of redundancy, particularly of data in one database that is not available to batch sequential programs. | very much | very much |
| 1.2 Difficulty | much | some |
| 2. Data Independence | | |
| 2.1 Reduced Effort in Application Development<br>Users have to know something about the database. | much | much |
| 2.2 Easier Maintenance and Modification | much | much |
| 3. High-level Queries | | |
| 3.1 Improved Timeliness and Accuracy | very much | very much |
| 3.2 Reduced Need for Custom Programs | some | some |
| 4. Degradation of System Performance | | |
| 4.1 Overall<br>Degradation comes from the number of users of the computer, not from the DBMS itself. | some | some |
| 4.1.1 Response Time | some | much |
| 4.1.2 Throughput | some | |
| 4.1.3 Number of Simultaneous Users | much | much |
| 4.1.4 Other | don't know | don't know |
| 5. Personnel | | |
| 5.1 Difficulty in Recruiting or Training<br>Good manuals and good sub-managers eliminate problems with training. | some | some |
| 5.2 Premium for DBMS Expertise | much | much |
| 5.3 Increased Need for Technical Support | some | some |
| 5.4 Amount of Vendor Training | some | some |
| 6. Start-up Costs | | |
| 6.1 DBMS procurement | little or none | little or none |
| 6.2 DBMS installation | don't know | don't know |

Responses to the Questionnaire

Food Safety and Quality Service -- Continued

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| 6.3 New support software | don't know | don't know |
| 6.4 New hardware | don't know | don't know |
| 6.5 New personnel | don't know | don't know |
| 6.6 Training | don't know | don't know |
| 6.7 Reorganization | don't know | don't know |
| 6.8 Conversion of existing data files and Programs | some | some |
| 6.9 Other | | |
| 6.10 TOTAL | | |

Responsee to the Questionnaire

Food Safety and Quality Service

| QUESTION | EXPERIENCE |
|---|---|
| Part Two | |
| 1. Cost/Benefit Analysis | |
| 1.1 Amount Done | don't know |
| 1.2 Amount Neceseary | don't know |
| 1.3 Difficulty in Identifying Costs and Benefits | little or none |
| 1.4 Difficulty in Quantifying Intangible Coete and Benefits | little or none |
| 1.5 Need of a Guideline | don't know |
| 1.6 Amount of Help Provided by Cost/Benefit Analysie in DBMS Procurement | don't know |
| 2. Requirements | |
| 2.1 Extent of Consultation with Users Before Procurement | very much |
| 2.2 Difficulty in Defining Application Requiremente | don't know |

Responses to the Questionnaire

Department of Agriculture

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| Part One | | |
| 1. Centralization | | |
| 1.1 Extent | little or none | little or none |
| 1.2 Difficulty | some | much |
| 2. Data Independence | | |
| 2.1 Reduced Effort in Application Development | some | much |
| 2.2 Easier Maintenance and Modification | much | much |
| 3. High-level Queries | | |
| 3.1 Improved Timeliness and Accuracy | much | much |
| 3.2 Reduced Need for Custom Programs | much | much |
| 4. Degradation of System Performance | | |
| 4.1 Overall | | |
| 4.1.1 Response Time | some | some |
| 4.1.2 Throughput | some | some |
| 4.1.3 Number of Simultaneous Users | some | much |
| 4.1.4 Other (database load) | some | much |
| 5. Personnel | | |
| 5.1 Difficulty in Recruiting or Training | some | much |
| 5.2 Premium for DBMS Expertise | don't know | don't know |
| 5.3 Increased Need for Technical Support | much | very much |
| 5.4 Amount of Vendor Training | some | some |
| 6. Start-up Costs | | |
| 6.1 DBMS procurement | some | some |
| 6.2 DBMS installation | some | much |
| 6.3 New support software (teleprocessor) | some | very much |
| 6.4 New hardware | some | some |

Responses to the Questionnaire

Department of Agriculture -- Continued

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| 6.5 New personnel | some | much |
| 6.6 Training | some | much |
| 6.7 Reorganization | some | some |
| 6.8 Conversion of existing data files and Programs<br>Most applications are new. | some | some |
| 6.9 Other | | |
| 6.10 TOTAL | | |

Responses to the Questionnaire

Department of Agriculture

| QUESTION | EXPERIENCE |
|---|---|
| Part Two | |
| 1. Cost/Benefit Analysis | |
| 1.1 Amount Done | some |
| 1.2 Amount Necessary | some |
| 1.3 Difficulty in Identifying Costs and Benefits | little or none |
| 1.1 Difficulty in Quantifying Intangible Costs and Benefits | very much |
| 1.5 Need of a Guideline | some |
| 1.6 Amount of Help Provided by Cost/Benefit Analysis in DBMS Procurement | little or none |
| 2. Requirements | |
| 2.1 Extent of Consultation with Users Before Procurement | much |
| 2.2 Difficulty in Defining Application Requirements | some |

# INTERVIEW REPORT

## U.S. GEOLOGICAL SURVEY

Early in 1974 the U.S. Geological Survey wrote functional and operational specifications for its Ground Water Site Inventory System. It then contracted with the Department of the Navy to perform a requirements analysis for a data management system that would operate on its existing computer system. The study considered 76 software packages, including report generators, file management systems, DBMS's, and others. It determined that only two database management systems, Product H and one other, came close to meeting organizational needs. The Navy then benchmarked the two systems and recommended Product H, which USGS acquired late in 1974 through a sole-source procurement.

USGS has been very pleased with Product H for the past 7 years and has not encountered any significant software problems. However, the system startup required a large amount of vendor/contractor assistance. The system now resides on a modern computer system which has 16 megabytes of virtual memory.

The Ground Water Site Inventory System, the primary application which USGS maintains on Product H, consists of approximately 760,000 records and requires approximately 500 million characters of online disk storage. USGS field workers interactively update and query the database on a national basis via telecommunications with state field offices. The system handles about 300 queries per week and about 50,000 updates (new card images) per week. Prior to converting to Product H, USGS maintained the data on 20-25 different 80 column punch card types at the state field offices and periodically forwarded the information to the USGS national office in Washington, D.C. At that time the agency had no way to process the data it was collecting.

USGS has recently acquired the vendor's data dictionary to use with Product H and is currently implementing it to manage data resources. A contract for approximately $400 per month provides USGS with updates and revisions of Product H.

During the next 5 years USGS plans to convert to a new DBMS. The agency is now conducting a study to determine the feasibility of installing minicomputers in the state field offices to process scientific data before transmitting it over conventional telephone lines to the Washington, D.C. office.

ADVANTAGES of a DBMS

1. Data independence
2. Ad hoc query facility

DISADVANTAGES of a DBMS

1. It requires a large amount of on-line disk.
2. It is unsuitable for scientific data, which is often not hierarchical
3. It requires continuous training.
4. It does not reduce programmer overhead.
5. It degrades system performance.

USGS emphasized that these disadvantages reflect their experience with a single DBMS. They encountered many of these problems during a period (mid-1970's) when the DBMS itself was undergoing change, and their statements would not necessarily be true of all DBMS's or of the current state of the art.

Responses to the Questionnaire

U. S. Geological Survey

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| Part One | | |
| 1. Centralization | | |
| 1.1 Extent | much | much |
| 1.2 Difficulty | some | very much |
| 2. Data Independence | | |
| 2.1 Reduced Effort in Application Development | very much | much |
| 2.2 Easier Maintenance and Modification<br>Bigger problem than expected. Modification is always necessary. | some | much |
| 3. High-level Queries | | |
| 3.1 Improved Timeliness and Accuracy | very much | very much |
| 3.2 Reduced Need for Custom Programs<br>The DBMS is good for queries, but less so for reports. | much | some |
| 4. Degradation of System Performance | | |
| 4.1 Overall | some | much |
| 4.1.1 Response Time | don't know | little or none |
| 4.1.2 Throughput | some | much |
| 4.1.3 Number of Simultaneous Users | little or none | little or none |
| 4.1.4 Other<br>Disk storage requirements; files require 100% overhead. | little or none | very much |
| 5. Personnel | | |
| 5.1 Difficulty in Recruiting or Training<br>The constraint is the personnel ceiling, not the lack of trained applicants. | some | little or none |
| 5.2 Premium for DBMS Expertise | little or none | little or none |
| 5.3 Increased Need for Technical Support<br>Initially | much | very much |
| 5.4 Amount of Vendor Training<br>A huge training effort for both users and programmers. | some | much |
| 6. Start-up Costs | | |
| 6.1 DBMS procurement ($400,000) | very much | very much |

Responses to the Questionnaire

U. S. Géological Survey -- Continued

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| 6.2 DBMS installation<br>Put the DBMS up in about a day. | some | some |
| 6.3 New eupport software | some | much |
| 6.4 New hardware (diek etorage) | some | some |
| 6.5 New personnel<br>The DBMS does not provide personnel savings as promised. | little or none | much |
| 6.6 Training<br>About $100,000 per year. | some | very much |
| 6.7 Reorganization | little or none | little or none |
| 6.8 Conversion of existing data files and Programs<br>Conversion was expensive ($250,000) and time-consuming. Some data were unique in each of 50 eitee and therefore required custom programming. | some | much |
| 6.9 Other (contractor eupport) | some | much |
| 6.10 TOTAL | some | much |

Responsee to the Questionnaire

U. S. Geological Survey

| QUESTION | EXPERIENCE |
|---|---|
| Part Two | |
| 1. Cost/Benefit Analysis | |
| 1.1 Amount Done | very much |
| 1.4 Amount Neceeeary | some |
| 1.3 Difficulty in Identifying Costs and Benefits | much |
| 1.4 Difficulty in Quantifying Intangible Coste and Benefits | very much |
| 1.5 Need of a Guideline | very much |
| 1.6 Amount of Help Provided by Coet/Benefit Analyeis in DBMS Procurement | some |
| 2. Requiremente | |
| 2.1 Extent of Consultation with Users Before Procurement | very much |
| 2.2 Difficulty in Defining Application Requirements | much |

QUESTIONNAIRE REPORT

AGENCY 9

This report summarizes the fill-in-the-blank questions from Agency 9's questionnaire. We did not interview anyone from the agency's central computing facility.

About 10 years ago Agency 9 acquired Product I for its computer center. The agency performed a requirements study, a cost/benefit analysis, and a workload study before deciding to buy a DBMS rather than acquire new application development techniques. It also performed a detailed feature evaluation, ran a benchmark, and evaluated the costs of all five products whose vendors bid on the RFP. Only Product I satisfied the mandatory requirements.

ADVANTAGES of a DBMS

1. The ease in creating a database and developing an application
2. The availability of all the resources and capabilities at hand without having to develop them
3. The capability of modifying applications because of number 1 above
4. The powerful user language

DISADVANTAGES of a DBMS

1. Maximum usage of resources. However, this is the trade-off for efficiency.

Responses to the Questionnaire

Agency 9

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| Part One | | |
| 1. Centralization | | |
| 1.1 Extent | some | some |
| 1.2 Difficulty | much | little or none |
| 2. Data Independence | | |
| 2.1 Reduced Effort in Application Development | some | much |
| 2.2 Easier Maintenance and Modification | some | much |
| 3. High-level Queries | | |
| 3.1 Improved Timeliness and Accuracy | much | much |
| 3.2 Reduced Need for Custom Programs | much | some |
| 4. Degradation of System Performance | | |
| 4.1 Overall<br>System degradation is partly a function of the configuration and tuning. | some | some |
| 4.1.1 Response Time | some | little or none |
| 4.1.2 Throughput | some | little or none |
| 4.1.3 Number of Simultaneous Users | some | little or none |
| 4.1.4 Other | | |
| 5. Personnel | | |
| 5.1 Difficulty in Recruiting or Training | some | little or none |
| 5.2 Premium for DBMS Expertise | some | some |
| 5.3 Increased Need for Technical Support | much | some |
| 5.4 Amount of Vendor Training | some | some |
| 6. Start-up Costs | | |
| 6.1 DBMS procurement | some | some |
| 6.2 DBMS installation | some | some |
| 6.3 New support software | some | little or none |

| QUESTION | EXPECTATIONS | EXPERIENCE |
|---|---|---|
| 6.4 New hardware | little or none | little or nons |
| 6.5 New personnel | some | some |
| 6.6 Training | some | some |
| 6.7 Reorganization | little or none | little or none |
| 6.8 Conversion of existing data files and Programs | littls or nons | little or none |
| 6.9 Other | | |
| 6.10 TOTAL | soms | soms |

Rssponses to ths Qusstionnaire

Agency 9

| QUESTION | EXPERIENCE |
|---|---|
| Part Two | |
| 1. Cost/Benefit Analysis | |
| 1.1 Amount Done | some |
| 1.2 Amount Necessary | some |
| 1.3 Difficulty in Identifying Costs and Benefits | soms |
| 1.4 Difficulty in Quantifying Intangible Costs and Bsnsfits | soms |
| 1.5 Need of a Guideline | much |
| 1.6 Amount of Help Provided by Cost/Benefit Analysis in DBMS Procuremsnt | soms |
| 2. Requirements | |
| 2.1 Extent of Consultation with Users Before Procuremsnt | much |
| 2.2 Difficulty in Defining Application Requirements | some |

# APPENDIX D

## DBMS COST/BENEFIT PARAMETERS

The National Bureau of Standards has contracted with the University of Florida to develop specifications for a DBMS cost/benefit decision model. As part of its first report, the contractor has composed a list of parameters for requirements that can be important to any organization considering software for a data management system (DMS), which is more general than a DBMS. These parameters, which form a tree structure, can help the organization evaluate alternative ways to manage its data. This appendix presents and discusses those parameters that affect the selection of a particular DMS, whether it be a file management system, a report generator, or a full-scale DBMS. The purpose here is to present detailed requirements that Federal agencies may want to consider in evaluating different systems. It would be beyond the scope of this report to explain the theory of the prospective decision model or to provide a technique for converting the parameters into costs and benefits. Federal managers will find in this appendix some detailed material not available in either the chapter on theoretical costs and benefits or the chapter on actual Federal experience with DBMS's.

### Parameter List

The relevant parameters constitute a subtree of the complete parameter tree. They all belong under one node that is numbered 3 and labeled "Data Managment System Software." This node focuses on technical issues that agencies may want to consider in defining DMS requirements and in comparing alternative DMS products. The numbering scheme is the usual one for a tree, with each node uniquely identified by concatenating the numbers for all its ancestor nodes with the integer that identifies it as a unique child of its parent node. For example, node # 3241 (Immunity to Structural Changes) is the first node under 324 (Data Independence), which is the fourth node under 32 (DMS Functional Capabilities). Following the list of parameters, which shows the tree structure by indentations, is a brief discussion of each parameter. The discussion should help a Federal manager decide if the parameter in question is relevant to the agency's requirements for data processing.

3 Data Management System Software

31 DMS Support Software

311 DMS Sort Utility

3111 Data Considerations for Sorting
3112 Relationship with Other DMS Programs
3113 Sorting Algorithms
3114 Internal Sorting Technique

312 Report Generation Facility

3121 Report Formatting
3122 Built-in Programs
3123 Report Generation Language
3124 Job Summarization

313 Data Dictionary

3131 Type of Interface
3132 Data Resource Requirements
3133 Data Structure and Design Decisions
3134 Recording of Application Data

314 Schema Design Aids

32 DMS Functional Capabilities

321 Security

3211 Access Control
3212 Processing Restriction
3213 Threat Monitoring
3214 Privacy Transformation

322 Availability of Data

323 Integrity

3231 Concurrency Control
3232 Data Accuracy
3233 Data Consistency

324 Data Independence

3241 Immunity to Structural Changes
3242 External-Conceptual Mapping
3243 Provision of Changes to the Conceptual Model

325 Backup, Recovery, and Reorganization

3251 Mechanism for Backup
3252 Mechanism for Recovery
3253 User Services
3254 Data Reorganization

326 Data Model
327 Operating System Compatibility

33 DMS Language Interface

331 Data Definition

3311 Logical Facilities
3312 Physical Facilities
3313 Subschema Definition Facilities
3314 Implementation
3315 Ease of Use
3316 Familiarity

332 DMS Data Manipulation

3321 Language Capabilities
3322 File Retrieval Capabilities
3323 Host Languages
3324 Ease of Use
3325 Ancillary Features
3326 Familiarity

34 DMS Data Organization and Access

341 Record Implementation

3411 Variable Length Records
3412 Repeating Groups
3313 Hierarchical Records
3414 Compression Techniques
3415 Duplicate Keys
3416 Record Size Limitations
3417 Null Values

342 Implementation Requirements

3421 File Implementation

3422 Linkset Implementation
3423 Index Implementation
3424 DMS Size Limitations
3425 File Placement

343 Retrieval Operations

3431 Primary Key Retrieval
3432 Secondary Key Retrieval
3433 Scan
3434 File Navigation

35 DMS Advanced Features

351 DMS Tuning

3511 Storage Structures
3512 Use of Main Storage
3513 Physical Storage of Data

352 Data Restructuring

3521 Logical Schema Changes
3522 Physical Schema Changes

353 Portability

3531 Adherence to Standards
3532 Compatible Hardware
3533 Compatible Software

354 Compatibility with Existing System

3541 File Compatibility
3542 Program Compatibility

36 DMS History of Use

361 Age and Use of System

3611 Number of Years in Use
3612 Number of Users

362 System Growth Rate

3621 Revenue Growth Rate by the Product
3622 User Growth Rate

363 User Organizations

3631 Age of User Group
3632 Sharing Common Application Programs
3633 Sharing Systems Modifications

364 Documentation on Experience

3641 Written Reports
3642 Verbal Communications

37 Operation Mode of DMS

371 Batched Operation Mode of DMS

3711 Retrieval/Query
3712 Updates
3713 Verification of Updates

372 On-line Operation Mode of DMS

3721 Processing Options
3722 Search Aids
3723 Tutorial Features

DMS Support Software

The Software Support category includes all software which supports the DMS: sorting utilities, report generators, data dictionaries, and schema design aids.

DMS Sort Utility. The evaluation of a sort utility depends on several criteria. The limitations on data volume, record length, and file size in the database must correspond to limitations set by the sort utility. Because the programs that call the sort routine sometimes determine the source and destination of data and allocate space, they may determine what sorting methods are appropriate.

Report Generation Facility. A report generator formats the results of user queries or computations. It usually includes programs to manipulate mathematical data and to keep records on particular jobs -- e.g., rejected data, validated transaction data, invoked procedures, error conditions, and job statistics.

Data Dictionary. A data dictionary is a software system which stores data about the database. It is an information system in its own right and as such does one or more of the following:

- Stores descriptions of each data element: its source, users, ownership, relationships, controls, and physical storage method;

- Enables the DBA to update data about the database;

- Produces reports or provides answers to ad hoc queries about the contents of the dictionary database;

- Accepts file or database descriptions from a programming language and generates proper entries in the dictionary database;

- Makes parts of the dictionary database available to programs at run time; and

- Preserves the integrity and security of the dictionary database.

One important factor in evaluating the data dictionary system is the type of interface between it and the DMS. An active data dictionary will accept data descriptions from programs or provide them to programs; a passive system will do neither. Some data dictionaries are stand-alone packages; others are integrated with the DMS.

Schema Design Aids. Schema design aids take as input information on data elements, application requirements, access frequencies, etc., and produce a DBMS schema with logical or physical structures. The aid may be a collection of separate tools or an integrated design system.

## DMS Functional Capabilities

Functional capabilities of data management systems include security against unauthorized access, availability of data, integrity, data independence, and backup and recovery. These capabilities depend on the data model of the DMS and the DMS's compatibility with the host computer's operating system.

Security. DMS security involves the protection of computer resources from illegal access, modification, and destruction. Protection mechanisms allow resources to be shared without violating an individual user's right of privacy. Kinds of security facilities include controls on access to data or operations, encryption processes, and auditing software that maintains records of system usage.

Availability of Data. A DMS makes its data more easily available to users than other systems do. Multiple users of a DMS can simultaneously access all portions of the integrated database, and a data manipulation language allows convenient access to and manipulation of the database.

Integrity. "Integrity" refers to the accuracy and consistency of data in a database. To maintain integrity, some DMS's verify the accuracy and meaningfulness of entered data and protect the data against invalid modifications. If the system allows multiple users simultaneous access to the same data, it should provide some kind of concurrency control. A DMS may also specify some general constraints that the data values and relationships in a database must always satisfy.

Data Independence. Many DMS's provide some level of data independence to ensure the following:

- Immunity of application programs to changes in storage structure and access strategy, and
- Provision of multiple users' views of the database.

Backup, Recovery, and Reorganization. Backup, recovery, and reorganization are functions that enable a DMS to cope with hardware and software failures without losing data. Backup facilities may involve any of the following: 1) duplicating a database entirely or partially and applying all changes to both copies; 2) periodically dumping the database either totally or selectively; or 3) dynamically saving parts of the database by temporarily locking them out. A DMS can also recover data in a number of ways: 1) system recovery from a checkpoint; 2) backout and restart; or 3) selective save and restart. Finally, a data reorganization facility enables the DMS to improve access and response times by changing the physical storage of data.

Data Models.

A data model consists of the logical data structures and the operations on those structures provided by a DMS. Each of the three principal data models --relational, network, and hierarchical -- offers particular benefits and imposes specific constraints on the user.

Operating System Compatibility. Many DMS's run under particular operating systems. Users should recognize that running the DMS in another environment may require changes to either the DMS or the operating system. Common changes involve recompiling some DMS source programs and changing

certain parameters or defaults of the operating system. Other changes could be significantly more difficult.

## DMS Language Interface

The Language Interface applies to languages for data definition, data manipulation, and queries.

Data Definition Language. The data definition language (DDL) consists of commands for defining the data in a database. Several important factors in evaluating the Data Definition Language are its logical and physical facilities, its subschema definition facilities, its ease of use, and the user's familiarity with it.

Data Manipulation. A Data Manipulation Language (DML) consists of commands for loading, updating, and querying the database. In evaluating a DML, a prospective user may want to consider several factors:

- what kinds of search expressions it allows;
- what operations it provides for summarizing and aggregating data, and for grouping queries and temporarily storing results;
- which programming languages have DML extensions to provide program access to the database; and
- how easy it is to learn and use.

## Data Organization and Access

This category covers both the organization and storage of data on secondary storage devices and the means of accessing and modifying data.

Record Implementation. Important features of record implementation include DMS support of variable length records, repeating groups, hierarchical records, data compression, duplicate keys, and null values.

Implementation Requirements. Implementation requirements include physical storage techniques and size limitations. File implementation methods include hash-based, indexed-sequential, inverted, multi-list, B-tree, transposed, and unordered files. Linksets are structures which link records of one file to those of another; common linkset structures

include pointer arrays and ringlists. Index implementations provide efficient access to records on secondary keys. Limitations on the size and number of files and relationships may be critical to some agencies.

Retrieval Operation. There are four basic types of retrieval operations: primary key retrieval, secondary key retrieval, scan (i.e., accessing all records of a file), and file navigation.

## Advanced Features

The advanced features of a DMS consist of facilities used in tuning and restructuring a DMS to reflect the changing nature of a database or to improve performance. Users may want to consider the compatibility of a DMS with commercial hardware and software products, and with the existing DMS.

DMS Tuning. Tuning a DMS often means changing database structures or physical storage techniques to improve performance or to adapt to changing application requirements. Some DMS's provide utilities that make such tuning relatively easy.

Data Restructuring Facility. Most commercial database management systems allow little modification of an existing database structure beyond rearrangements of data fields, adding or deleting fields from records, or other relatively straightforward changes. To make major changes in logical storage structures, the user may have to redefine and reload the database. Any DMS utility that makes reloading unnecessary could be important to an agency with large databases or rapidly changing data requirements.

Portability. Portability is the capability of a software package to run in more than one computer environment. Commercial DMS's vary considerably in portability, and agencies contemplating future hardware procurements may want to consider carefully the portability of available systems.

Compatibility with Existing Systems. Incompatibility between present and future DMS's could cause large costs in converting files and programs.

DMS History

A DMS's history of use is an important indicator of its viability as a product. Its age, its customer base, its growth, and the availability of documents reporting experiences with the product indicate the likelihood both of an active user organization and of continuing support, maintenance, and enhancements,

Operation Mode of the DMS

Whether a DMS operates interactively or in batch mode is a major factor in its appropriateness for a given organization.

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

| U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET (See instructions) | 1. PUBLICATION OR REPORT NO. NBS SP 500-84 | 2. Performing Organ. Report No. | 3. Publication Date November 1981 |
|---|---|---|---|

4. TITLE AND SUBTITLE

Costs and Benefits of Database Management: Federal Experience

5. AUTHOR(S)

Jesse M. Draper

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

NATIONAL BUREAU OF STANDARDS
DEPARTMENT OF COMMERCE
WASHINGTON, D.C. 20234

7. Contract/Grant No.

8. Type of Report & Period Covered

Final

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

Same as item 6.

10. SUPPLEMENTARY NOTES

Library of Congress Catalog Card Number: 81-600152

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

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

The Federal Government has a large investment in a wide variety of database management systems (DBMS's) and in diverse applications using those systems. Data managers from eight Federal agencies report that they are pleased with the power, flexibility, and cost effectiveness of DBMS's, whether they use their system as sophisticated access software or make it the focal point of a large, integrated data processing system. While DBMS's may save an agency money in the long run, their major immediate benefits are almost always increased functionality and productivity, not reduced cost. Centralized data and high-level user languages reduce the time for developing and maintaining computer programs, and the latter enable nonprogrammers to use computers effectively.

The amount of cost/benefit analysis an agency needs before deciding to buy a DBMS increases with the complexity of the application. The experiences of the interviewed agencies, together with a structured list of cost/benefit parameters, should help Federal managers in understanding the potential value of DBMS technology and in defining their requirements for data management.

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

application development; computer software; cost/benefit analysis; database management systems; data management; program maintenance; requirements studies; software procurement.

13. AVAILABILITY

☒ Unlimited

☐ For Official Distribution. Do Not Release to NTIS

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

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

14. NO. OF PRINTED PAGES

106

15. Price

$5.00

# ANNOUNCEMENT OF NEW PUBLICATIONS ON COMPUTER SCIENCE & TECHNOLOGY

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

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

Name _____

Company _____

Address _____

City _____ State _____ Zip Code _____

(Notification key N-503)

*U.S. GOVERNMENT PRINTING OFFICE : 1981 0-360-997/1854

# NBS TECHNICAL PUBLICATIONS

## PERIODICALS

**JOURNAL OF RESEARCH**—The Journal of Research of the National Bureau of Standards reports NBS research and development in those disciplines of the physical and engineering sciences in which the Bureau is active. These include physics, chemistry, engineering, mathematics, and computer sciences. Papers cover a broad range of subjects, with major emphasis on measurement methodology and the basic technology underlying standardization. Also included from time to time are survey articles on topics closely related to the Bureau's technical and scientific programs. As a special service to subscribers each issue contains complete citations to all recent Bureau publications in both NBS and non-NBS media. Issued six times a year. Annual subscription: domestic $16; foreign $20. Single copy, $3.75 domestic; $4.70 foreign.

NOTE: The Journal was formerly published in two sections: Section A "Physics and Chemistry" and Section B "Mathematical Sciences."

**DIMENSIONS/NBS**—This monthly magazine is published to inform scientists, engineers, business and industry leaders, teachers, students, and consumers of the latest advances in science and technology, with primary emphasis on work at NBS. The magazine highlights and reviews such issues as energy research, fire protection, building technology, metric conversion, pollution abatement, health and safety, and consumer product performance. In addition, it reports the results of Bureau programs in measurement standards and techniques, properties of matter and materials, engineering standards and services, instrumentation, and automatic data processing. Annual subscription: domestic $11; foreign $13.75.

## NONPERIODICALS

**Monographs**—Major contributions to the technical literature on various subjects related to the Bureau's scientific and technical activities.

**Handbooks**—Recommended codes of engineering and industrial practice (including safety codes) developed in cooperation with interested industries, professional organizations, and regulatory bodies.

**Special Publications**—Include proceedings of conferences sponsored by NBS, NBS annual reports, and other special publications appropriate to this grouping such as wall charts, pocket cards, and bibliographies.

**Applied Mathematics Series**—Mathematical tables, manuals, and studies of special interest to physicists, engineers, chemists, biologists, mathematicians, computer programmers, and others engaged in scientific and technical work.

**National Standard Reference Data Series**—Provides quantitative data on the physical and chemical properties of materials, compiled from the world's literature and critically evaluated. Developed under a worldwide program coordinated by NBS under the authority of the National Standard Data Act (Public Law 90-396).

NOTE: The principal publication outlet for the foregoing data is the Journal of Physical and Chemical Reference Data (JPCRD) published quarterly for NBS by the American Chemical Society (ACS) and the American Institute of Physics (AIP). Subscriptions, reprints, and supplements available from ACS, 1155 Sixteenth St., NW, Washington, DC 20056.

**Building Science Series**—Disseminates technical information developed at the Bureau on building materials, components, systems, and whole structures. The series presents research results, test methods, and performance criteria related to the structural and environmental functions and the durability and safety characteristics of building elements and systems.

**Technical Notes**—Studies or reports which are complete in themselves but restrictive in their treatment of a subject. Analogous to monographs but not so comprehensive in scope or definitive in treatment of the subject area. Often serve as a vehicle for final reports of work performed at NBS under the sponsorship of other government agencies.

**Voluntary Product Standards**—Developed under procedures published by the Department of Commerce in Part 10, Title 15, of the Code of Federal Regulations. The standards establish nationally recognized requirements for products, and provide all concerned interests with a basis for common understanding of the characteristics of the products. NBS administers this program as a supplement to the activities of the private sector standardizing organizations.

**Consumer Information Series**—Practical information, based on NBS research and experience, covering areas of interest to the consumer. Easily understandable language and illustrations provide useful background knowledge for shopping in today's technological marketplace.

*Order the* **above** *NBS publications from: Superintendent of Documents, Government Printing Office, Washington, DC 20402.*

*Order the* **following** *NBS publications—FIPS and NBSIR's—from the National Technical Information Services, Springfield, VA 22161.*

**Federal Information Processing Standards Publications (FIPS PUB)**—Publications in this series collectively constitute the Federal Information Processing Standards Register. The Register serves as the official source of information in the Federal Government regarding standards issued by NBS pursuant to the Federal Property and Administrative Services Act of 1949 as amended, Public Law 89-306 (79 Stat. 1127), and as implemented by Executive Order 11717 (38 FR 12315, dated May 11, 1973) and Part 6 of Title 15 CFR (Code of Federal Regulations).

**NBS Interagency Reports (NBSIR)**—A special series of interim or final reports on work performed by NBS for outside sponsors (both government and non-government). In general, initial distribution is handled by the sponsor; public distribution is by the National Technical Information Services, Springfield, VA 22161, in paper copy or microfiche form.

**U.S. DEPARTMENT OF COMMERCE**
**National Bureau of Standards**
Washington, DC 20234

OFFICIAL BUSINESS

Penalty for Private Use, $300

POSTAGE AND FEES PAID
U.S. DEPARTMENT OF COMMERCE
COM—215


THIRD CLASS