A DBMS SELECTION MODEL

by

RICHARD WILLIAM FOX

B.A., University of Tampa, 1975

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Computer Science

KANSAS STATE UNIVERSITY
Manhattan, Kansas
1979

Approved by:

Elizabeth Unger

TABLE OF CONTENTS

ILLEGIBLE DOCUMENT

THE FOLLOWING DOCUMENT(S) IS OF POOR LEGIBILITY IN THE ORIGINAL

THIS IS THE BEST COPY AVAILABLE

THIS BOOK CONTAINS NUMEROUS PAGES WITH THE ORIGINAL PRINTING BEING SKEWED DIFFERENTLY FROM THE TOP OF THE PAGE TO THE BOTTOM.

THIS IS AS RECEIVED FROM THE CUSTOMER.

THIS BOOK CONTAINS NUMEROUS PAGES WITH MULTIPLE PENCIL MARKS THROUGHOUT THE TEXT.

THIS IS THE BEST IMAGE AVAILABLE.

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 Problem

There is, at the present time, no definitive guide for the selection of a Data Base Management System (DBMS) by a non-technically-oriented manager (Til, 78). The problem is thus the development of a model which will provide a basic instrument for use in selecting a DBMS; one that assumes at least a working knowledge of the use of computers at the management level, but that does not presuppose an extensive computer-science level background.

The model should allow the manager to apply weights or deviations to differentiate the intended application from the "norm", yet provide an adequate coverage of the many areas which need to be considered in the selection of a DBMS. The model should be of sufficient technical depth to allow the manager to ask appropriate and realistic questions of the vendor representatives, yet at a high enough level that a manager can use the model without a significant amount of outside aid.

The model must recognize that there are a very large number of DBMS packages presently on, or soon to be on, the commercial market and thus must give the manager the capability of eliminating -- with a minimal expenditure of time -- those candidate packages which are obviously not

suitable. It must, however, allow for an investigation in depth of those DBMS packages which appear to be reasonable and desirable alternatives.

Finally, realizing that the user of the model may be a middle-level manager who must justify the selection process for the final package to a higher-level manager, the model must provide at least minimal capability for quantification of the various selection criteria, plus enough supportive data on the selection methodology to allow the manager to adequately defend the product of the selection cycle.

1.2 Solution

The original concept entailed a single-phased selection model plus a follow-up user-questionnaire to aid in the validation of the model. Initial attempts to work with a one-level model proved unsatisfactory, since there was no provision for a minimal expenditure of time on obviously unsuitable DBMS packages and the dedication of the maximum time on those few DBMS's which appeared to best meet the user's needs.

The model evolved into a two-phased proposal: the first phase, designated the "filter", was designed to allow an almost cursory look at all of the available packages and a relatively rapid elimination of those DBMS's which were unsuitable; the second phase, which was designated the

"discriminator" was designed to allow an in-depth study of those packages which survived the culling process of the filter. As work continued on the two-phased approach, it became obvious that there was an insufficient base of data on the criteria applied by the "real-world" user in the selection of a DBMS for a commercial application. There was also a lack of data on those factors that a present DBMS user would consider, in perspective of the experience with a working DBMS, in the process of DBMS selection.

The prospective-user selection model was set aside and a concentrated effort made to 1) ascertain what factors influenced present users to select the DBMS they were now using, and 2) those factors which they would consider in the re-selection process, having gained familiarity with a DBMS in a working environment. A user questionnaire was designed by the author, and, after some restructuring and re-evaluation of the format as a result of discussions with the various committee members, the questionnaire was sent to twenty-one present users of DBMS packages.

In order to minimize the number of independent variables interacting within the selection process, the users to whom the questionnaires were sent were selected from the lists made available by the vendors of three major currently-available DBMS packages: MRI Systems Corporation [System 2000], Cullinane Corporation [IDMS], and CINCOM Systems, Inc. [TOTAL]. These three systems are listed in

the Auerbach Information Management Series on Data Base Management as three of the "top five" DBMS's currently on the commercial market, and they represent a relatively diverse data base organization: IDMS representing the DBTG-type network with CODASYL-type standardization; TOTAL representing a DBTG-type network without the rigid CODASYL standards; and System 2000 representing the hierarchical DBMS with major inverted-list structuring.

After a review of the responses to the questionnaire by the twelve users who returned the form, plus several telephone conversations with the Data Base Administrator or equivalent at the user sites, a moderate restructure of the prospective-user model was made. A few questions were found to be either inappropriate to the commercial user or of minor significance in the selection process. Several of the five-point rating scale descriptors were either reworded or changed in value, and two new questions were added to the model to fill in previous gaps in the user discriminator areas.

Further work with the revised model indicated that the second phase would have to be further broken down into subordinate parts. The model, at that point, did not adequately reflect the difference between user necessities and user desires or allow any intermediate-level rating of user requirement criteria. With the expansion of the model into an extended format, it was determined that the user

could not be expected to remember and compare comments provided in nearly twenty different areas. An intermediate compaction stage was developed to narrow the critical factors to be considered by the user down to five limited groups. This intermediate stage allows the user to limit the span of decision considerations, while providing a general explanation of the impact of various user needs on the selection process.

The final model thus evolved into a two-phased, seven part structure, as follows:

I. The Filter

A set of twelve questions to aid in the rapid elimination of DBMS's considered unsuitable for the specific user application

II. The Discriminator

A. User Requirements Section

1. Identification of User Requirements
2. Requirements Consolidation Form
3. Explanation Section [to explain the impact of various user requirements on selection criteria
4. Vendor Capability Form

B. System Features Section

1. Features Evaluation Form
2. System Comparison Form

1.3 Assumptions

1. The model is targetted for a user who has already made a decision to secure a DBMS. Nothing in the report is related to "selling" either the concept of computer use or the advantages of using a DBMS. The model presumes that the management is at least generally familiar with the DBMS concept, has made a decision to install a DBMS, and is now in the process of selecting the best DBMS for a particular application.

2. The model is intended for the manager who is not technically proficient in the details of computer hardware or software. It is specifically not designed to be of sufficient scope, technically, for the computer analyst who must conduct an in-depth technical research of various DBMS's. It assumes a minimal working knowledge of the use of computers at the management level. The manager of a data processing facility should be able to use the model, without the aid of technical analysts, either as an independent management-level selection tool, or as an aid in making a final selection from an analyst-prepared report on several DBMS's under consideration.

3. The model is system independent; it does not assume any particular hardware, nor does it consciously add weight to a particular DBMS package or engineering concept. It will allow managers with diverse vendor systems to assess the DBMS's presently available for use on their

present or proposed hardware -- with their present or proposed staff.

4. The model is intended to be used as a guide; it demands several value decisions on the part of the user in the assignment of weight values for areas considered key to the enterprise application. The manager is expected to be familiar enough with the intended application to logically and realistically determine the appropriate value for the various decision factors presented.

5. The present-user questionnaire chapter is provided as an indicator of what some present users of DBMS's have reported about the type of application in which a particular DBMS is being used, the relative merits of various DBMS features, and the value attached to selection criteria at the time of purchase and subsequent to a reasonable period of familiarization with the system.

6. To provide a reasonable set of boundaries to the area of study, many of the newer or peripheral concepts in data base technology were specifically excluded from the report, such as the concept of distributed data bases, back-end processors, DBMS's in the developmental stage, and consideration of RAP and other special interest areas not presently available to the general public.

CHAPTER 2
A DBMS SELECTION MODEL

2.1 Introduction

This Data Base Management System (DBMS) Selection Model is designed to aid the non-technical manager in the selection of the most appropriate DBMS for a specific application. The model consists of two phases: Phase I is a rapid evaluation "filter"; Phase II is a detailed "discriminator". The user should be able to eliminate the majority of DBMS's on the commercial market, with a minimal expenditure of time, by use of Phase I. All of the DBMS's which survive the filtering process of Phase I are then evaluated in Phase II in an attempt to make a final selection among the limited number of alternatives which appear to be suitable.

Passage through the filter is primarily a "go/no-go" proposition; any DBMS which does not meet the minimal criteria established in the filter is dropped from all future consideration. The second-phase discriminator section allows the user to tailor the DBMS requirements specifications to a specific intended application and provides a guide for the quantification of user-specific requirements to be applied in the final selection process.

It is expected that the user has had some management experience with computers in the intended application

environment and is familiar with the more common computer management terminology. Prior to beginning the selection process via the model, the user must be familiar with at least the following areas for the intended application:

A. Present or proposed "mainframe" hardware
B. Present or proposed auxiliary memory devices
C. The compiler language capability available
D. The skill level of prospective end-users, analysts, and system maintenance personnel
E. If applicable, the absolute capital budget ceiling for the intended application.

The general selection process consists of two phases and a total of seven different areas of consideration. It is suggested that the user study the model flowchart [See Figure 2.1] and the flowchart explanation prior to the application of the model.

2.2 Explanation of Model Flowchart

The manager may apply the filter to all DBMS's which appear to be reasonable alternatives for the intended application. Any DBMS's which do not meet all of the criteria within the filter are eliminated from further consideration; all DBMS's. All DBMS's which satisfy all of the criteria in the filter are passed to the discriminator.

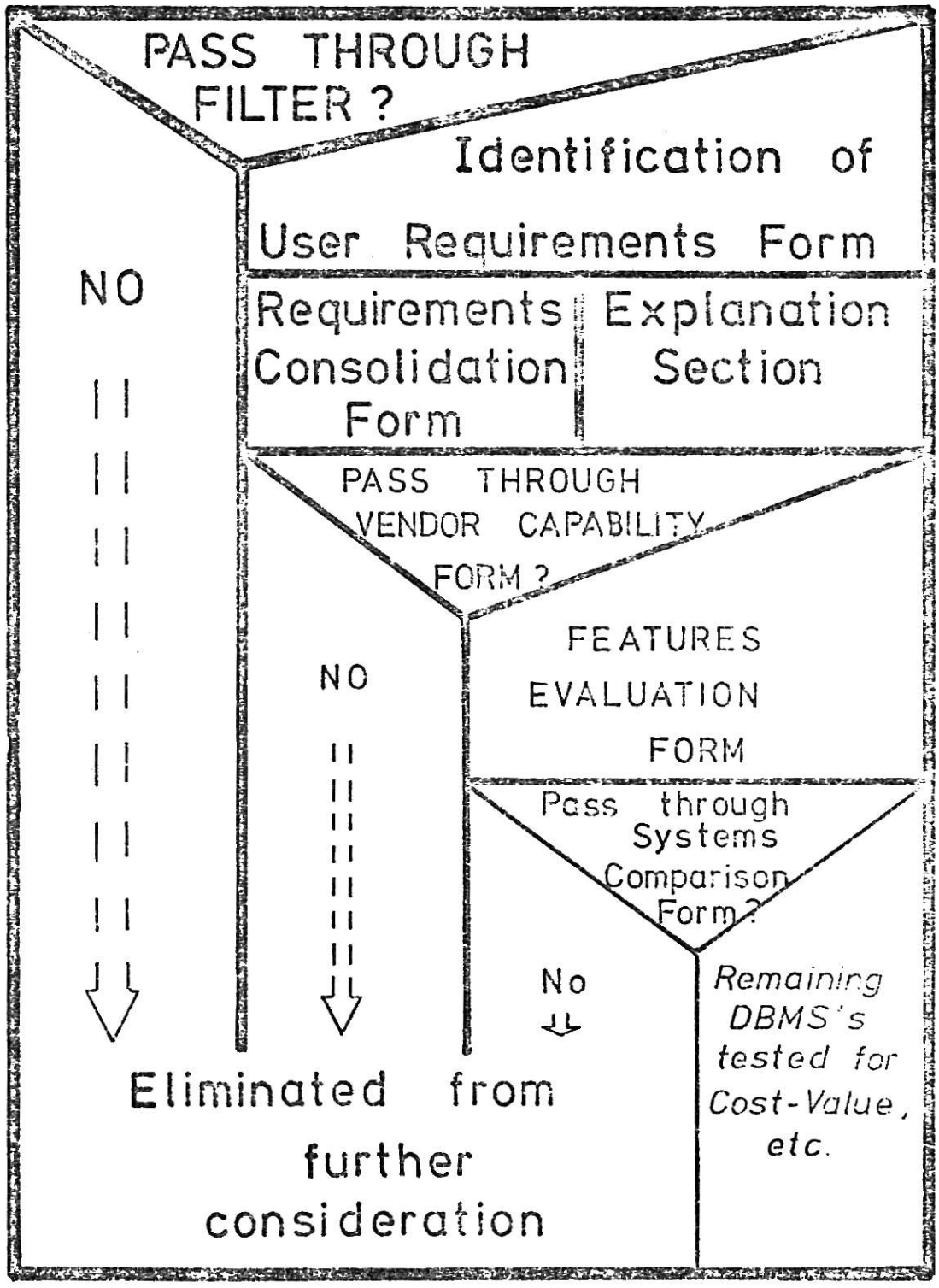

Figure 2.1 MODEL FLOWCHART

The discriminator phase consists of two general areas of interest: the User Requirements Section and the System Features Section. The User Requirements Section allows the manager to tailor general specification characteristics to the application and then -- through the use of the Vendor Capability Form -- to eliminate those DBMS's which do not meet the minimum tailoring specifications. The System Features Section provides the manager with a formatted evaluation scheme by which desired system features may be assigned values and then the vendors' capability to comply with user demands is rated using the manager-tailored System Comparison Form.

All DBMS's which have not been eliminated by the model may then be submitted for a final cost-value analysis if appropriate to the application.

2.3 Model Application Alternatives

The manager may choose to apply the model in one of two alternative manners:

A. Independent Manager-level Selection Process

If the manager intends to conduct the entire selection process, the model should be used in its entirety. After securing a list of those DBMS's currently available [Slo, 77; Tan, 75; Boy, 77; Com, 78], the manager can apply the filter to eliminate all unsuitable candidate DBMS's. The discriminator is then used to provide a framework for

an in-depth study of those DBMS's which pass through the filter phase of the model.

B. Decision Aid Process

If the subordinate-level analysts have made the initial study of DBMS's available for use in the intended application, the manager may delete the entire filter process. Phase II may be used to provide the framework for a decision study for final selection of a DBMS from a limited number of candidate packages. Since the manager can apply application-specific values to the various selection criteria, the use of Phase II will provide the manager with some measure of quantifiable support for the presentation of the final recommendation at the higher management level.

2.4 The Filter (See Figure 2.2)

A. Hardware

Is the system presently in use, in a commercial environment, on your vendor mainframe ?

If you now have, or are planning a change to one of the IBM 360/370-series mainframes, there is a relatively wide range of available DBMS packages. If your mainframe is from another vendor, or if you have older equipment or

The FILTER

		VENDOR / SYSTEM IDENTIFICATION NUMBER									
HARDWARE	SYSTEM IS PRESENTLY IN USE, IN A COMMERCIAL ENVIRONMENT, ON YOUR TYPE OF MAINFRAME										
	SYSTEM IS COMPATIBLE WITH THE LANGUAGE COMPILER ON YOUR PRESENT SYSTEM										
	SYSTEM REQUIRES LESS THAN ____k OF MAIN OR AUXILIARY MEMORY (YOUR MAXIMUM AVAILABLE)										
STAFF	ACTUAL END-USER SKILL REQUIRED IS WITHIN THE CAPABILITY OF PERSONNEL WHO QUERY DATA BASE										
	SYSTEM MAINTENANCE SKILL REQUIRED IS WITHIN CAPABILITY OF YOUR SYSTEMS STAFF										
	TOTAL MAN-HOURS PER MONTHS REQUIRED TO MAINTAIN THE SYSTEM										
CAPITAL	MINIMUM PURCHASE / LEASE PRICE FOR BASIC PACKAGE FOR YOUR APPLICATION										
	ANNUAL MAINTENANCE -- IF MAINTENANCE CONTRACT AVAILABLE FROM VENDOR										
	ADDITIONAL EXPENSE FOR PERIPHERAL EQUIPMENT OR TRAINING OF PERSONNEL FOR APPLICATION										
VENDOR	THIS IS MAIN LINE BUSINESS FOR VENDOR -- NOT MINOR SIDELINE -- OR ALL NEW SYSTEM LINE										
	VENDOR HAS BEEN IN BUSINESS FOR MORE THAN FOUR YEARS										
	VENDOR LISTED -- WITH HIGH CREDIT RATING -- IN DUNN & BRADSTREET, LOCAL BBB , ETC.										

Figure 2.2

are planning to use a mini-computer, this question is one of the primary eliminators of the majority of candidate DBMS packages.

In order of discriminatory difficulty, the question is three part: is the DBMS compatible with your type of vendor mainframe; presently running on that mainframe; and meeting acceptable demands in a commercial environment. It is not uncommon for smaller software houses to advertise compatibility on equipment based on a carefully-controlled limited-user, laboratory simulation. Some vendors have advertised compatibility for hardware for which few other systems are available when they believed that they could get their package to work on that hardware. Unless there is sufficient time and money to subsidize this research, demand the names of commercial establishments who are presently using the DBMS on the intended application type hardware -- preferably for the same type application.

Does the system demand peripheral disk or drum memory ?

This is oriented at the smaller user who may not have disk or drum peripheral memory devices. If the intended system uses magnetic tape or any of the floppy disk units, demand the name of a commercial user who is using the DBMS with that specific type of device. Any system which

advertises "rapid retrieval" [less than five seconds] can be expected to demand a random access memory [RAM] type peripheral device -- or a very large amount of add-on main memory capability. If the DBMS is to be used in a multiple user, simultaneous operation, or there is a demand for rapid retrieval on any but the very smallest size data base, the system will almost certainly require a RAM device to avoid tedious search routines and extended search times.

Is the system compatible with the compiler language you now have ?

If there is a significant present investment in active programs or if the programming staff would require an extensive retraining program in a new language to work with a particular DBMS, the inability of that DBMS to interface with the current language should be cause for elimination from future consideration. The conversion costs for data and storage media, plus the application programs, is likely to outweigh any benefits derived from the DBMS -- unless the acquisition of a DBMS is a planned portion of a major system upgrade or overhaul. Efficiency of retrieval would be seriously hampered by any interface package or by any "patching". Select a DBMS for which the vendor can provide a list of users who are limited to your present compiler languages.

System requires less than ---k of main or auxiliary memory (with your maximum space as the criterion).

Calculate the maximum memory storage space which can be realistically devoted to the DBMS on a full-time basis. Inverted lists, chains and rings tend to expand as the user becomes familiar with the capability of the DBMS; make sure that there is sufficient memory space for the expected future expansion of system-required pointers and links as well as the application data expansion.

Binary-search systems may demand that all of the data be maintained in core to reduce retrieval times; the system may degrade to near-uselessness if it is required to look up indices in main memory and transfer control to auxiliary memory devices to locate the actual data. The constant switching of control between main memory and auxiliary memory devices frequently results in "thrashing", which can nearly stop the processing of data for the user. Of note, however, is that the relative memory costs are shrinking rapidly. If the intended application is for a small data base which will expand over a period of years, simple memory requirements may be more cost-effective on a "sliding-scale" acquisition plan: use a smaller memory now and add on memory in a few years when it is much cheaper.

B. Staff Experience

Actual end-user skill required is within the capability of personnel who will query the data base.

You must know the lowest level of skill likely to be possessed by a potential end-user in the application for which the DBMS is to be used. If only analysts will access the data base, even in the future, the end-user skill required may be realistically allowed to stay high. If there is the possibility of allowing personnel who do not have an extensive computer background to have access to the system, an English-like, simple and straight-forward query language facility may be a requirement. If you are shown a "new" query language, check with present users for system efficiency. If the query language demonstration is part of the vendor sales pitch, demand a demonstration of a non-standard query; look at the information available in the sample data base, and ask for a response which would not be one of the standard vendor-prepared replies.

Even if your application end-users are relatively unsophisticated as to computer experience, if there is a limited set of queries which will always satisfy demands placed on the system, there may not be a need for an expensive "user-friendly" language facility. If there is the intent to use the DBMS to solicit non-standard data

combinations from the data base in order to aid in some decision process, it may be best to limit selection to a very high-level query language or a relational-type data base.

System maintenance skill required is within the capability of the present staff.

This area is almost totally application-dependent. If the application demands a large number of insertions or deletions of data in a relatively short time period, find out the skill level required for optimizing the insertion process or compacting the data base after a series of deletions. If there is a requirement for retrieving some data very frequently and maintaining some other data for only infrequent retrieval, determine what skill level is required for optimization of data storage location. If there is a fluctuating requirement for frequency of data retrieval for the same set of data, determine the skill level to effect an algorithm whereby data can "migrate" from quicker-access areas to slower-access areas or the reverse. (A good example of this requirement would be an airlines reservation system, which demands a moderately low frequency of retrieval for flights next month; a very high frequency for the same flight "tomorrow"; and an almost non-existant need for retrieval after the flight has been

completed.]

Total man-hours required per month to maintain the system.

Because the filter is arranged in a relatively larger-to-smaller-mesh order, it is unlikely that you will have a significant number of systems to evaluate by this point, so talk to some users of the DBMS. Contrary to general memory costs, labor costs are rising rapidly, and an extensive maintenance requirement is likely to become a major cost-effectiveness factor in future budgetary considerations.

C. Capital Expenditure

Minimum purchase or lease price for the basic package for your specific application.

At the filter level you should not be overly concerned with exact price comparisons or specialized add-on features pricing, but you must be aware of the budget limitations under which the selection process is being carried out. This question does not deal with the cost-value or even cost-effectiveness consideration; it is oriented at the elimination of any DBMS which is totally unsuitable for the intended application as a result of major differences between the minimal system cost and the maximum cost

ceiling for the application.

Annual maintenance costs, if contract available ?

Can a maintenance contract be purchased from the vendor, and, if so, at what cost? Many of the smaller vendors sell a software product for which service is not available. Evaluate the staff which will support the DBMS in the application environment: if the staff can maintain the system without any help from either the vendor or the software designer, the maintenance contract may not be a necessity. If there is no maintenance contract, however, make sure that there is a contractual provision for your receipt of a copy of the source code so that an in-house maintenance system can be developed.

Additional expense for peripheral equipment or training of personnel for the application ?

Does the system demand knowledge of RPG or a similar report generator in order to get a formatted output? Does it interface with the peripherals persently available, or will there be a need to purchase additional hardware to accommodate the system? If the system is a CODASYL-type DBMS, are the application personnel sufficiently familiar with COBOL that they can easily adapt to the COBOL-like

structure, or will there be a major retraining requirement to satisfy the language-related requirements of the system?

D. Vendor

This is a main line business for the vendor, not a new line or a minor side line.

Software production is not limited to the "normal" computer hardware firms. A tractor company with a large computer installation may decide to market a DBMS designed by its own DP staff. The DBMS may be a very good one, and it may be reasonably priced, but the potential buyer must be aware that the software package is only incidental to the main business of making tractors. A failure to achieve satisfactory sales in the software market may result in the dropping of all interest in the DBMS. Do not look at the parent organization size, but at the size of the enterprise which provides the DBMS service. Even an established software house may have a relatively untested product which proves to be unacceptable on the commercial market. Even if the DBMS is a product of a major hardware vendor, there is no guarantee that the DBMS will continue in the product-line. The large hardware vendor has the major vested interest in the hardware products; if it is determined that a new DBMS is both more efficient and demands four more disk drives than the previous data base package, the old line of software may be quickly dropped.

Vendor has been in business for more than four years.

This discriminator is based on the relative failure rate of new businesses. It is very difficult for a new business to weather the first few years. Any business that has remained in existence for more than four years has a higher probability of remaining in business than one which is just starting out. Again, however, this is application-dependent: if the staff is fully capable of debugging any problems in the future, and the source code is available, a new business may provide an efficient and less-expensive alternative for your application. If there is any doubt, at least consider an agreement whereby a third party that is acceptable to you holds a copy of the source code -- to be released to you in the event of a business failure by the vendor.

Vendor listed with high credit rating in Dunn and Bradstreet, local Better Business Bureau, etc.

This is a different approach to the same problem of dealer reliability, and the decision as to the relative importance of a good credit rating on the part of the vendor is based on staff capability. If the staff can maintain and debug a system that is available for half the price of the average major-vendor package, it may be a good

bargain. Consider the additional problem of loss of service through civil action by another firm if there is any doubt about the reliability of the vendor with whom you are dealing. If a theft or infringement of another vendor's software is proven, you may be faced -- at a later date -- with the option of surrendering the DBMS or buying the system rights to the system you have been already using from the lawful owner of the software.

2.5 Explanation of Terminology

A. Introduction

An implicit assumption of the DBMS selection model is that the potential user of the model is at least moderately familiar with general hardware terms -- bytes and words, disk or drum auxiliary storage devices, relative and direct machine addressing, etc. It cannot be assumed, however, that the prospective user is totally familiar with more advanced concepts in the data base area or the newer field of Data Base Management Systems. One of the primary target users of the model is the manager who has a minimal level of experience with a DBMS.

This section describes in general terms the three broad classifications of data base organization -- network, hierarchy and relational -- and explains some of the higher level problem areas associated with each organization. If the reader is familiar with DBMS, this section may be

skipped, although a light scan would confirm that the reader and author are using terminology in a compatible sense. If the reader is a new DBMS user, this section should be studied in depth; it is not intended to be a comprehensive definition of data base or DBMS terminology, but it may point out areas with which the reader is sufficiently unfamiliar to warrant some further study prior to the use of the "discriminator" phase of the selection model.

B. File Management vs. DBMS

A description of the data base sometimes given is that of a "pool" of information in which the various users go fishing for information. For a casually-interested observer this definition may suffice, but, because of the demand for a logical structure for the data imposed by the storage and retrieval operations conducted by the computer, a more definitive concept is required by the manager.

In the data base not serviced by a DBMS, each user would access the data pool through some set of common access methods. Any user familiar with the common access method set could access any data residing in the data pool. There is a total lack of data security or independence, plus an obviously unacceptable assumption that every person who needs access to the data is familiar enough with the computer structure to select the appropriate access method for that specific data.

The DBMS is an intermediary; it bridges the gap between the user and the way in which the data is actually stored in the computer. It consists of a set of procedures and data structures which isolate the user application from all of the computer-operator details of storage, retrieval security and hardware characteristics. The DBMS allows a person unfamiliar with computers to access information held in the computer with a minimum of knowledge. It also provides a set of safeguards that preclude users from accessing information which should not be released to them. In a graphic representation, the common access method and the DBMS method would look like Figure 2.3 below:

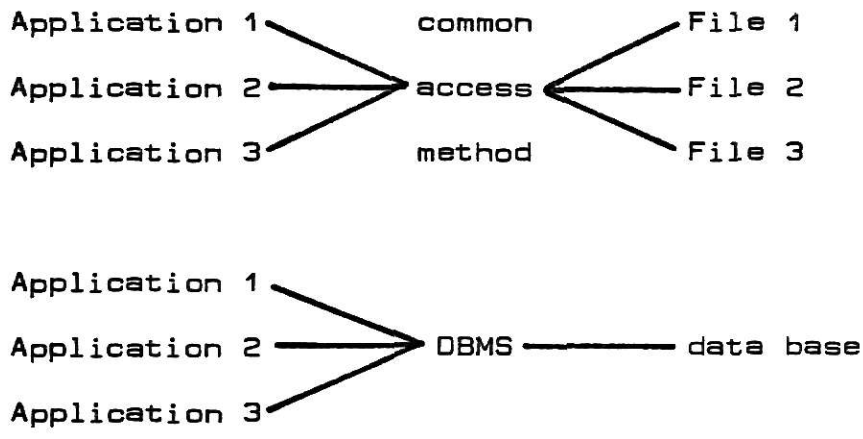

Figure 2.3

The DBMS introduces the concept of data independence from the physical memory organization; the user does not

need to know anything about internal computer architecture or computer data logic. This physical data independence makes it easier for a user to access data, while hiding data which is unauthorized for that user.

C. General DBMS Organization

Any DBMS provides some measure of physical data independence, but the manner in which the logical view of the data is presented to the user differentiates the types of DBMS's into two different conceptual categories. In the first category, the computer looks at the actual content of each record and retrieves every record which contains data matching some user-defined set of criteria; in the second category, the computer traverses a set of links or pathways through the records which has been previously mapped by the user and retrieves the data at a user-specified location, regardless of the content of the data at that location.

The first category, computer search of actual data content of each record, is normally called a relational data base architecture; the retrieval of data is predicated on some relationship between the searched-for record and some user-defined guidelines. The second general type of architectural category, in which the computer traverses some pre-defined path and retrieves the data residing at a user-specified location, is more complex for the user. It demands some understanding of how information is

physically stored in the computer, how the computer jumps from record to record (traverses the interconnecting paths) and the ability to initially construct the paths so that there is some correct route through the data base to all of the required data.

Martin uses a taxi driver and passenger analogy to explain this relationship. If the passenger desires to go to the movie "the Godfather", the only command necessary for an experienced driver would be, "Take me to 'the Godfather'". Even if the movie had been changed to another theater, the passenger would not need to be aware of that information; the driver would find the correct theater. With a less-knowledgeable driver, the passenger might be required to know the theater where the movie was presently playing and ask for the specific theater; the driver would then find the address for that theater. With a third driver, the passenger might be required to look up the address of the theater and tell the driver to go to a specific street address. Finally, with the totally inexperienced driver, the passenger would be required to give full directions: "Turn left at the next corner ..... drive five blocks .... turn right .... etc." [Mar, 78]

The passenger in the analogy is the application programmer or user, and the driver is the software which interprets requests to the data base and finds the "address" of the data desired. In the relational data base

the user need only know what type of data is desired; the DBMS will take the appropriate actions to find the record in which the data is currently residing. In the second architectural type of data base, the passenger -- or user of the data base -- must have, as a minimum, the name of the memory area in which the data resides. At the most inefficient DBMS level, the final driver version, the user must have the ability to specify the entire linkage path through the data base to the physical storage location of the data desired.

This second general type of data base, in which the user must have some knowledge of the manner in which the data is stored, is normally broken down into two named types of data bases, with the type based on the structure of the linkage paths. If the data is arranged so that there is generally only one path into a record area, the data base is said to have a "tree" or "hierarchical" structure. If the data is arranged so that there may be many paths into any given record area, the structure is called a "network" or "plex" structure.

D. The Relational Data Base

Probably the most familiar way in which data is represented to a viewer is the table or "flat file". The relational data base uses a tabular or flat file format, with the modification that each "column" must have a

distinct name and that no two "rows" may contain the exact same data. This last requirement, the distinctive or unique row, is based on the "entity" concept central to the relational method of data retrieval. Since the relational model looks at data content, not at record identification or address, the only way a record can be unique is for the combination of data contained within the record to be unique. An example of a simple relational data base would be as shown in Figure 2.4 below:

Part Number	Item	Color	Price	On Hand
334-4587-2	Widget	Blue	$1.45	2,000
334-4587-4	Widget	Green	$1.32	450
334-4594-6	Fangle	Blue	$1.45	1,800

Figure 2.4

Many relational data bases use a "user-friendly" set of English-like commands to access data. Through a series of either "relational algebra" or "relational calculus" operations, the software associated with the DBMS allows the user to make queries based on the relationships between the data desired and the content of the records stored in the data base. For the simple data base shown in Figure 2.4, the user could ask: "Give me all the information about Part Number 334-4587-2"; "Give me all of the Part

Numbers of Items whose Price is over $1.40; or "How many Widgets are presently On Hand?" The relational DBMS offers a capability for flexibility not matched by any of the other systems, since no user needs to be aware of the connections between records, but only of the data name about which information is being requested. The relational DBMS thus has the capability of offering the user the freedom to make almost "free-form" requests: "Give me the Item name of all Items which are Green, with a Price over $1.75 and with a present On Hand value of over 1500." The user-level simplicity is the major advantage of the relational data base; no other type of data base can offer this user-friendly capability for queries without a complex query front-end package as an optional feature.

The relational data base has, however, a significant disadvantage: the extra storage space requirement demanded as a result of the need for every record to contain all of the pertinent information. If there were five hundred different combinations of colors, prices and parts numbers for "Widgets", every unique combination record would need the sotrage space allocated for the repeating name of "Widget". If the parts inventory of a large concern contained a hundred thousand line numbers, and the space allocated for the item name was only thirty characters, the extra storage required for the redundancy of that single item name would be, for a standard mini-computer disk drive

system of 10 megabytes, two extra disk drives. Additional redundancies could normally be expected in a working data base, further increasing the demand for extra storage space.

The logical layout for a relational data base is at the most simplistic level, and the skill level required of the programmer or user is relatively minimal; the extra storage space demanded for a large number of different types of items may be prohibitive. Although there are no relational DBMS's available on the commercial market at the present time, primarily as a result of the tremendous overhead costs, the cost of memory is decreasing rapidly and the relational data base is felt by many to be the data base of the 1980's. And, as pointed out by Stocker, as the data redundancy is raised [value is stored instead of the address reference], searching becomes cheaper [Nij, 77]. For a firm which spends a major portion of the computer time searching the data base, particularly if requests are unique and not a repetitive standard query, the decreasing costs of memory are likely to make the relational data base the best choice in the near future.

E. The Hierarchical Data Base

The hierarchical model can be visualized as an upside-down "tree". The trunk of the tree is the general data base; the major branches of the tree are the major

subdivisions of the data base; the smaller branches which radiate from the larger branches are the records contained in each subdivision; and the leaves are the data areas contained in each record. A simple hierarchical data base could look like Figure 2.5 below:

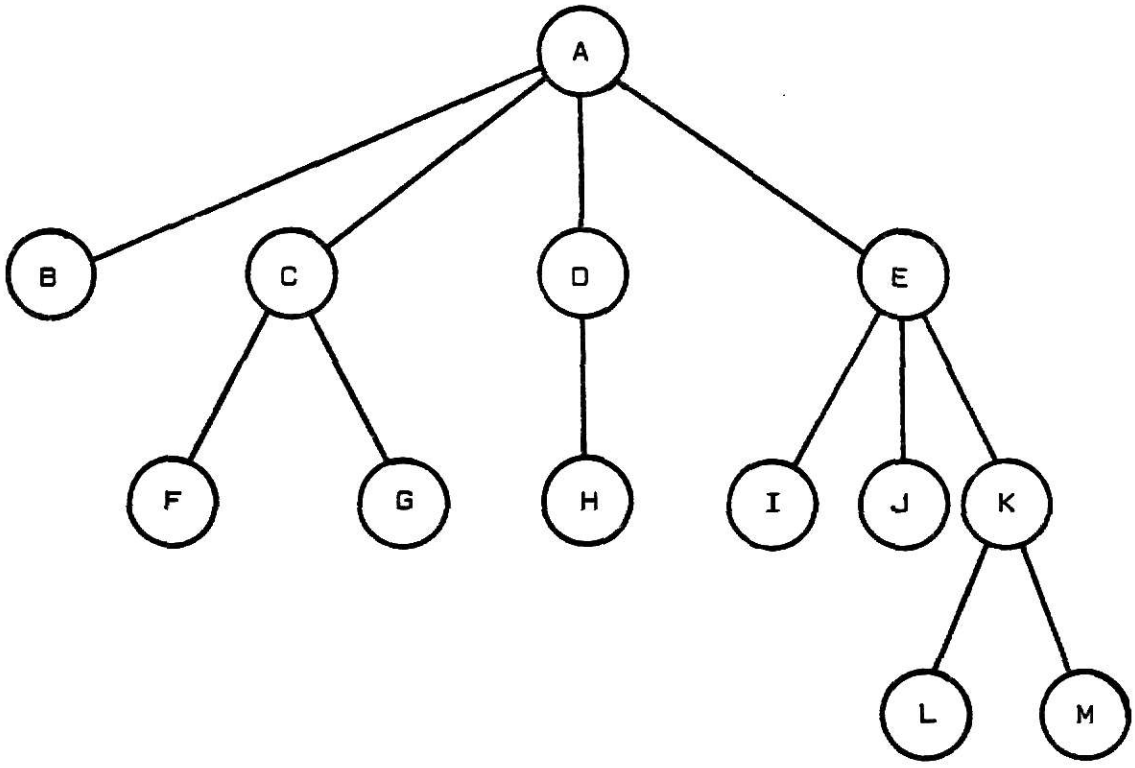

Figure 2.5
Hierarchical Data Base Structure

Each level is the "parent" of the next lower level; each "child" of a higher level then becomes a "parent" of the level below itself. The major advantage over the relational data base is that the data contained in any

given "parent" record is only stored once for the entire parent-child set, which eliminates a considerable amount of required storage space. The programmer who constructs the data base, however, must be familiar with all of the possible paths into and out of the specific data areas: there is no way to get to "M" without going through -- in order -- "A", then "E" and finally "K".

The hierarchical data base is relatively easy to visualize in small pieces, but may be difficult to conceptualize as a whole if the data base is very large. It is easy to see that "L" is the first piece of data found in the "K" record; it may be difficult to follow a twenty level tree to get down to the "K" record. And, since the DBMS retrieves whatever data is at the specified location, a request designed to retrieve "L" which incorrectly specified a route through "J" instead of through "K" would return a "no data present" response, since there is no child below "J". As a consequence, the programmer who designed a utilization program would need to be aware of exactly where each of the record nodes resided within the data base.

F. The Network Data Base

The network or "plex" structured data base is similar to the hierarchical data base model in that there is some form of "parent-child" relationship between records

which is expressed at the physical storage level as a pointer or link connecting the records. The major factor which distinguishes the network-type data base from the hierarchical data base is that the network allows a "child" to have any number of parents; there can be any number of links between two record types. Thus, the hierarchical data base from Figure 2.5 could be connected -- as a network model -- as shown in Figure 2.6 below:

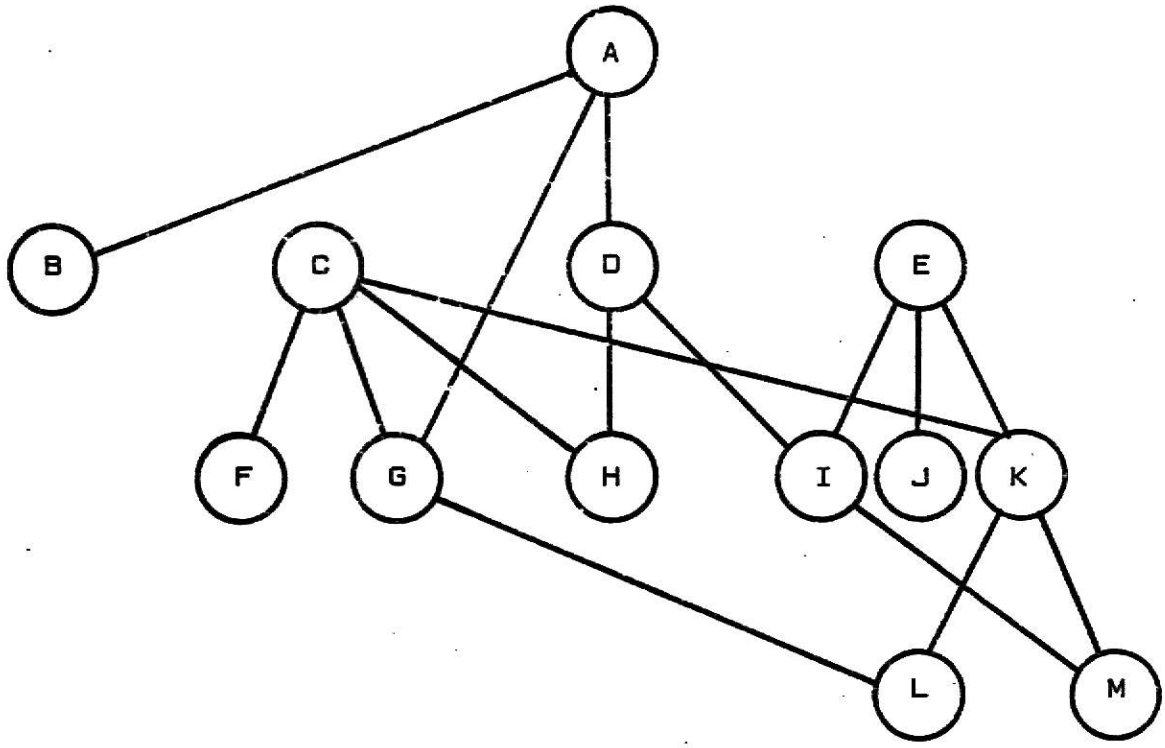

Figure 2.6
Network Data Base Structure

Note that "E" is not longer connected with "A" as in the more strict hierarchical model, but that "I" has two

"parents" -- both "D" and "E"; that "G" is a subordinate member of the "A" set as well as the "C" set; and that none of the records are restricted in the superior (parent) or subordinate (child) records with which they can then be connected.

Networks have nearly arbitrary connections, and, as a direct result, there is not infrequently a serious problem in "navigating" through the structure of the data base. If the program specified that all of the "child" records for "C" in Figure 2.6 should be retrieved, there would be a total of four records returned: "F", "G", "H" and "K". It is likely that the programmer would not be considering records "H" and "K" as members of the set subordinate to record "C", since they appear to be subordinate to sets "D" and "E"; the program, however, would follow the access links and retrieve all records linked to "C". There is almost always the requirement for the combination of an "intelligent program" plus a knowledgeable user [Wie, 77]. There is also an increased probability that a minor error on the part of the programmer may misroute the computer path to the point that the entire data base may "crash", requiring a complete reload of the data base -- an expensive and time-consuming indicator of program errors. The added requirement for multi-level pointers also slows down the retrieval time if an incorrect access path is chosen.

A sub-set of the network data base is the DBTG or CODASYL-type network data base. The Data Base Task Group [DBTG] of CODASYL has established guidelines under which a "standardized" network data base should operate. To be considered a CODASYL-type network, the data base must be modified in several different areas: 1) it must include the concept of a Data Base Administrator [DBA], who designs the entire system and sets the limits of the data base visible to specific applications [the sub-schemas] via a Data Definition Language unavailable to the individual application; 2) there must be a total separation of the overall view of the data base [or schema] from the individual user view or sub-schema, with users restricted to a separate Data Manipulation Language [DML]; 3) there is a restriction on the inter-record relationships allowed, somewhat similar to the hierarchical model, in that record links may flow in only one direction [e.g. from "parent" to "child"]. Unlike the hierarchical data base restriction, however, the CODASYL-standard restrictions do not limit a record to a single parent as long as the parent record is from a set other than the one in which the child is a member. The CODASYL concept also includes provisions for DBA-controlled privacy locks down to the field level [the actual data field contained in any given record], and the capability for utilizing virtual as well as actual results for procedural mathematics.

CODASYL-type networks thus limit some of the forms a data base may assume, and, in the case of a complex (N:M) relationship between records, may require the addition of new records to circumvent the linkage restrictions. The advantage of the CODASYL network is that there is at least an attempted industry standard around which any vendor can structure a new data base. The programming language is the same for all machines; programmers who know COBOL and the CODASYL restrictions can change jobs without retraining. Additional long-term advantages of CODASYL-type data bases are that hardware changes do not require a change in application programs; parts are interchangeable (magnetic tapes and plug-to-plug compatibility of storage devices); data is interchangeable over a network; and procedural differences between different data bases are minimized (Fry, 76). Also, because of both the COBOL-type commands and the requirement for a world-wide standardization, the US government has adopted CODASYL as the governmental standard for data bases.

G. Higher-level File Organizations

The demand for more efficient and more rapid storage and retrieval mechanisms has, in many instances, outstripped the capabilities of present technology to provide the service demanded by the user. As a result, several second-file or higher-level file organizations have

been developed as extensions to both the hierarchical and the network data base concepts. The major recognized file organization techniques in use today are:

1) List and multilist organizations

2) Partially or fully inverted organizations

3) Ring and chained-tree organizations

(Car, 79).

The simple list is a structure which links data elements or records of like type through the use of pointers included within the record itself. Once the program has located the first record of any given list, all of the remaining members of the list can be retrieved by following the "next-record-in-the-list" pointer contained within the record presently accessed. Normally a special end-of-list "flag" indicates when no more records follow the record last retrieved. The multilist organization involves several lists, and enables the user to access any record in one of the lists and then "thread" through the other lists as necessary for data retrieval; multilist organizations are sometimes called "threaded" organizations as a result. (See Figure 2.7)

The inverted-list technique involves the building of exterior index lists to those sets of records frequently accessed. If a manufacturing company frequently required a listing of sub-assemblies manufactured by the various suppliers, an inverted list for each supplier could be

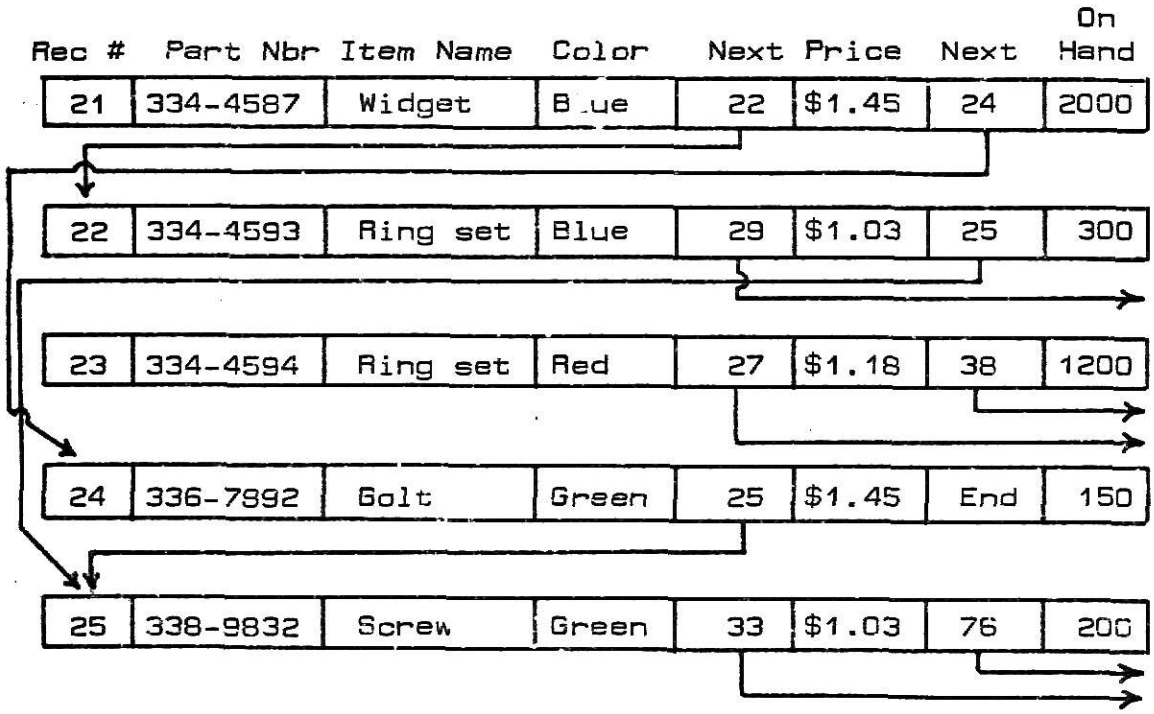

Figure 2.7
Multilist Organization

constructed which contained pointers to all of the parts provided by each individual supplier. The retrieval time for an inverted list is limited to the exact number of records presently on the list, since the search stops at the end of the list. If a portion of the data base is inverted, the structure is called a partially-inverted data base; if all of the records within the data base are on some inverted list, the data base is said to be fully inverted. Inverted lists require additional space for the index structure and pointers to the various records ; therefore, a fully-inverted data base is likely to be a

wasteful extravagance. [See Figure 2.8]

Blue Items	Green Items	$1.45 Items	$1.03 Items
21	24	21	22
22	25	24	25
29	33	end	76
etc.	etc.		etc.

Figure 2.8
Inverted List Organization

The ring structure is primarily a list in which the pointer in the final record points back to the record at the head of the list. The ring has a minor disadvantage in that there must be an extra mechanism to identify when one complete circuit has been made of all the records; without this "all-records-read" indicator, the retrieval process would stay in an infinite loop. The major advantage of the ring is that the loop can be entered at any point. In the list structure, the retrieval process is required to find the first record in the list before proceeding to read any other records. In the ring structure, once any record in the ring is located, all of the remaining records in the ring are available through the pointers. [See Figure 2.9]

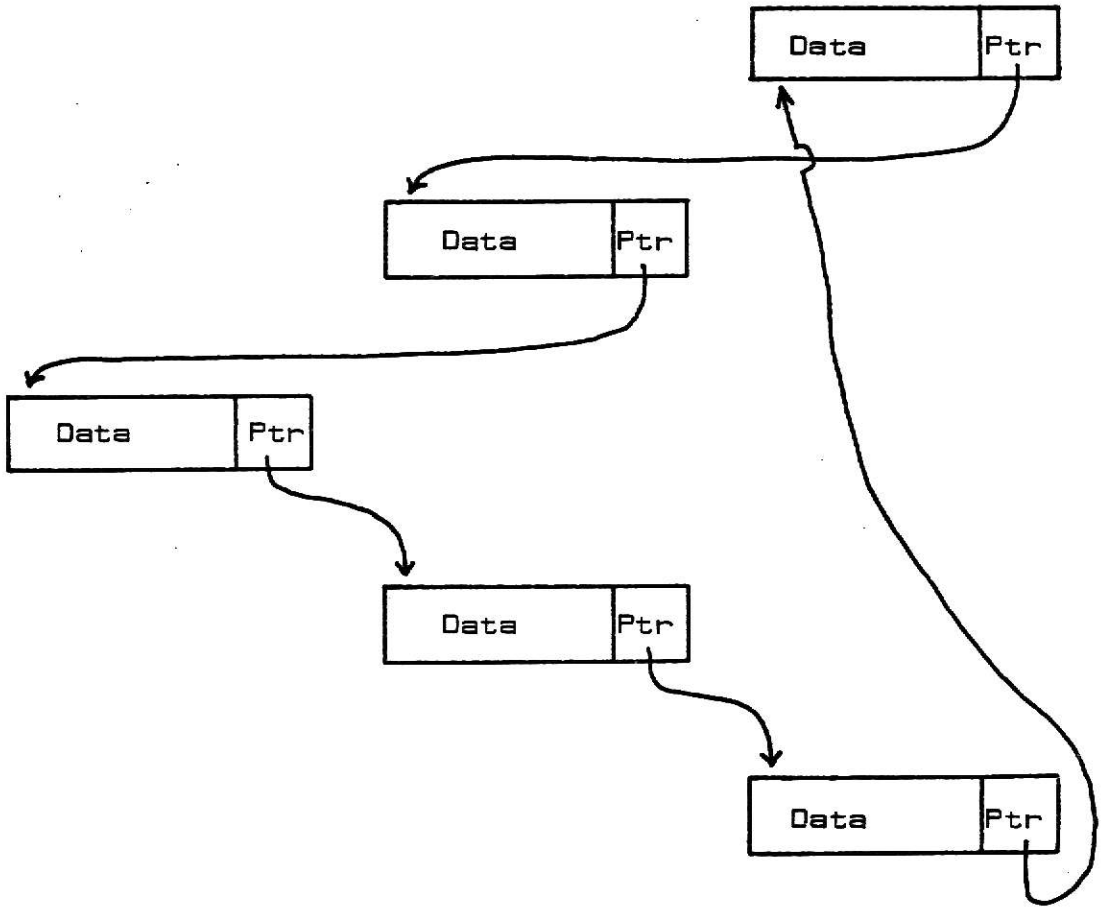

Figure 2.9
Ring File Organization Structure

The chained-tree structure is a complex structure in which pointers are arranged in a hierarchy such that each pointer set includes a pointer to both the subordinate generation and to the next member of the same generation. The lowest-level of pointers in the chained-tree points to the actual machine location of the records contained in the chained-tree; this eliminates the the need for any key

or index mechanism to access the records.

An additional extension is the use of doubly-chained or bi-directional pointers ir the lists, rings or trees. One of the shortcomings of the single-direction pointer is that a hardware malfunction in the read-the-pointer sequence may cause the loss of all subsequent records in the list. In the doubly-linked list or ring, there are two pointers stored in each record field area: one to the record which follows and one to the record which preceeds it in the list. If a hardware malfunction drops one of the pointers, a recovery mechanism can use the remaining link to rebuild the entire structure. (See Figure 2.10)

From a performance evaluation standpoint, the ring increases the capability of the "read-next-record" sequence but may degrade noticeably if all of the records in several interlocking rings are needed. Since each pair of rings in an interlocking ring structure intersects at a record, the record at which the intersection occurs must be read <u>each</u> time a transfer is made from one ring to another [Wie, 77]. The multilist organization is not as cost-effective as either the inverted-list or doubly-chained organization; it takes as much storage space as an inverted file and more storage space than a doubly-chained structure, yet its access time is slower than either [Car, 79]. The doubly-chained tree is more cost-effective than the inverted list, however, and may, under special circumstances actually

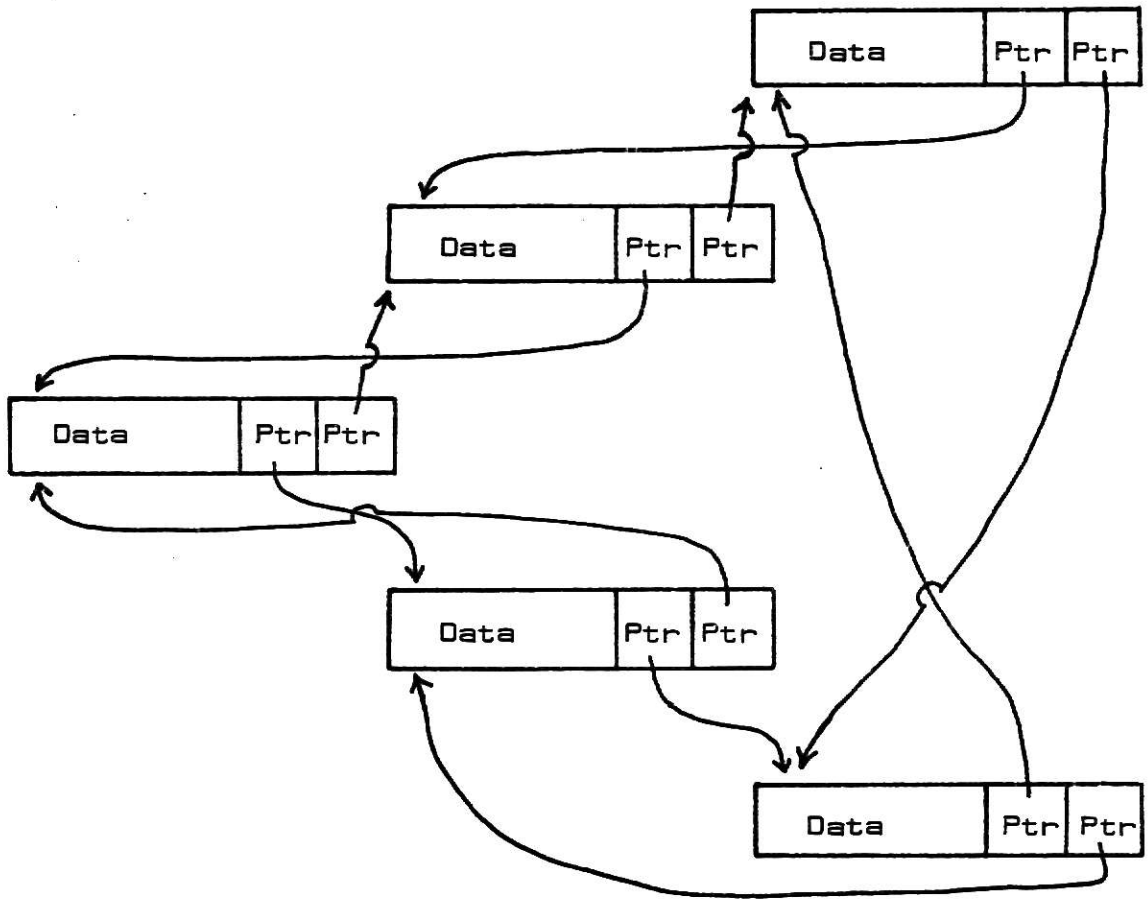

Figure 2.10
Doubly-linked Ring Organization Structure

require less actual space than the original sequence of records [Car, 79]. It also has a significant advantage during an insertion or deletion sequence, as the inverted list requires that updates be reflected in both the index and in the record if any key is altered.

The inverted list is more efficient for insertions or deletions than the doubly-chained record if there is a change of "non-key" data elements, since the only action

required to add a new record is to find the last member of the inverted list and insert the pointer to the new record following that last record. Recent sophisticated packages using the inverted list have allowed an additional capability for making very high-level user queries. In a sequence very similar to the one in the relational data base system, a relatively inexperienced user could request the "Widgets available in Blue, at less than a certain Price, and available in an On Hand quantity of over 1,000". The retrieval sequence would then fetch the three named inverted lists -- Widgets in Blue, Less than X Dollars, and On Hand quantity over 1,000 -- and produce a listing of all parts which were common to all three lists. The major difference between the relational system and the network or hierarchy with the inverted list option, is that the inverted list system only works for those qualities which have previously been considered of enough significance to construct an inverted list for that type of query.

2.6 The Discriminator

A. Introduction

If the model is being used as an independent tool to aid in the selection of the most suitable DBMS for a particular application, the majority of the DBMS's which were found to be available should have been eliminated

prior to moving into the Discriminator phase. If the model is being employed as a decision aid to select the best DBMS from a limited list of generally suitable DBMS's which has been prepared by a subordinate analyst, the number of DBMS's to be considered should be relatively small. In either case, the function of the Discriminator is to either select the best DBMS for the application, or to limit the number of DBMS's for final consideration under a cost-value or other evaluation scheme as required by the application.

The Discriminator phase is sub-divided into two areas: a User Requirements Section and a System Features Section. The User Requirements Section employs an Identification of User Requirements Form to aid the user in the establishment of minimum tailoring requirements for the application. A Requirements Consolidation Form is then used to demonstrate the composite effects of user requirements on the desired architecture. The final portion of the User Requirements Section is the Vendor Capability Form, which is a formatted evaluation sheet to aid in the elimination of those DBMS's which do not meet the previously established minimum user requirements.

The second stage of the Discriminator, the System Features Section, provides a methodology to make a final elimination of all DBMS's which do not have the specific features desired by the user. The Features Evaluation Form

is a suggested value scheme by which the available system features may be evaluated and assigned weights to reflect the value of certain features to the specific application. The final component of the Discriminator phase is the System Comparison Form, which is a formatted device to aid in the comparison of vendor compliance with the features selected as desirable by the user.

B. Identification of User Requirements [See Figure 2.11]

1. What is the maximum number of users who may need to access the data base at any one time ?

One	2-4	5-20	20-100	over 100

A single-user system is significantly different than a two-user system. If the work is purely research, and you are working with information that no one else would ever need to access in parallel, a single-user system may be sufficient; almost any other consideration demands more than the single-user capability. If the present budget allocated for the system limits the entry points at the present time to one, make sure that there will not be a later need for a second or third entry site; once a single-user system is accepted, it may be impossible to

IDENTIFICATION OF USER REQUIREMENTS FORM

1.	WHAT IS THE MAXIMUM NUMBER OF USERS WHO MAY NEED TO ACCESS THE DATA BASE AT ONE TIME ?	[1] ONE	[3] 2-4	[5] 5-20	[7] 20-100	[10] over 100
2.	WHAT IS THE MAXIMUM NUMBER OF RECORDS TO BE MAINTAINED ON FILE IN THE DATA BASE ?	[1] < 1 k	[3] 1-10 k	[5] 10-100k	[7] 100 k - ½ mil	[10] over ½ mil
3.	WHAT IS THE MAXIMUM RETRIEVAL TIME ACCEPTABLE TO MAINTAIN FULL OPERATIONAL EFFICIENCY ?	[1] < 1 sec	[2] 2-4 sec	[5] 4-15 sec	[7] 15 sec - 1½ min	[10] over 1½ min
4.	WHAT IS THE LONGEST PERIOD THE DATA CAN BE INACCESSIBLE AND MAINTAIN FULL EFFICIENCY ?	[1] over 24 hrs	[3] 8-24 hr	[5] 4-8 hr	[7] 2-4 hr	[10] < 2 hr
5.	HOW FREQUENTLY WILL YOU NEED TO INSERT OR DELETE ITEMS IN THE DATA BASE ?	[1] Monthly	[1] Weekly	[5] 10/day	[7] 500/day	[10] over 500/day
6.	HOW OFTEN WILL YOU NEED TO CHANGE THE ORGANIZATION OF THE DATA BASE ?	[1] Never	[3] < Once Year	[5] Once a Year	[7] Semi-Annual	[10] Once a Quarter
7.	DO YOU NEED TO MAINTAIN DATA SECURITY-SENSITIVE TO PERSONNEL OUTSIDE THE FIRM ?	[1] Never	[2] Almost Never	[3] On Occasion	[4] Freq	[5] All of Time
8.	DO YOU NEED TO MAINTAIN DATA TO BE KEPT SECURE FROM FIRM PERSONNEL WITH ACCESS TO DATA BASE ?	[1] Never	[5] Almost Never	[10] On Occasion	[15] Freq'ly	[20] All of Time
9.	HOW IMPORTANT IS STANDARDIZATION IN YOUR SPECIFIC APPLICATION ?	[1] None	[3] Little	[5] Minor Impact	[7] Moderate Impact	[10] Major Impact
10.	WHAT IS THE SKILL LEVEL OF YOUR EXPECTED END-USER ?	[10] 8th grade	[7] HS diploma	[5] College Degree	[1] Cptr Degree	[-1] Degree in Cptr Sc
11.	WHAT IS THE SKILL LEVEL OF YOUR SYSTEMS PERSONNEL ?	[10] HS Diploma	[7] College Degree	[5] Cptr Degree	[1] Degree Cpt Sc	[-1] Adv Deg. Cptr Sc.
12.	HOW MUCH COBOL EXPERIENCE DO YOUR PROGRAMMERS AND SYSTEMS PERSONNEL HAVE NOW ?	[7] None	[5] < 1 mo	[3] 1-3 mo	[2] 3 mo - 1 year	[1] over 1 year

Figure 2.11

add even one more parallel access capability at a later date.

Two to four entry sites are considered a very small system; significantly different from the single-user system, but demanding much less data control than a ten or twelve entry site application. Five to twenty is a small business application. For both of these categories, however, you must be sure that you will not need to expand to the next category in the future.

Twenty to a hundred users seems to include all but the very large systems, which would require a significant overhead to maintain security, subschemas, and some type of efficiency algorithms.

2. What is the maximum number of records to be maintained on file in the data base ?

less than 1000	1 k - 10 k	10 k - 100 k	100 k - ½ mil	over ½ mil

This question may depend on the relative size of the average record to be stored in the data base. If the average record length is over 240 characters, you should divide the average record size by 240 and then multiply the estimated number of records by that factor.

Less than one thousand records is considered a very minimal size data base. A thousand records, in typical 80-

character card format, would require approximately 640 k bytes of data storage area in the unformattec mode; the entire set of records could thus be kept on two standard cassettes. Some of the newer cassette vendors are now advertising 120- 150 k character storage, in 88-character blocks, which would mean that anything less than 1400 records would fit on one of the newer type cassettes. The same size data base would also fit on 2½ floppy disks (at an average capacity of 300 k bytes per floppy disk), or one side of a double-sided diskette (at approximately 1.25 megabytes, formatted, per side), or only nine mini-floppy disks (at an average of about 70 k bytes per mini-floppy disk). Retrieval time would be degraded on either the cassette (because of required tape travel time) or the mini-floppy disks (because of the necessity to change the individual disks), but, on that small a system, it is unlikely that retrieval time is the critical factor in the selection criteria.

A thousand records, in an 80-character format, would require about 640 k bytes of data space; ten thousand records, in the same format, would require 6½ megabytes of space. This means that the 1,000 to 10,000-record data base would fit on a mini-computer type disk unit (at an average capacity of 10 megabytes), or could, if necessary, be stored on four or five diskette units (such as the Shugart 850-line); this appears to be a reasonable break

point for a moderate-sized data base.

A hundred thousand records, in an 80-character format, would require approximately 64 megabytes of data storage space: the data base could fit on one 100 megabyte disk unit, such as the IBM 2314, or on six 10 megabyte mini-computer class disk units such as the Dec 3860. As a consequence of the extra controller hardware to allow ganged mainframe-type disk units or a large number of mini-computer disk units, any data base requiring more than 100 megabytes of data storage area is considered to be a large data base.

A data base storage capacity for over a half million records would require in excess of three of the 2314-type disk drive units; for the type of user likely to utilize this model, a greater-than-half-million-record requirement places that user in the very large system category.

3. What is the maximum retrieval time acceptable to maintain full operational efficiency ?

less than 1 sec	2-4 sec	4-15 sec	15 sec- 1½ min	over 1½ min

At the time this model was originally designed, the industry-standard retrieval time response boundaries were established as less than three seconds being a "query"-type

retrieval category. In November 1978 the standards were changed, primarily as a result of the new hardware capabilities and the consequent user demands for faster and faster response times. The categories shown in the present model are the current industry standards for "excellent", "very good", "average", and, at anything over 15 seconds, "poor" response times. Because this model is oriented toward the smaller data base user, the fourth category is set at 15 seconds to 1½ minutes, which reflects a time that could be reasonably expected from a small cassette-type system. The final category, "over 1½ minutes", is designed to define the lack of query-response time as a primary factor in the DBMS selection process.

For a "normal" user, only a limited few would ever need a response time of less than one second. Aerospace applications or aircraft traffic control for a large commercial airport might require a less-than-one-second response time, but even then it must be realized that the user is demanding that the computer fetch and display the answer in significantly less time that it would take the average end-user to make an entry for even the most simple query. Rapid response time systems cost a lot of money, with the system cost rising in a rapid exponential curve based on the query response requirements. Only the most irresponsible manager would demand a very rapid response system without a serious consideration of whether the need

is a mandatory requirement of the application or just the desire to have the same capability of the system owned by a competitor.

4. What is the longest period the data can be inaccessible and still maintain full system efficiency ?

over 24 hrs	8- 24 hours	4- 8 hrs	2- 4 hrs	less than 2 hours

This question deals with the rebuild and reload time factors. If the system crashes, what is the longest period that the enterprise can reasonably operate without the data base before losing a measurable amount of the system efficiency? Discussions with committee members, and with Jacob Slonim, the previous project manager of the DOMESTIC Project, have established that the average inverted-list and query-language data base, such as System 2000, would need slightly over 24 hours for rebuild; that the average CODASYL-type network data base, such as IDMS, could be expected to allow a complete rebuild cycle in approximately 16- 24 hours; and that the typical DBTG-based, non-CODASYL network type data base might complete a total rebuild cycle in four to eight hours. It must be noted that rebuild time is almost totally size-dependent. A very small data base may be expected to be ready to run in an hour or less; a very large data base may take days to rebuild and reload.

The final category, "less than two hours", was chosen as an indicator that the entire capability of the user to conduct the business depends on access to the data base. If degradation of the entire enterprise occurs in less than two hours without the data base, the primary concern of the user is the providing of information from the data base.

Conversations with the Data Base Administrators at several DBMS sites have demonstrated two important characteristics of DBMS rebuild times: 1) rebuild and reload times reported to the author by actual users of the DBMS's in a commercial environment are frequently multiples of the time advertised by the vendor, and 2) selection of too short a reload period as a mandatory criterion in the selection process can be an excessively restrictive factor in the search for a satisfactory DBMS.

5. How frequently will you need to add or delete items on the data base ?

Monthly	Weekly	10 Per Day	500 Per Day	Over 500 Per Day

This is a difficult area to quantify at the extreme ends of the scale, yet a very high or very low rate of insertions or deletions can be a significant factor in the selection of a DBMS.

A once-monthly insertion or deletion rate indicates

that a programmer could probably write a special program for any necessary changes to the data base. A once-weekly rate could probably be satisfied by a program stored on magnetic tape or in card deck form and brought out of storage for any needed insertions or deletions.

Ten changes a day exceeds an average rate of one change an hour for an 8-hour shift, so the user would require at least a moderately-efficient, easy-to-access insertion and deletion program. The final break point was chosen as approximately 500 a day, since this rate exceeds a change a minute. A rate of more than one change a minute requires more than just an efficient program for insertions and deletions; it requires a data base architecture which is optimized for group insertions or deletions. It was decided that something less than this rate might be satisfied by local optimization procedures, but that anything over that rate would require special consideration in the selection of an appropriate underlying architecture for the data base.

6. How often will you need to change the data base organization ?

Never	Less than Yearly	Yearly	Semi-Annual	Once a Quarter

This question is not concerned with how often data residing within the data base is changed, but how often the actual schema or major subschemas of the data base are altered. Changing the *entire* inventory line numbering system might be done at the data field level without any change to the data base organization; changing to a new product line which had subassemblies or subaccounts unique to the present organizational structure *would probably* require a data base organization change.

The average user with whom the author has had contact reports an observed rate of yearly or less changes to the data base organization structure. With most present-day data base organizations, however, the performance of the data base system could be improved measurably by a careful study of the present or intended heavy-traffic usage areas of the data base and a corresponding reorganization of the data base. Data bases are, almost by definition, dynamic; if response time is a consideration for the application, reorganization should be conducted on a more frequent basis than is done by most of the users who presently own DBMS's.

7. Do you need to maintain data that is security sensitive to personnel outside the firm ?

8. Do you maintain data that must be kept secure from personnel in your own firm who have access to the general data base ?

(for both)	Never	Almost Never	On Occasion	Frequently	All of the Time

These two questions are designed to separate the security requirements that can be met easily with some device or software facility -- independent of the DBMS, and those security requirements that will need either special DBMS features or else a complex and tailored software to meet the user's requirements.

Question 7 refers to the necessity of keeping persons outside of the company from accessing any portion of the data base. At the lowest end, no security of any kind is required; at the highest end, no one should be able to sign on the system and gain access to the data base without first passing some rigorous access parameter, such as a multiple keyword sequence or an ID password and account number match.

Question 8 refers to the requirement for an in-house security check by the user. There is a requirement to prevent some personnel who have access to the data base from accessing certain sensitive data stored in another part of the data base. In a bank application, the tellers might be authorized to check if there were sufficient funds in a person's account to honor a withdrawal of a specified amount. If the bank records were constructed on a record-per-customer basis, it is likely that any loans owed to the

bank would be in one of the data fields of the record of the person making the withdrawal. Privacy restrictions would dictate that the teller not be allowed to check the loan status on everyone who desired to cash a check or make a savings withdrawal, so a field-level lock -- a lock on that portion of the record contents that dealt with all of the loan transactions -- would be inserted into each bank customer's record. Even if the teller were able to find out the label name of the field, or the navigational path to the data area, a software feature of the DBMS would refuse to retrieve data about loans for any user with a teller identification number.

Although many users may feel that they need field-type security, they may not have the justification for spending the extra charge for this feature. The opposite view, however, is that with the rapidly-growing concern about the inadvertant divulging of personal information, it may soon be an economic fact of life that record-level or even the field-level security is the best insurance against civil suit.

9. How important is standardization to your application ?

None	Little	A Consideration	Moderate Concern	Major Concern

As discussed before [Section 2.5, f], standardization may be a requirement if your firm intends to conduct any business with the US government involving data base compatibility. Even if there is no intention of working with the government in a "shared software" environment , standardization will eventually be an economic necessity. Vendor equipment becomes a secondary consideration if all of the programs are written in a standardized format. All of the programmers are "interchangeable", as are storage devices. The only argument against standardization is that the present standardization agreement [CODASYL] may not include the best set of restrictions, and the cost for a change-over at the present time may be wasted if there is a change in the standards in the near future. Arguments tendered by the major hardware vendors should be dismissed in most cases, since the vendors are more interested in limiting purchasers of the mainframe to only that specific vendor's line of peripherals.

There are many areas of CODASYL which are not considered state-of-the-art in software techniques, but the relative independence from specific vendor hardware may be an economic factor which more than subsidizes any minor changes in the present CODASYL standardization agreement.

10. What is the skill level of your expected end-user ?

8th grade	HS diploma	college degree	computer degree	degree in computer sc.

This consideration should include a very careful projection of future needs; the eventual user group will probably extend downward in skill from the original intended target group. If the data base is only for the high-level management, the secretaries may be asked to retrieve data for their bosses; if the data base is set up for inventory control, a lower-level stock clerk may be required to retrieve part numbers from the data base. If a moderately-inexpensive training program provided by the vendor can allow adequate access to the data base for all of your personnel, at least consider the future consequence of the necessity for a vendor-provided training program for users. Training time for a new hiree includes both lost time for the new hiree and for the experienced person conducting the training. Two years after the installation of the DBMS, the vendor will probably not be willing to send a company representative to your site for training -- even though the training program is a new-sales feature.

11. What is the skill level of your systems personnel ?

HS diploma	college degree	computer degree	degree cptr sc.	adv. degree in cptr sc.

Using the same evaluation as for Question 10 above, select the skill level likely to be the lowest for any possible future systems personnel. In the future, you may not be able to afford a data base that demands an advanced degree in computer science as a prerequisite for systems *maintenance*; personnel costs are rising much more rapidly than are computer hardware costs.

12. How much COBOL experience do the programmers and analysts have at the present time ?

None	Less than one month	1- 3 months	3 mo - 1 year	over 1 year

If all of your programming is done in COBOL, even a *new trainee* could be lumped into the highest category, since COBOL is not an additional requirement for the DBMS beyond that of the normal application work load. If only a small portion of your staff, such as the accounting and budget technicians, are familiar with COBOL, use the level of experience for those personnel *not in* that area.

C. Explanation of Assigned Point Values

The scale values on the User Requirements Form range, on the average, from 1 to 10, with a two-point differential between the adjacent categories. This point spread differential was chosen to give adequate weight to the difference between a very high and very low choice of category descriptors. If the scales were valued a straight 1- 5 and the user selected all but one response in the middle of the scale, the total score would be only two points different from the "average" even if the category which was not in the middle were at the very high or very low end. With the present 1- 10 point scale, if the user selects only one response at the high or low end, it makes a five-point difference in the total; this should attract some attention to the deviation from the norm.

Several of the assigned point values have been altered from the average weighting scheme to reflect particular differences in the values considered appropriate to that specific question. The first two responses for Question 3, which is concerned with retrieval time criteria, are considered significantly different from the remaining three responses. The difference between a "less than 1 second" and a two-second retrieval time is significant, but, for the expected user of this model, not as significant as the difference between a four-second and a fifteen-second retrieval requirement. The "2-4 second" category value has

been adjusted to "2" to show this closer relationship to the low end of the scale than to the middle of the scale.

Question 5 has been adjusted to demonstrate the minimal difference between a monthly and a weekly rate of insertions. Neither of these two categories would require the elimination of a data base structure which is generally insertion-inefficient.

Questions 7 and 8 are weighted differently than the average since they are an interactive set. The high end of Question 7 is a "5"; the middle point of Question 8 is a "10". Any user who needs exterior security all of the time and field-level security at least part of the time is roughly equivalent to the user who needs some field-level security on occasion. The adjusted weighting reflects this equivalency.

Questions 10 and 11 were adjusted at the high skill level end of the scale to emphasize that a requirement for a very high level of computer skill may be a major deficit in future budgetary considerations. Question 12 was altered from the average to slightly minimize the effect of just one language on the entire selection process. Since a programmer with more than three months COBOL experience is considered to be familiar enough with the general concepts of COBOL that there would not be a problem in learning a CODASYL-type data base language structure, the differences between the two highest categories were minimized.

D. User Requirements Consolidation Form (See Figure 2.12)

After you have reviewed the point values assigned for each response, adjust any point values which do not appear to reflect critical areas of consideration for your specific application. Transfer the point values above the selected category for each question to the two-column box on the left side of the Consolidation Form.

The Consolidation Form has five special areas of consideration on the right side of the form. Transfer the assigned point values to the box provided below the appropriate question and add up the totals for each of the five areas. When you have completed the Consolidation Form, refer to the Explanation Section (Section 2.6-E) for a discussion of the effects of user requirements on general DBMS architectural considerations.

E. User Requirements Explanation Section

1. Type of Organization

As a general rule of thumb, if the total score for this area is above 40 you should be looking at either a relational-type data base or else a network or hierarchical data base which has <u>both</u> an extensive English-like query language and some form of higher-level file

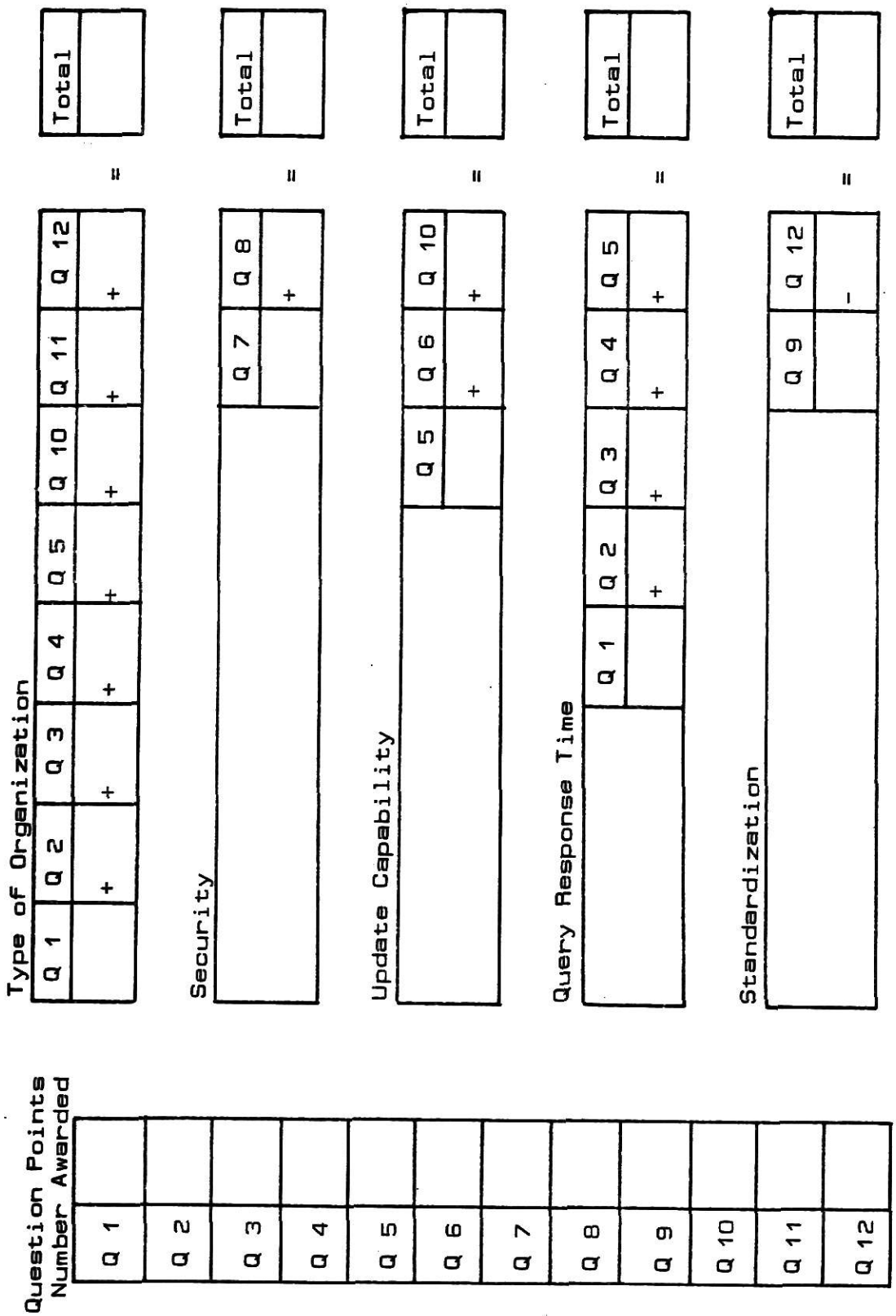

Figure 2.12 Requirements Consolidation Form

organization method. If Questions 2, 3 and 12 received high values and this is the primary reason the total is above 40, the best choice would probably be a CODASYL-type network with the high-level query language and high-level file organization method; the advantages of a relational data base decrease at the high point value for all three of these areas. If the total score is below 25 it is likely that almost any type of data base organization will satisfy your needs, although the relational data base would probably be a cost-ineffective choice. If there are any values at either the very high or very low end of the scale you should read the explanation which follows for that particular area and modify the selection to accommodate any significant differences noted.

Question 1: Number of Users

CODASYL-type networks and hierarchical data bases demand specific user actions to open and close a file Volume Table of Contents, or VTOC, in order to access a data area. The higher the number of users, the greater the likelihood of temporary or permanent deadlock, and the greater the probability of a programming error on the part of one user causing severe problems to all other users. The inverted-list data bases may cause problems if the number of users is very high, since they demand a series of

index-type lists. A large number of concurrent users may increase the chances of deadlock or data error problems as more than one user attempts to access the same index-list.

As the number of users increases, there is less a chance of deadlock or excessive waiting period for access for the relational-type data base than for any other type of data base, since the relational data base operates at the individual record level. Unless more than one user attempts to access the exact same record, there is no problem inherent in the relational concept that increases as the number of users rises.

The prospective buyer of a relational data base must understand that this lack of restriction applies only to a "true" relational data base. Because of the high level of interest in the relational concept, many of the major computer vendors are now in process of putting a new "relational" data base on the commercial market, but the data base is "relational" only at the query level. The new relational data bases are actually network or hierarchical data bases with a very sophisticated front-end query language package which allows the user to interact with the data base in a free-form manner similar to the relational data base. The underlying architecture still demands the analyst-level familiarity with the appropriate opening and closing of files and still possesses the previously stated problem areas in the access of lists and files as the

number of users increases. If a DBMS is advertised as a relational DBMS, make sure that the data base is relational -- not just the front-end query package interaction set.

Question 2: Number of Records

If there is a very small number of records to be maintained on the data base, any of the organizational types may be appropriate. As the number of records rises those hierarchical and network-type organizations without some higher-level file organization begin to deteriorate because of the extensive navigation requirement necessary to access any given record. If a large number of records is an implicit indicator of a complex organizational structure, an inverted list structure is inappropriate; the performance of an inverted list may be severely degraded, since inverted lists do not handle complex structures well [Mar, 77].

Inverted lists are well suited for large numbers of records in an "informational" system, but generally not suited for a "working operational" system [Mar, 77]. The inverted list structure is exceptionally well suited for a large data base where queries involve the confirmation of record presence or absence or a check of the number of a specific type of record presently on file in the data base. If the general type of query is of the type "Are there any

records which contain Blue Widgets?" or "How many Blue Widgets are there?", the answer can be determined by a simple counting of the index list -- without ever accessing a single record. A potential problem, however, is the inclusion of too large a percentage of the total data base on the indexed lists; if too many records are maintained on the index lists, more time is spent in "check the list -- get the pointer -- find the record" than would be spent in a less sophisticated sequential scan. A general rule is that the average number of records qualified by the average query must be less than 10% of the total records in the data base to justify the inverted list approach [Car, 79].

Relational data base performance also breaks down as the number of records increases. Since a relational data base searches individual records for actual data content, more records to search means more time required to return a response. In the near future it is expected that faster memory retrieval techniques and developments in the area of associative memory will lower retrieval rates to the point that relative differences between the relational and other type data bases are not apparent to the user.

Question 3: Retrieval Times

As a very simplified rule, retrieval times for the three major organization types vary -- from slowest to

fastest -- in the order: relational, network and then hierarchical systems [Wie, 77]. At the present level of technology, relational data bases are quite inefficient if the major consideration is retrieval time. Chain and ring structures are also on the slow end of the spectrum of retrieval times [Wie, 77]. With the restrictions mentioned in the discussion for Question 2, the inverted list tends to be one of the faster structures for pure "speed of retrieval". The primary shortcoming of the inverted list, for some applications, is the inability of inverted lists to carry up-to-the-minute data. Most inverted list DBMS's depend on off-line updates as a result of the necessity to update both the record and all lists on which the record appears. If a high rate of insertions and deletions is combined with the requirement for rapid retrieval, the update complexity of the inverted list degrades the system to the point that the user will not be satisfied [Mar, 77].

Question 4: Data Base Inaccessibility

Rebuild time for a data base is very size-dependent, but, as a general rule, the relative rebuild times -- from shortest to longest -- should be: relational, inverted list and hierarchy or network. The crucial discriminator is the number of pointers to be updated and the manner in which the pointers are constructed. The more pointers that

are needed by a DBMS, the longer will be the requirement to reinsert all of the pointers and link them with the appropriate records. The double-linked lists and chains do offer additional security, since a minor malfunction of the hardware will not lose a complete list of records. The reloading process for a doubly-linked structure is a very complex and time-consuming function, however, placing any of the doubly-linked chains and rings at the very slowest end of the rebuild range.

Question 5: Rate of Insertions/Deletions

Relational data bases would appear to be designed expressly for the optimization of insertions and deletions. There are no pointers to update, and there is no path to navigate to ensure that the record reaches the necessary physical location for later retrieval.

Inverted lists are very poor for insertions and deletions, and any hierarchical data base tends to be a poor choice if the files are volatile (Mar, 77). If the frequency of update divided by rate of retrieval is high, the physical linkage systems would be a good choice, since the update capabilities of the physical linkage structures outweigh the limitations of retrieval speed (Ross, 78). Or, in simpler terms, a highly volatile file that is accessed on a moderately infrequent basis would be a good

application for which the user should consider any of the rings, chains or more complex doubly-linked structures.

Question 10: End-user Skill Level

If the skill level of the end-user is very low, the obvious choice is the relational data base, and the next best choice would be a network system with a proven and comprehensive query language which allows an English-like interaction with the system. If the system is hierarchical or network without the front-end query package, there is almost no capability to handle the "casual user"; the user must know how the data base is laid out in order to access the data desired. Network data bases appear to be the worst for the less experienced user, since the multiple relationships available between records increases the probability of user-induced navigation error.

Question 11: Staff Skill Level

As with the end-user skill level question, relational data bases favor the less experienced programmer. Since relational data bases are arranged in tabular or flat-file format, a format with which most people are familiar, they are much more easily visualized. The logic-type errors produced by programmers are normally lower for relational

data bases than for either the network or hierarchical type data bases [Loc, 77]. The Lochovsky study also indicated that the hierarchical-type structure appeared to induce the most logic-type errors, since most programmers had severe difficulties with the position pointer, the "get next" command common to the hierarchical data base, and the method of retrieval to guarantee that all data which met a specified criterion was actually selected [Loc, 77].

Question 12: COBOL Experience

The category values for this question were weighted so that even a high level of COBOL experience would have a somewhat limited effect on the selection process. A high level of COBOL experience does tend to weight the choice of a CODASYL-type network data base, however, and a very low level of COBOL experience would detract from the value of the CODASYL-standard data base.

CODASYL-type networks are based on COBOL-like language commands; the original concept of the DBTG was literally to design a standardized set of rules for network data bases that would use the generalized COBOL language structure. If the staff has no experience in COBOL, there may be a significant requirement for language structure retraining in order to use a CODASYL-standard data base.

2. Security

If the total for the security area is over 12 you either presently have the need for, or soon will have the need for, a DBMS which offers security locks down to the field level. The relational data bases have the greatest flexibility for specific field level locks. The next best choice would be one of the CODASYL-standard data bases. If there is an indication that field-level locks may be required, it is best to get a DBMS that has that specific feature; tailored software is expensive and would be a "patch" if installed at a later date.

3. Update Capability

If the total score for this area is over 10, the most likely choice is a relational data base; the least likely choice is any data base which uses sophisticated pointers or linkage structures. If the total score is very low, the most cost-effective data base structure may be one involving rings and chains. There is an indication that you have the on-board skill to deal with the complex structures and linkage paths, and there is little need for a highly-efficient insertion and deletion program or a sophisticated data base reorganization algorithm.

4. Query Response Time

Generally, if the total for this area is

above 15 you should be looking at a relational type data base, and if the total is less than 10 you should be considering some form of inverted list structure. Look at the specifics of the application: if there are a large number of "Are there" or "How many are there"-type queries, the inverted list is tailor-made for the application. If the queries are primarily non-standard in nature, no matter what the size of the data base or the expected query response times, your best choice is a relational data base, and the second-best category would be a DBMS with a high-level query language feature. Note that Questions 4 and 5 are a counter-active pair for this consideration; if one goes up significantly, the other must come down.

5. Standardization

If you need standardization because of present or proposed contracts with the government which involve data base interoperability -- or if you are planning to integrate into a data base network, you don't need to look at the total: get a CODASYL-type data base. If standardization is only a consideration and the total score for this area is 5 or higher, a CODASYL-type data base is also a good choice. [Note that Question 12 is a subtractive factor; the highest possible score is nine.]

F. Vendor Capability Form (See Figure 2.13)

The Vendor Capability Form is designed to provide a reasonable format for aiding the prospective user in rating the various vendor packages. It is primarily a "go/no-go" design. You should have already identified your minimum requirements in previous stages of the model; you may insert either the vendor feature offered or a simple "yes/no" response as to whether the minimal needs are satisfied by that particular DBMS package.

1. Maximum number of users to access system

Vendors can be expected to use the maximum capability figures, under optimal conditions, so make sure that the vendor-quoted number of users exceeds your requirement. If you desire a large system (more than 30 simultaneous users), get a list of present users who are now using the system successfully with at least that number of users.

2. Maximum number of records to be stored

The primary area of concern here is that the vendor-quoted figure applies to the same type of storage media to be used in the intended application. Do not accept storage space figures quoted in "words"; different vendors use different word lengths.

3. Maximum guaranteed retrieval time

Retrieval time is a factor inextricably

VENDOR CAPABILITY FORM		User Need	Vendor/ System Identification Number					
1	Maximum number of Users to access system							
2	Maximum number of records to be stored							
3	Maximum guaranteed retrieval time							
4	Maximum expected rebuild time							
5	Ease of adding/deleting items on data base							
6	Ease of changing data base organization							
7	Exterior security locks							
8	Security locks down to field level							
9	Meets CODASYL standards							
SYSTEMS SATISFYING MINIMUM USER REQUIREMENTS								

Figure 2.13

bound to the size of the data base. A DBMS that has a vendor-quoted response time of less than a second may, if the test data is based on a data base of 1,000 records, return answers for a million-record data base in ten minutes. Determine the size of the system for which the retrieval times are valid; if the size of the test system is much different than your intended application, check a user who has a data base size approximating your intended application. The multiplicative factor for multiples of a given size data base vs. expected retrieval times is likely to be logrithmic or worse, so be very judicious in any attempts at extrapolation from vendor-quoted data.

4. Maximum rebuild time

One of the most frequent complaints made to the author by present users of DBMS's was the difference between the vendor-quoted rebuild time and the actual time required in a commercial environment. The Davis book also supports this prevalence of user complaint concerning the requirement for extensive rebuild times for a commercial application (Dav, 75). This area should be validated by contact with a present user of the system with your type of application before much credence is placed in vendor claims.

5. Ease of inserting or deleting items

The manner in which items are inserted or deleted may alter the capabilities of the system to

efficiently handle the problem. Chaining, for instance, may be very efficient for an item-at-a-time insertion, but breaks down drastically if the user attempts to make a large number of insertions at one time. Make sure that the "ease of insertion" claim applies to your specific expected pattern of effecting this task.

6. Ease of changing data base organization

This is a frequently overlooked area of concern. Most of the users contacted during the early part of the study reported an exceptionally infrequent restructuring of the data base organization, so do not base your requirements on present-user comments. If the DBMS does not have some facility to periodically compact the data base and purge the system of all unnecessary data, the efficiency of the DBMS is likely to degrade noticeably over time. As familiarity with the DBMS increases, it is likely that some restructuring would be beneficial. Data bases are dynamic, and there should be the capability of adapting the method of accessing the data base to best meet the changing structure of the data base itself. Many present users do not restructure because of the complexity and time requirement of the process for their DBMS package -- not because of a lack of need to reorganize.

7. Exterior security locks

If the system uses the schema/ sub-schema concept, there will be some measure of logical data

independence even without some added security features. If your present system has an adequate security lock-out feature as part of the operating system, the built-in security lock as a DBMS feature may be a needless expense.

8. Security lock down to field level

Hand-tailored software is expensive; if there is even a possible future need for field-level locks, get a DBMS which offers the locks as part of an integrated approach package.

9. Meets CODASYL standards

The issues have been discussed; if you need standardization, or if you believe that the benefits of a single-language, single-protocol training are worthwhile, look for a CODASYL-type DBMS.

G. Features Evaluation Form [See Figure 2.14]

This form is a check of those minimum areas which need to be addressed with the vendors. The point values appearing in parentheses are only a guide, based on the previous general range of 1-10, with a two-point difference between levels to adequately demonstrate a very high or very low rating.

1. What is the skill level required by the "end-user" for your specific application ?

FEATURES EVALUATION FORM

1	WHAT IS THE SKILL LEVEL REQUIRED BY THE "END-USER" FOR YOUR SPECIFIC APPLICATION ?	8th grade ---------- (10) HS diploma --------- (7) College degree ----- (5) Computer degree ---- (3) Degree Computer Sc.- (1)	_____
2	WHAT IS THE SKILL LEVEL REQUIRED BY YOUR SYSTEM MAINTENANCE AND SERVICE PERSONNEL ?	HS diploma --------- (10) College degree ----- (7) Computer degree ---- (5) Degree Computer Sc.- (3) Adv deg/ Cptr Sc. -- (-1)	_____
3	WHAT IS THE SKILL LEVEL REQUIRED BY YOUR STAFF TO MODIFY THE SYSTEM FOR FUTURE EXPANSION ?	HS diploma --------- (10) College degree ----- (7) Computer degree ---- (5) Degree Computer Sc.- (3) Adv deg/ Cptr Sc. -- (-1)	_____
4	HOW MUCH TRAINING WILL YOUR OWN "END-USERS" NEED TO INTERACT SUCCESSFULLY WITH THE SYSTEM ?	Less than 3 hours -- (10) 3 to 8 hours ------- (7) 2 to 3 days -------- (5) 4 to 5 days -------- (3) More than one week - (1)	_____
5	HOW MUCH TRAINING WILL YOUR OWN SYSTEM MAINTENANCE STAFF NEED TO SERVICE THE SYSTEM ?	3 to 8 hours ------- (10) 2 to 3 days -------- (7) 4 to 5 days -------- (5) 6 to 10 days ------- (3) More than 2 weeks -- (1)	_____
6	ARE COMPREHENSIVE USER MANUALS AND USER TRAINING PROGRAMS AVAILABLE FOR THE SYSTEM ?	Included with system (10) For moderate fee --- (7) One area limited --- (5) Both areas limited - (3) No "good" manuals -- (-1)	_____
7	WHAT IS THE EXPECTED RESPONSE TIME FOR TELEPHONIC SERVICE REQUESTS ?	24-hour service ---- (10) 1st shift only svc - (7) return call w/i 1 day (5) return call w/i 2 day (3)	_____
8	WHAT TYPE OF VENDOR DIRECT-CONTACT IS AVAILABLE FOR SERVICE PROBLEMS ?	Major office in city (10) Many large branches w/experience base (7) Few branches /or/ rel. new system -- (5) Single office /or/ new system ------- (3) No direct contact -- (-1)	_____
9	HOW MANY USER NAMES WILL THE VENDOR SUPPLY FOR CONFIRMATION OF VENDOR/ SYSTEM RELIABILITY ?	5+ users/ same appl. (10) 3-5 users/ same appl (7) 10 general users --- (5) 3-9 general users -- (3) less than 3 users -- (-1)	_____

Figure 2.14

The key is "for your specific application". If the intended application is primarily research, it may be realistic to accept a requirement for a very high level of user skill in order to circumvent the additional expense of a sophisticated high-level query language. If the intended application is for less-sophisticated end-users, make sure that the lowest skill level person you are likely to hire will be capable of meeting the vendor-specified minimum end-user qualifications.

2. What is the skill level required by your system maintenance and service personnel ?

3. What is the skill level required by your staff to modify the system for future expansion ?

Both questions are user-dependent and should be addressed by the vendor only after the exact application is understood. Note that an advanced degree in computer science is given a rating of a negative 1 -- indicating that this may be a detrimental requirement in the future. It may be of major interest to check on your legal capability to make system modifications, since many of the vendors demand a contractual agreement which precludes any modification of the software.

4. How much training will your own "end-users" need to successfully interact with the system ?

5. How much training will your own system maintenance staff need to service the system ?

This is a factor which must be considered in view of expected personnel turn-over for the specific application. If personnel turn-over is very low, a one-time training program may be a valid expense; if the turn-over is very high, find out how new personnel will be able to attend vendor training sessions.

6. Are comprehensive user manuals and user training programs available for the system ?

Many of the smaller systems on the market include a thirty-page or less user manual. Ask to see the user manual if there is any doubt as to the quality of information. Lack of a reasonably comprehensive manual for the user is considered a deficit to the system rating, as is the absence of a vendor-supported training program.

7. What is the expected response time for telephonic service requests ?

Present-user response has indicated that users seldom

if ever require a service representative on site after the initial installation. The capability of reaching a knowledgeable service representative on short notice, however, is a major selection factor, since the user may have a totally disabled system until the vendor replies with the corrective action technique. If the system is widely used and the vendor has a large base of experience with user problems, a 24-hour answering service of semi-skilled user-oriented technicians may be adequate to solve any problems. If the system is new to the vendor it may take some detailed study by more sophisticated technicians to uncover the user problem and repair it. Find out what the actual response time is from a user who is working in your general application field.

8. What type of vendor direct contact is available for major service problems ?

Although geographic locale has little to do with vendor responsiveness, the highest level of vendor response can still probably be expected from a major vendor which has its main offices in the same city as the prospective user. A high number of branch offices will be satisfactory for most users, since the majority of user problems which would normally be encountered have probably been solved many times over by the vendor. It should be noted that

many users felt that as a system became more accepted, the level of response to the user dropped severely [Dav, 75]. It may be advisable to check with new users of the widespread systems; the vendor may have provided excellent service while in the initial acceptance stage, but may not be providing the same level of service now that the product is well-established.

3. How many user names will the vendor supply for confirmation of vendor or system reliability ?

The primary interest is the type of users whose names are provided. Every attempt should be made to secure a list of users who have a similar application to the one for which the DBMS is being selected. Note that less than three users is assigned a negative value, since it is an indication that either the system is so new that your firm may be helping in the "shake-down run" for the vendor, or else there is a lack of satisfied users.

H. System Comparison Form [See Figure 2.15]

After you have adjusted the values of the System Evaluation Form to reflect a realistic appraisal of the specific application for which the DBMS is intended, the use of the System Comparison Form is primarily a decision

SYSTEMS COMPARISON FORM		Vendor/ System Identification Number					
1	Skill level required by "end-user"						
2	Skill level required for system maintenance						
3	Skill level needed to modify system						
4	Training for "end-user" to use the system						
5	Training for staff to maintain the system						
6	User manuals and training program availability						
7	Direct contact with vendor for service problems						
8	Response time for telephonic service requests						
9	User names supplied for reliability check						
TOTAL POINTS AWARDED EACH SYSTEM							

Figure 2.15

as to the level of each area which applies to the system being rated and the assignment of the appropriate points to that system feature. If, for example, the system being rated has a requirement for a high school diploma skill level for the end-user and you have decided that the high school diploma level is adequately weighted at 7 points, insert a "7" in the Question 1 box for that system. When you have assessed the vendor capability to meet all of your requirements for each of the nine areas, add up the total points and eliminate those DBMS's which have the least number of total points.

If there are no major differences between the nine areas for your intended application, one possible way to emphasize relative weights for your application would be to employ a weighted-ranking scheme. Rank the nine areas in order of their relative importance to your application. After you have rank ordered the areas, assign nine points to the most important area, eight points to the second most important area, etc., down to one point for the least important area. Using the numerical point value assigned as a multiplicative factor of relative importance to your application, multiply the vendor feature points from the Features Evaluation Form by that relative value. [If the prospective end-user skill required was the most critical consideration and the rated system demanded at least a high school education, the points awarded for that system for

Question 1 would be 9 x 7 or 63.)

CHAPTER 3
USER QUESTIONNAIRE

3.1 Introduction

The User Questionnaire (See Figure 3.1) was used as a validation instrument for the DBMS selection model to determine what present users of DBMS's considered as the important criteria in the DBMS selection process. Any data quoted is based on a non-statistically sound sample; no attempt was made to gather enough data to provide a sample which would have provided statistically significant results.

3.2 By-Question Analysis

What is the maximum number of users who could access your data base at any one given time ?

The question was used to investigate the possibility of one of the three DBMS packages being favored over either of the other two for a large or small number of users, and to help define what the reasonable bounds of the number of users on the selection model should be.

Response from the users, discussions with the Data Base Administrators, and consultations with the committee members have indicated that these figures are valid as an

USER QUESTIONNAIRE

WHAT IS THE MAXIMUM NUMBER OF USERS WHO COULD ACCESS YOUR DATA BASE AT ANY ONE GIVEN TIME ?

One	2 - 4	5 - 20	20-100	over 100

WHAT IS THE *AVERAGE NUMBER OF RECORDS* YOU MAINTAIN ON FILE IN THE DATA BASE ?

< 1000	1000-10,000	10,000-100,000	100,000-500,000	>½ million

WHAT IS THE MAXIMUM RETRIEVAL TIME YOU CAN ACCEPT AND STILL REMAIN AT FULL OPERATIONAL EFFICIENCY ?

< 3 sec. (query)	3-10 sec.	10 sec-1½ min.	1½ min-30 min.	>30 min.

DO YOU MAINTAIN DATA IN THE DATA BASE THAT IS SECURITY-SENSITIVE TO PERSONNEL OUTSIDE THE FIRM ?

never	almost never	on occasion	frequently	all of the time

DOES ANY DEPARTMENT MAINTAIN DATA WHICH MUST BE KEPT SECURE FROM OTHER PERSONNEL WITH ACCESS TO THE DATA BASE?

never	almost never	on occasion	frequently	all of the time

WHAT WAS YOUR ACTUAL PURCHASE PRICE [ ] Purchased at
OR ANNUAL LEASE PRICE ? [ ] Leased for

less $1000.00	$1000 - $5000	$5000 - $15,000	$15,000 - $50,000	over $50,000

WHAT IS THE *AVERAGE ANNUAL MAINTENANCE COST* [DBMS ONLY, BUT INCLUDE IN-HOUSE COSTS IF IN VENDOR CONTRACT] ?

less than $250	$250-$1000	$1000-$5000	$5000-$15,000	over $15,000

WHAT IS THE SKILL LEVEL REQUIRED BY THE REAL "END-USER" IN YOUR SPECIFIC APPLICATION ?

8th grade	HS diploma	college degree	computer experience	degree in computers

WHAT IS THE SKILL LEVEL REQUIRED BY THE PERSONNEL WHO MAINTAIN OR SERVICE THE SYSTEM FOR YOU ?

HS diploma	college degree	computer degree	computer science degree	advanced degree in computers

HOW FREQUENTLY DO YOU ADD ITEMS TO THE DATA BASE OR DELETE ITEMS FROM THE DATA BASE ?

monthly	weekly	< 10/day	< 500 a day	over 500/ day

HOW FREQUENTLY DO YOU CHANGE THE GENERAL STRUCTURE OF THE DATA BASE ?

never	less than yearly	yearly	twice yearly	quarterly or less

CAN YOUR OWN STAFF MODIFY THE SYSTEM TO ALLOW AN EXPANSION OF THE PRESENT APPLICATION ?

No, even with help	With help	Probably	For limited expension	For any desired task

WHAT IS THE AVERAGE RESPONSE TIME WHEN YOU REQUEST HELP FROM THE SERVICE REPRESENTATIVE ?

No rep available	return call w/i 2 days	Personal visit from rep: w/i a week	w/i 2 days	within ½ day

HOW IMPORTANT IS STANDARDIZATION TO YOUR APPLICATION ? (GOVERNMENT CONTRACTS; NETWORK CAPATABILITY; Etc.)

no impact	little impact	a consideration	accerated impact	major impact

HOW MANY USERS OF THE DBMS YOU CHOSE DID YOUR FIRM CONTACT PRIOR TO SELECTING THAT SYSTEM ?

none	1 - 2	3 - 5	5 - 10	over 10

FACTORS

PLEASE RANK-ORDER THE SEVEN FACTORS ON THE RIGHT TO SHOW THE PROPORTIONAL WEIGHT ACCORDED EACH AREA AT THE TIME THE DBMS WAS SELECTED. (1= MOST IMPORTANT AREA; 7= LEAST IMPORTANT AREA)

____ - - - - COST - - - - ____
____ HARDWARE AVAILABLE ____
____ PEER RECOMMENDATION ____
____ - SECURITY OPTIONS - ____
____ - USER SKILL-LEVEL - ____
REQUIREMENTS
____ STAFF SKILL-LEVEL ____
REQUIREMENTS
____ - STANDARDIZATION - ____

PLEASE RANK-ORDER THE SAME SEVEN FACTORS, GIVEN THAT YOU ARE FAMILIAR WITH AT LEAST ONE DBMS AND HAVE BEEN HIRED TO ACT AS A CONSULTANT FOR A FIRM WHICH IS INTERESTED IN SELECTING A DBMS. (1= AREA WHICH SHOULD BE GIVEN MOST CONSIDERATION IN YOUR TYPE OF APPLICATION; 7= AREA OF LEAST SIGNIFICANCE FOR YOUR TYPE OF APPLICATION)

Figure 3.1

indicator of relative user size, and that the break points were acceptable for the prospective-user model.

What is the average number of records you maintain on file in the data base ?

The question was designed to determine if the relative number of records on file -- or the resulting amount of auxiliary memory storage demanded by the number of records on file -- was a determinant in the choice of a DBMS.

The majority of users who replied maintain either large or very large data bases, so the present users' replies may indicate that the scale is too low. It was decided that the target user may need a scale which allows the choice of a smaller data base, so the scale category break points were retained for the model. Since the vendors have provided names of primarily large data base users, the bias reflected by the present users' replies is not considered a reflection of the primary audience of the model.

A few members of the committee felt that the question might be reworded to reflect "byte" size instead of the number of records. This would be a valid criticism of the model for a manager with extensive computer technological background, but it is likely that the manager for whom the model is designed is more familiar with records than the

computer-science terminology of bits, bytes and word lengths. Even the most non-technically oriented manager is going to be familiar with the gross number of records which are presently maintained on file by the application for which the DBMS is being selected. In order to maintain this reference criterion for the target user, the category descriptors using records instead of bytes were maintained for the selection model.

What is the maximum retrieval time you can accept and still remain at full operational efficiency ?

The industry standard changed during the interim between the design of the user questionnaire and the final production of the selection model. The author believes the new standards may be unrealistically biased toward the extremely short response times, but the committee decided that the model should reflect the industry standard. As a compromise, the author included an explanation of how the prospective user should realistically appraise the scale figures in the verbal explanation of the model.

Do you maintain data in the data base that is security sensitive to personnel outside the firm ?

Does any department maintain data which must be kept secure from other personnel with access to the data base ?

These two questions were designed to establish the difference between general data base access rights and the requirement for field-level security. Response to the user questionnaire and discussion with the committee indicated that these were valid questions, and that the range of values offered is satisfactory to establish this difference in user requirement.

What was your actual purchase price or annual lease price ?

This question was designed to find out how closely the vendor-advertised purchase and lease prices reflected the contractual prices paid by the users. The prices used as a base line were selected from the DBMS Technical Report [Slo, 78] and discussions with the committee. Of note is that the prospective user would probably make some type of cost-effectiveness or cost-value study of the final number of contending vendor packages, so this question was eliminated from the final model. The only question in the model which refers to cost is the area in the filter which is designed to eliminate those DBMS's which are obviously

unsuitable for the intended application as a result of a major difference between the system cost and the budget ceiling established for the application.

What is the skill level required by the real "end-user" in your specific application ?

What is the skill level required by the personnel who maintain or service the system for you ?

These two questions were designed to check the validity of vendor claims for minimum skill levels required for acceptable interaction with the system -- both at the end-user and systems personnel levels. The author felt the levels might be low-biased, but comments by the present users indicated that the ranges are realistic, or possibly even slightly high-biased. The majority of present users reported required end-user skill level at the 8th grade to high school diploma level, and the required systems personnel skill at the high school diploma to college degree (non-computer science) level. The ranges were left intact for the model.

How frequently do you add items to the data base or delete items from the data base ?

There was some feeling that this scale might also be

moderately low-biased, but user response indicates that the scales are generally realistic for the commercial user.

How frequently do you change the general structure of the data base ?

Discussion with committee members, particularly Dr. Maryanski, indicated that the commercial users of DBMS's seldom restructured their data bases. The academic experience of the author indicated that a restructuring should be conducted on an approximately semi-annual basis for any large data base. User response was nearly unanimous, however, in that restructuring was very seldom conducted on large data bases. This is an area which would be of significant interest for another study (e.g. how much efficiency is lost as a result of either the unwillingness or the incapability of the commercial user to restructure the data base on a regular interval schedule), but it was determined that this area was peripheral to the model. The ranges for the restructuring question were maintained in the selection model.

Can your staff modify the system to allow an expansion of the present application ?

When the questionnaire was written, the author was

unaware of the prevalence of vendor contracts specifically prohibiting users from modifying the system software, or the general practice of prohibiting users from securing a copy of the source code. The users contacted were all using major-vendor DBMS's, so there is less probability of those particular vendors going out of business abruptly. The prospective user of the model may not be interested in one of the larger vendor packages, however, so the model includes the question plus some explanation concerning the necessity for gaining access to the source code if the vendor is very small, new to the business, or if there is any indication that the vendor might not be available in the reasonable future.

What is the average response time when you request help from the service representative ?

This scale had to be restructured because of the replies from present users. The scale was originally based on a geographic span-of-control concept, since the author assumed that the vendor representatives would make personal visits in case of software failure. Response to the user questionnaire indicated that the majority of present users had never had a vendor representative visit the user site after the initial installation phase was completed. The

present users reported a preponderance of vendor-user contact by telephone, so the model was modified to reflect this finding.

How important is standardization to your application ?

The requirement for standardization is difficult to define in quantitative terms. The users appeared to view the provided scale as a reasonable measure of the area, so the scale was left intact.

How many users of the DBMS you chose did your firm contact prior to selecting the system ?

A few of the users who replied to the questionnaire were in the initial installation phase of that particular DBMS by the primary vendor. There was some indication that most of the present users who were not one of the initial installations used "other-user satisfaction" as one of the primary selection criterion for their application. Study of the type of DBMS selected by the present users indicates that the type of DBMS which was selected was not the optimal DBMS available at the time of selection. The dependence on comments from other users is not a realistic sole criterion for DBMS selection, although other user

comments are a valuable source of information about the service reputation of a vendor. This question was altered in the final model to demonstrate that user environment is a major determinant of DBMS field performance and to encourage prospective users to solicit comments primarily from users who have an application similar to the one for which the DBMS is being selected.

3.3 By-Respondent Analysis

The author expected that there would be some sort of correlation between the type of DBMS chosen and the general type of application for which the DBMS was chosen. The study of present users did not demonstrate any definable correlation between DBMS chosen and type of application, except for the obvious exception where the type of hardware possessed limited the choice of DBMS's to a single package.

Discussions with the present users provided much more data than appears on the questionnaires returned by the users. In order to provide background data for future work in this area, a general verbal description of each user is provided to demonstrate some of the factors which were used as primary selection criteria by actual users. It is very obvious that most of the users did not use a model such as this one during the selection of the DBMS. This may appear to contradict the validity of developing a rigorous

selection model. The author feels that the failure of the users to employ a selection model graphically demonstrates the need for a comprehensive model at the present time. If the present users had had access to a good selection model, there is an increased probability that there would have been more correlation between the DBMS selected and the user requirements.

USER A -- TOTAL (See Figure 3.2)

A large university, with over 100 entry sites and in excess of one million records on file. With primarily an academic requirement, retrieval times which might be slow for commercial applications are considered adequate by the institution. Privacy laws and the amount of data on students maintained in the data base demand some internal data base security options (for which TOTAL was not the best choice). The purchase price is within the ranges in the DBMS Technical Report (Slo, 78), and the annual cost of maintenance is within the range advertised by CINCOM.

End-user and staff skill level are about average, with the user stating a high school diploma requirement for both areas. The rate of insertions and deletions is moderately high, and the user reports a general restructuring of the data base on a yearly basis because of the need to reflect yearly academic class schedule changes.

With the university computer personnel available, this user indicated the staff could modify the system for almost any task, although CINCOM does not provide a source listing and excludes the user from modifying the system.

Standardization is of moderate impact (also not a good reason for selecting TOTAL); the hardware available at the time of purchase was listed as the second most important criterion in the DBMS selection process.

WHAT IS THE MAXIMUM NUMBER OF USERS WHO COULD ACCESS YOUR DATA BASE AT ANY ONE GIVEN TIME ?

One	2 - 4	5 - 20	20-100	over 100 ✓

WHAT IS THE AVERAGE NUMBER OF RECORDS YOU MAINTAIN ON FILE IN THE DATA BASE ?

< 1000	1000-10,000	10,000-100,000	100,000-500,000	>½ million ✓

WHAT IS THE MAXIMUM RETRIEVAL TIME YOU CAN ACCEPT AND STILL REMAIN AT FULL OPERATIONAL EFFICIENCY ?

< 3 sec. (query)	3-10 sec.	10 sec-1½ min. ✓	1½ min-30 min.	>30 min.

DO YOU MAINTAIN DATA IN THE DATA BASE THAT IS SECURITY-SENSITIVE TO PERSONNEL OUTSIDE THE FIRM ?

never ✓	almost never	on occasion	frequently	all of the time

DOES ANY DEPARTMENT MAINTAIN DATA WHICH MUST BE KEPT SECURE FROM OTHER PERSONNEL WITH ACCESS TO THE DATA BASE?

never	almost never	on occasion	frequently	all of the time ✓

WHAT WAS YOUR ACTUAL PURCHASE PRICE [ ] Purchased at
OR ANNUAL LEASE PRICE ? [ ] Leased for

less $1000.00	$1000 - $5000	$5000 - $15,000	$15,000 - $50,000 ✓	over $50,000

WHAT IS THE AVERAGE ANNUAL MAINTENANCE COST (DBMS ONLY, BUT INCLUDE IN-HOUSE COSTS IF IN VENDOR CONTRACT) ?

less than $250	$250-$1000	$1000-$5000 ✓	$5000-$15,000	over $15,000

WHAT IS THE SKILL LEVEL REQUIRED BY THE REAL "END-USER" IN YOUR SPECIFIC APPLICATION ?

8th grade	HS diploma ✓	college degree	computer experience	degree in computers

WHAT IS THE SKILL LEVEL REQUIRED BY THE PERSONNEL WHO MAINTAIN OR SERVICE THE SYSTEM FOR YOU ?

HS diploma ✓	college degree	computer degree	degree computer science	advanced degree in computers

HOW FREQUENTLY DO YOU ADD ITEMS TO THE DATA BASE OR DELETE ITEMS FROM THE DATA BASE ?

monthly	weekly	< 10/day	< 500 a day ✓	over 500/ day

HOW FREQUENTLY DO YOU CHANGE THE GENERAL STRUCTURE OF THE DATA BASE ?

never	less than yearly	yearly ✓	twice yearly	quarterly or less

CAN YOUR OWN STAFF MODIFY THE SYSTEM TO ALLOW AN EXPANSION OF THE PRESENT APPLICATION ?

No, even with help	With help	Probably	For limited expansion	For any desired task ✓

WHAT IS THE AVERAGE RESPONSE TIME WHEN YOU REQUEST HELP FROM THE SERVICE REPRESENTATIVE ?

No rep available	return call w/i 2 days ✓	Personal visit w/i a week	visit w/i 2 days	visit within ½ day

HOW IMPORTANT IS STANDARDIZATION TO YOUR APPLICATION ? [GOVERNMENT CONTRACTS; NETWORK CAPATABILITY; Etc.]

no impact	little impact	a consideration	moderate impact ✓	major impact

HOW MANY USERS OF THE DBMS YOU CHOSE DID YOUR FIRM CONTACT PRIOR TO SELECTING THAT SYSTEM ?

none	1 - 2	3 - 5 ✓	5 - 10	over 10

PLEASE RANK-ORDER THE SEVEN FACTORS ON THE RIGHT TO SHOW THE PROPORTIONAL WEIGHT ACCORDED EACH AREA AT THE TIME THE DBMS WAS SELECTED. (1= MOST IMPORTANT AREA; 7= LEAST IMPORTANT AREA)

FACTORS	
6	- - - - COST - - - - 6
2	HARDWARE AVAILABLE 2
3	PEER RECOMMENDATION 3
4	- SECURITY OPTIONS - 4
7	- USER SKILL-LEVEL - 7
	REQUIREMENTS
1	STAFF SKILL-LEVEL 1
	REQUIREMENTS
5	- STANDARDIZATION - 5

PLEASE RANK-ORDER THE SAME SEVEN FACTORS, GIVEN THAT YOU ARE FAMILIAR WITH AT LEAST ONE DBMS AND HAVE BEEN HIRED TO ACT AS A CONSULTANT FOR A FIRM WHICH IS INTERESTED IN SELECTING A DBMS. (1= AREA WHICH SHOULD BE GIVEN MOST CONSIDERATION IN YOUR TYPE OF APPLICATION; 7= AREA OF LEAST SIGNIFICANCE FOR YOUR TYPE OF APPLICATION)

Figure 3.2 User A

USER B -- TOTAL [See Figure 3.3]

A large government institution, with over a hundred remote entry sites and nearly 12 million records. Maximum acceptable retrieval times was in the 3010 second range,but there was a requirement for security lock-outs and record-level security all of the time.

The purchase price was $38,000 in 1972, with an annual maintenance contract of $1,000. End-user skills required were assessed at the 8th grade level, but the systems skill level was college degree. Additions and deletions are now processed using a macro data language, and structural reorganization is done on a less-than-yearly basis as a result of the major expense involved.

Some modifications have been performed to meet the security requirements, but there is only a limited on-board modification capability within the staff. The DBA stated that there had never been a contact with the vendor after the installation, although service was available if needed.

Only 3 or 4 users were contacted prior to selection, but this is one of the first CINCOM installations. The user has IBM 360/65's and 370/168's now, but, at the time of purchase, TOTAL was the only package that would fit the hardware possessed. Hardware was thus the most important consideration at the time of purchase, but would be the second least important criterion now. The user said a DBMS with better security features would be a better choice.

WHAT IS THE MAXIMUM NUMBER OF USERS WHO COULD ACCESS YOUR DATA BASE AT ANY ONE GIVEN TIME ?

One	2 - 4	5 - 20	20-100	over 100 ✓

WHAT IS THE AVERAGE NUMBER OF RECORDS YOU MAINTAIN ON FILE IN THE DATA BASE ?

< 1000	1000-10,000	10,000-100,000	100,000-500,000	>½ million ✓

WHAT IS THE MAXIMUM RETRIEVAL TIME YOU CAN ACCEPT AND STILL REMAIN AT FULL OPERATIONAL EFFICIENCY ?

< 3 sec. (query)	3-10 sec. ✓	10 sec-1½ min.	1½ min-30 min.	>30 min.

DO YOU MAINTAIN DATA IN THE DATA BASE THAT IS SECURITY-SENSITIVE TO PERSONNEL OUTSIDE THE FIRM ?

never	almost never	on occasion	frequently	all of the time ✓

DOES ANY DEPARTMENT MAINTAIN DATA WHICH MUST BE KEPT SECURE FROM OTHER PERSONNEL WITH ACCESS TO THE DATA BASE?

never	almost never	on occasion	frequently	all of the time ✓

WHAT WAS YOUR ACTUAL PURCHASE PRICE (✓) Purchased at
OR ANNUAL LEASE PRICE ? ( ) Leased for

less $1000.00	$1000 - $5000	$5000-$15,000	$15,000-$50,000 ✓	over $50,000

WHAT IS THE AVERAGE ANNUAL MAINTENANCE COST (DBMS ONLY, BUT INCLUDE IN-HOUSE COSTS IF IN VENDOR CONTRACT) ?

less than $250	$250-$1000	$1000-$5000 ✓	$5000-$15,000	over $15,000

WHAT IS THE SKILL LEVEL REQUIRED BY THE REAL "END-USER" IN YOUR SPECIFIC APPLICATION ?

8th grade ✓	HS diploma	college degree	computer experience	degree in computers

WHAT IS THE SKILL LEVEL REQUIRED BY THE PERSONNEL WHO MAINTAIN OR SERVICE THE SYSTEM FOR YOU ?

HS diploma	college degree	computer degree ✓	degree computer science	advanced degree in computers

HOW FREQUENTLY DO YOU ADD ITEMS TO THE DATA BASE OR DELETE ITEMS FROM THE DATA BASE ?

monthly	weekly	< 10/day	< 500 a day ✓	over 500/ day

HOW FREQUENTLY DO YOU CHANGE THE GENERAL STRUCTURE OF THE DATA BASE ?

never	less than yearly ✓	yearly	twice yearly	quarterly or less

CAN YOUR OWN STAFF MODIFY THE SYSTEM TO ALLOW AN EXPANSION OF THE PRESENT APPLICATION ?

No, even with help	With help	Probably	For limited expansion ✓	For any desired task

WHAT IS THE AVERAGE RESPONSE TIME WHEN YOU REQUEST HELP FROM THE SERVICE REPRESENTATIVE ?

No rep available	return call w/i 2 days	Personal visit from rep: w/i a week	w/i 2 days ✓	within ½ day

HOW IMPORTANT IS STANDARDIZATION TO YOUR APPLICATION ? (GOVERNMENT CONTRACTS; NETWORK CAPATABILITY; Etc.)

no impact	little impact	a consideration	moderate impact ✓	major impact

HOW MANY USERS OF THE DBMS YOU CHOSE DID YOUR FIRM CONTACT PRIOR TO SELECTING THAT SYSTEM ?

none	1 - 2	3 - 5 ✓	5 - 10	over 10

FACTORS

PLEASE RANK-ORDER THE SEVEN FACTORS ON THE RIGHT TO SHOW THE PROPORTIONAL WEIGHT ACCORDED EACH AREA AT THE TIME THE DBMS WAS SELECTED. (1= MOST IMPORTANT AREA; 7= LEAST IMPORTANT AREA)

2	COST	3
1	HARDWARE AVAILABLE	6
7	PEER RECOMMENDATION	7
6	SECURITY OPTIONS	5
5	USER SKILL-LEVEL REQUIREMENTS	2
3	STAFF SKILL-LEVEL REQUIREMENTS	1
4	STANDARDIZATION	4

PLEASE RANK-ORDER THE SAME SEVEN FACTORS, GIVEN THAT YOU ARE FAMILIAR WITH AT LEAST ONE DBMS AND HAVE BEEN HIRED TO ACT AS A CONSULTANT FOR A FIRM WHICH IS INTERESTED IN SELECTING A DBMS. (1= AREA WHICH SHOULD BE GIVEN MOST CONSIDERATION IN YOUR TYPE OF APPLICATION; 7= AREA OF LEAST SIGNIFICANCE FOR YOUR TYPE OF APPLICATION)

Figure 3.3 User B

USER C -- TOTAL [See Figure 3.4]

A large eastern city system, with 16 entry sites and approximately 140,000 records on file. Maximum retrieval was listed as 5 seconds, and no DBMS security was required at the data base or lower levels. [The DBA stated that the privacy restrictions do not affect municipal governments in that state.]

The system is leased at $1,000 per month, with a maintenance agreement as part of the package agreement from NCR. End-user skill requirement is low [8th grade] but the systems staff skill requirement is high [degree in computer science]. Insertions and deletions are done at the rate of approximately 400 a day, and the system has never been reorganized. The user stated that modifications could be made with minimal help from the vendor.

Vendor on-site contact has been non-existent, but any telephonic requests are returned within 2 days. Only three users were contacted prior to the selection of the system, and standardization is considered to have no impact.

Hardware was considered the primary criterion at the time of acquisition, but would be the least important factor now. Because of a desire for future interaction with the federal government, standardization would be the primary factor in any future selections.

WHAT IS THE MAXIMUM NUMBER OF USERS WHO COULD ACCESS YOUR DATA BASE AT ANY ONE GIVEN TIME ?

One	2 - 4	5 - 20 ✓	20-100	over 100

WHAT IS THE AVERAGE NUMBER OF RECORDS YOU MAINTAIN ON FILE IN THE DATA BASE ?

< 1000	1000-10,000	10,000-100,000	100,000 ✓ 500,000	>½ million

WHAT IS THE MAXIMUM RETRIEVAL TIME YOU CAN ACCEPT AND STILL REMAIN AT FULL OPERATIONAL EFFICIENCY ?

< 3 sec. (query)	3-10 ✓ sec.	10 sec-1½ min.	1½ min-30 min.	>30 min.

DO YOU MAINTAIN DATA IN THE DATA BASE THAT IS SECURITY-SENSITIVE TO PERSONNEL OUTSIDE THE FIRM ?

never ✓	almost never	on occasion	frequently	all of the time

DOES ANY DEPARTMENT MAINTAIN DATA WHICH MUST BE KEPT SECURE FROM OTHER PERSONNEL WITH ACCESS TO THE DATA BASE?

never	almost never ✓	on occasion	frequently	all of the time

WHAT WAS YOUR ACTUAL PURCHASE PRICE ( ) Purchased at
OR ANNUAL LEASE PRICE ? ( ) Leased for

less $1000.00	$1000 - $5000 ✓	$5000-$15,000	$15,000-$50,000	over $50,000

WHAT IS THE AVERAGE ANNUAL MAINTENANCE COST (DBMS ONLY, BUT INCLUDE IN-HOUSE COSTS IF IN VENDOR CONTRACT) ?

NA

less than $250	$250-$1000	$1000-$5000	$5000-$15,000	over $15,000

WHAT IS THE SKILL LEVEL REQUIRED BY THE REAL "END-USER" IN YOUR SPECIFIC APPLICATION ?

8th grade ✓	HS diploma	college degree	computer experience	degree in computers

WHAT IS THE SKILL LEVEL REQUIRED BY THE PERSONNEL WHO MAINTAIN OR SERVICE THE SYSTEM FOR YOU ?

HS diploma	college degree	computer degree	degree ✓ computer science	advanced degree in computers

HOW FREQUENTLY DO YOU ADD ITEMS TO THE DATA BASE OR DELETE ITEMS FROM THE DATA BASE ?

monthly	weekly	< 10/day	< 500 a day ✓	over 500/ day

HOW FREQUENTLY DO YOU CHANGE THE GENERAL STRUCTURE OF THE DATA BASE ?

never ✓	less than yearly	yearly	twice yearly	quarterly or less

CAN YOUR OWN STAFF MODIFY THE SYSTEM TO ALLOW AN EXPANSION OF THE PRESENT APPLICATION ?

No, even with help	With help ✓	Probably	For limited expansion	For any desired task

WHAT IS THE AVERAGE RESPONSE TIME WHEN YOU REQUEST HELP FROM THE SERVICE REPRESENTATIVE ?

No rep available	return call w/i 2 days ✓	Personal visit from rep: w/i a week	w/i 2 days	within ½ day

HOW IMPORTANT IS STANDARDIZATION TO YOUR APPLICATION ? (GOVERNMENT CONTRACTS; NETWORK CAPATABILITY; Etc.)

no impact ✓	little impact	a considerat-ion	accerated im-pact	major impact

HOW MANY USERS OF THE DBMS YOU CHOSE DID YOUR FIRM CONTACT PRIOR TO SELECTING THAT SYSTEM ?

none	1 - 2	3 - 5 ✓	5 - 10	over 10

FACTORS

PLEASE RANK-ORDER THE SEVEN FACTORS ON THE RIGHT TO SHOW THE PROPORTIONAL WEIGHT ACCORDED EACH AREA AT THE TIME THE DBMS WAS SELECTED. (1= MOST IMPORTANT AREA; 7= LEAST IMPORTANT AREA)

4	COST	5
1	HARDWARE AVAILABLE	7
5	PEER RECOMMENDATION	6
6	SECURITY OPTIONS	3
2	USER SKILL-LEVEL REQUIREMENTS	4
3	STAFF SKILL-LEVEL REQUIREMENTS	2
7	STANDARDIZATION	1

PLEASE RANK-ORDER THE SAME SEVEN FACTORS, GIVEN THAT YOU ARE FAMILIAR WITH AT LEAST ONE DBMS AND HAVE BEEN HIRED TO ACT AS A CONSULTANT FOR A FIRM WHICH IS INTERESTED IN SELECTING A DBMS. (1= AREA WHICH SHOULD BE GIVEN MOST CONSIDERATION IN YOUR TYPE OF APPLICATION; 7= AREA OF LEAST SIGNIFICANCE FOR YOUR TYPE OF APPLICATION)

Figure 3.4 User C

USER D -- TOTAL [See Figure 3.5]

A major southern wholesale house, with 450 entry sites and over a million records on file. Query response time requirements are 4-5 seconds maximum, and exterior security is considered mandatory all of the time, with interior security a frequent consideration.

The system was purchased for $30,000, with an included maintenance contract for $1,500 a year. The system cost is moderately low; the maintenance agreement is slightly high.

Both end-user and systems skill level were judged to be at the high school diploma level. Insertions and deletions (primarily inventory adjustments) are conducted at a rate of over 500 a day, but data base reorganization is done less than yearly. Some systems modification has already been performed in order to interface with DATBAS, which is another DBMS in use by the firm.

Vendor service response was said to be good, with a return telephone call within one day. Standardization is of no concern, and only two previous users were contacted prior to selection of the system.

The primary selection criterion at the time of system purchase was hardware; for a reselection process, the staff skill level requirement would be the most important point.

WHAT IS THE MAXIMUM NUMBER OF USERS WHO COULD ACCESS YOUR DATA BASE AT ANY ONE GIVEN TIME ?

One	2 - 4	5 - 20	20-100	over 100 ✓

WHAT IS THE AVERAGE NUMBER OF RECORDS YOU MAINTAIN ON FILE IN THE DATA BASE ?

< 1000	1000-10,000	10,000-100,000	100,000-500,000	>½ million ✓

WHAT IS THE MAXIMUM RETRIEVAL TIME YOU CAN ACCEPT AND STILL REMAIN AT FULL OPERATIONAL EFFICIENCY ?

< 3 sec. (query)	3-10 sec. ✓	10 sec-1½ min.	1½ min-30 min.	>30 min.

DO YOU MAINTAIN DATA IN THE DATA BASE THAT IS SECURITY-SENSITIVE TO PERSONNEL OUTSIDE THE FIRM ?

never	almost never	on occasion	freq-uently	all of the time ✓

DOES ANY DEPARTMENT MAINTAIN DATA WHICH MUST BE KEPT SECURE FROM OTHER PERSONNEL WITH ACCESS TO THE DATA BASE?

never	almost never	on occasion	freq-uently ✓	all of the time

WHAT WAS YOUR ACTUAL PURCHASE PRICE (✓) Purchased at
OR ANNUAL LEASE PRICE ? [ ] Leased for

less $1000.00	$1000 - $5000	$5000-$15,000	$15,000-$50,000 ✓	over $50,000

WHAT IS THE AVERAGE ANNUAL MAINTENANCE COST (DBMS ONLY, BUT INCLUDE IN-HOUSE COSTS IF IN VENDOR CONTRACT) ?

less than $250	$250-$1000	$1000-$5000 ✓	$5000-$15,000	over $15,000

WHAT IS THE SKILL LEVEL REQUIRED BY THE REAL "END-USER" IN YOUR SPECIFIC APPLICATION ?

8th grade	HS diploma ✓	college degree	computer exper-ience	degree in com-puters

WHAT IS THE SKILL LEVEL REQUIRED BY THE PERSONNEL WHO MAINTAIN OR SERVICE THE SYSTEM FOR YOU ?

HS diploma ✓	college degree	computer degree	computer science	advanced degree in computers

HOW FREQUENTLY DO YOU ADD ITEMS TO THE DATA BASE OR DELETE ITEMS FROM THE DATA BASE ?

monthly	weekly	< 10/day	< 500 a day	over 500/ day ✓

HOW FREQUENTLY DO YOU CHANGE THE GENERAL STRUCTURE OF THE DATA BASE ?

never	less than yearly ✓	yearly	twice yearly	quarterly or less

CAN YOUR OWN STAFF MODIFY THE SYSTEM TO ALLOW AN EXPANSION OF THE PRESENT APPLICATION ?

No, even with help	With help	Probably	For limited expansion	For any desired task ✓

WHAT IS THE AVERAGE RESPONSE TIME WHEN YOU REQUEST HELP FROM THE SERVICE REPRESENTATIVE ?

No rep available	return call w/i 2 days	Personal visit w/i a week	visit from rep w/i 2 days ✓	within ½ day

HOW IMPORTANT IS STANDARDIZATION TO YOUR APPLICATION ?
(GOVERNMENT CONTRACTS; NETWORK CAPATABILITY; Etc.)

no impact ✓	little impact	a consider-ation	moderate im-pact	major impact

HOW MANY USERS OF THE DBMS YOU CHOSE DID YOUR FIRM CONTACT PRIOR TO SELECTING THAT SYSTEM ?

none	1 - 2 ✓	3 - 5	5 - 10	over 10

FACTORS

PLEASE RANK-ORDER THE SEVEN FACTORS ON THE RIGHT TO SHOW THE PROPORTIONAL WEIGHT ACCORDED EACH AREA AT THE TIME THE DBMS WAS SELECTED. (1= MOST IMPORTANT AREA; 7= LEAST IMPORTANT AREA)

5	COST	5
1	HARDWARE AVAILABLE	2
4	PEER RECOMMENDATION	4
6	SECURITY OPTIONS	6
7	USER SKILL-LEVEL REQUIREMENTS	7
2	STAFF SKILL-LEVEL REQUIREMENTS	1
3	STANDARDIZATION	3

PLEASE RANK-ORDER THE SAME SEVEN FACTORS, GIVEN THAT YOU ARE FAMILIAR WITH AT LEAST ONE DBMS AND HAVE BEEN HIRED TO ACT AS A CONSULTANT FOR A FIRM WHICH IS INTERESTED IN SELECTING A DBMS. (1= AREA WHICH SHOULD BE GIVEN MOST CONSIDERATION IN YOUR TYPE OF APPLICATION; 7= AREA OF LEAST SIGNIFICANCE FOR YOUR TYPE OF APPLICATION)

Figure 3.5 User D

USER E -- System 2000 (See Figure 3.6)

A large southwestern manufacturing company, with only three entry sites, but over a half million records. The maximum response time to queries was 3 to 10 seconds.

Exterior security was needed on occasion, but interior security was almost never demanded -- primarily as a result of the limited number of entry sites.

The system was purchased for less than $50,000, with a maintenance contract of over $1,000 a year. The end-user skill requirement is a high school diploma, with a systems skill level of a non-computer science college degree. The insertion and deletion rate exceeds 500 a day, and, in an unusual response, the user stated that reorganization was conducted semi-annually. The staff could modify the system for any desired result (questionable with MRI's normal restrictions); standardization was of moderate importance (not a good selection criterion for choosing System 2000).

Four other users were contacted prior to the selection of the system, and hardware possessed was the major point of interest at the time of purchase and is still considered the most important selection criterion (even though the firm has a 370/158, which allows many different options). The second-most important area for selection was the skill level required for staff personnel, with the third- most important criterion as peer recommendation (higher than is normally admitted.)

WHAT IS THE MAXIMUM NUMBER OF USERS WHO COULD ACCESS YOUR DATA BASE AT ANY ONE GIVEN TIME ?

One | 2 - 4 ✓ | 5 - 20 | 20-100 | over 100

WHAT IS THE AVERAGE NUMBER OF RECORDS YOU MAINTAIN ON FILE IN THE DATA BASE ?

< 1000 | 1000-10,000 | 10,000-100,000 | 100,000-500,000 | >½ million ✓

WHAT IS THE MAXIMUM RETRIEVAL TIME YOU CAN ACCEPT AND STILL REMAIN AT FULL OPERATIONAL EFFICIENCY ?

< 3 sec. (query) | 3-10 sec. ✓ | 10 sec-1½ min. | 1½ min-30 min. | >30 min.

DO YOU MAINTAIN DATA IN THE DATA BASE THAT IS SECURITY-SENSITIVE TO PERSONNEL OUTSIDE THE FIRM ?

never | almost never | on occasion ✓ | frequently | all of the time

DOES ANY DEPARTMENT MAINTAIN DATA WHICH MUST BE KEPT SECURE FROM OTHER PERSONNEL WITH ACCESS TO THE DATA BASE?

never | almost never ✓ | on occasion | frequently | all of the time

WHAT WAS YOUR ACTUAL PURCHASE PRICE (✓) Purchased at
OR ANNUAL LEASE PRICE ? ( ) Leased for

less $1000.00 | $1000-$5000 | $5000-$15,000 ✓ | $15,000-$50,000 | over $50,000

WHAT IS THE AVERAGE ANNUAL MAINTENANCE COST (DBMS ONLY, BUT INCLUDE IN-HOUSE COSTS IF IN VENDOR CONTRACT) ?

less than $250 | $250-$1000 ✓ | $1000-$5000 | $5000-$15,000 | over $15,000

WHAT IS THE SKILL LEVEL REQUIRED BY THE REAL "END-USER" IN YOUR SPECIFIC APPLICATION ?

8th grade ✓ | HS diploma | college degree | computer experience | degree in computers

WHAT IS THE SKILL LEVEL REQUIRED BY THE PERSONNEL WHO MAINTAIN OR SERVICE THE SYSTEM FOR YOU ?

HS diploma | college degree ✓ | computer degree | degree computer science | advanced degree in computers

HOW FREQUENTLY DO YOU ADD ITEMS TO THE DATA BASE OR DELETE ITEMS FROM THE DATA BASE ?

monthly | weekly | < 10/day | < 500 a day | over 500/ day ✓

HOW FREQUENTLY DO YOU CHANGE THE GENERAL STRUCTURE OF THE DATA BASE ?

never | less than yearly | yearly | twice yearly ✓ | quarterly or less

CAN YOUR OWN STAFF MODIFY THE SYSTEM TO ALLOW AN EXPANSION OF THE PRESENT APPLICATION ?

No, even with help | With help | Probably | For limited expansion | For any desired task ✓

WHAT IS THE AVERAGE RESPONSE TIME WHEN YOU REQUEST HELP FROM THE SERVICE REPRESENTATIVE ?

No rep available | return call w/i 2 days ✓ | Personal visit from rep w/i a week | w/i 2 days | within ½ day

HOW IMPORTANT IS STANDARDIZATION TO YOUR APPLICATION ? (GOVERNMENT CONTRACTS; NETWORK CAPATABILITY; Etc.)

no impact | little impact | a consideration | moderate impact ✓ | major impact

HOW MANY USERS OF THE DBMS YOU CHOSE DID YOUR FIRM CONTACT PRIOR TO SELECTING THAT SYSTEM ?

none | 1 - 2 | 3 - 5 ✓ | 5 - 10 | over 10

FACTORS

PLEASE RANK-ORDER THE SEVEN FACTORS ON THE RIGHT TO SHOW THE PROPORTIONAL WEIGHT ACCORDED EACH AREA AT THE TIME THE DBMS WAS SELECTED. (1= MOST IMPORTANT AREA; 7= LEAST IMPORTANT AREA)

5	COST	4
1	HARDWARE AVAILABLE	1
3	PEER RECOMMENDATION	5
6	SECURITY OPTIONS	6
4	USER SKILL-LEVEL REQUIREMENTS	3
2	STAFF SKILL-LEVEL REQUIREMENTS	2
7	STANDARDIZATION	7

PLEASE RANK-ORDER THE SAME SEVEN FACTORS, GIVEN THAT YOU ARE FAMILIAR WITH AT LEAST ONE DBMS AND HAVE BEEN HIRED TO ACT AS A CONSULTANT FOR A FIRM WHICH IS INTERESTED IN SELECTING A DBMS. (1= AREA WHICH SHOULD BE GIVEN MOST CONSIDERATION IN YOUR TYPE OF APPLICATION; 7= AREA OF LEAST SIGNIFICANCE FOR YOUR TYPE OF APPLICATION)

Figure 3.6 User E

USER F -- System 2000 (See Figure 3.7)

A large manufacturing concern, with several data bases operating within the system; nearly a hundred users; and an average of 14,000 records on any given data base. The data base retrieval time considered satisfactory was 10 seconds to one minute.

External security was almost never a consideration, but the user stated that internal security was done on a frequent basis (unresolved by the author). Purchase price was $95,000, with an annual maintenance contract of $8,500.

End-user skill level was listed as requiring a high school diploma, and staff skill level was listed at the college degree level. Insertions and deletions were made at a rate of less than 500 a day, but the system was reorganized quarterly -- a unique response.

The primary weight for the selection process was the necessity to port the system to both CDC and Univac; the user contacted twenty present users of the system prior to making the final DBMS selection. The user also stated that flexibility and expandability would be the major considered criteria in any future selection cycle.

WHAT IS THE MAXIMUM NUMBER OF USERS WHO COULD ACCESS YOUR DATA BASE AT ANY ONE GIVEN TIME ?

One	2 - 4	5 - 20	20-100 ✓	over 100

WHAT IS THE AVERAGE NUMBER OF RECORDS YOU MAINTAIN ON FILE IN THE DATA BASE ?

< 1000	1000-10,000	10,000-100,000 ✓	100,000-500,000	>½ million

WHAT IS THE MAXIMUM RETRIEVAL TIME YOU CAN ACCEPT AND STILL REMAIN AT FULL OPERATIONAL EFFICIENCY ?

< 3 sec. (query)	3-10 sec.	10 sec-1½ min. ✓	1½ min-30 min.	>30 min.

DO YOU MAINTAIN DATA IN THE DATA BASE THAT IS SECURITY-SENSITIVE TO PERSONNEL OUTSIDE THE FIRM ?

never	almost never ✓	on occasion	frequently	all of the time

DOES ANY DEPARTMENT MAINTAIN DATA WHICH MUST BE KEPT SECURE FROM OTHER PERSONNEL WITH ACCESS TO THE DATA BASE?

never	almost never	on occasion ✓	frequently	all of the time

WHAT WAS YOUR ACTUAL PURCHASE PRICE (✓) Purchased at
OR ANNUAL LEASE PRICE ? ( ) Leased for

less $1000.00	$1000 - $5000	$5000-$15,000	$15,000-$50,000 ✓	over $50,000

WHAT IS THE AVERAGE ANNUAL MAINTENANCE COST (DBMS ONLY, BUT INCLUDE IN-HOUSE COSTS IF IN VENDOR CONTRACT) ?

less than $250	$250-$1000	$1000-$5000 ✓	$5000-$15,000	over $15,000

WHAT IS THE SKILL LEVEL REQUIRED BY THE REAL "END-USER" IN YOUR SPECIFIC APPLICATION ?

8th grade	HS diploma ✓	college degree	computer experience	degree in computers

WHAT IS THE SKILL LEVEL REQUIRED BY THE PERSONNEL WHO MAINTAIN OR SERVICE THE SYSTEM FOR YOU ?

HS diploma	college degree ✓	computer degree	computer science	advanced degree in computers

HOW FREQUENTLY DO YOU ADD ITEMS TO THE DATA BASE OR DELETE ITEMS FROM THE DATA BASE ?

monthly	weekly	< 10/day	< 500 a day	over 500/ day ✓

HOW FREQUENTLY DO YOU CHANGE THE GENERAL STRUCTURE OF THE DATA BASE ?

never	less than yearly	yearly	twice yearly	qua terly or less ✓

CAN YOUR OWN STAFF MODIFY THE SYSTEM TO ALLOW AN EXPANSION OF THE PRESENT APPLICATION ?

No, even with help	With help	Probably	For limited expansion ✓	For any desired task

WHAT IS THE AVERAGE RESPONSE TIME WHEN YOU REQUEST HELP FROM THE SERVICE REPRESENTATIVE ?

No rep available	return call w/i 2 days	Person w/i a week ✓	visit from rep w/i 2 days	within ½ day

HOW IMPORTANT IS STANDARDIZATION TO YOUR APPLICATION ? (GOVERNMENT CONTRACTS; NETWORK CAPATABILITY; Etc.)

no impact	little impact	a consider-ation	moderate im-pact ✓	major impact

HOW MANY USERS OF THE DBMS YOU CHOSE DID YOUR FIRM CONTACT PRIOR TO SELECTING THAT SYSTEM ?

none	1 - 2	3 - 5	5 - 10	over 10 ✓

PLEASE RANK-ORDER THE SEVEN FACTORS ON THE RIGHT TO SHOW THE PROPORTIONAL WEIGHT ACCORDED EACH AREA AT THE TIME THE DBMS WAS SELECTED. (1= MOST IMPORTANT AREA; 7= LEAST IMPORTANT AREA)

FACTORS

7	- - - - COST - - - -	7
2	HARDWARE AVAILABLE	2
6	PEER RECOMMENDATION	6
5	- SECURITY OPTIONS -	5
4	- USER SKILL-LEVEL REQUIREMENTS -	4
3	STAFF SKILL-LEVEL REQUIREMENTS	3
1	- STANDARDIZATION -	1

PLEASE RANK-ORDER THE SAME SEVEN FACTORS, GIVEN THAT YOU ARE FAMILIAR WITH AT LEAST ONE DBMS AND HAVE BEEN HIRED TO ACT AS A CONSULTANT FOR A FIRM WHICH IS INTERESTED IN SELECTING A DBMS. (1= AREA WHICH SHOULD BE GIVEN MOST CONSIDERATION IN YOUR TYPE OF APPLICATION; 7= AREA OF LEAST SIGNIFICANCE FOR YOUR TYPE OF APPLICATION)

Figure 3.7 User F

USER G -- System 2000 [See Figure 3.8]

A large university, with over 100 entry sites and over 450,000 records. Because of the primarily academic type operations, retrieval times of ten seconds were considered adequate. Since the registrar maintains seven levels of security, there was an interface with a CICS TP monitor to insure exterior security.

Price was not disclosed, but there was no maintenance contract with the vendor. End-user and staff skill levels were about average at high school diploma and college degree respectively. Insertions and deletions were made at a rate of nearly 10,000 a day, and the system was reorganized on a yearly basis. The DBA felt that the staff could modify the system for any purpose.

The system is resident on an IBM 370/158, and the user talked to more than ten users of TOTAL, DATBAS and IMS prior to selecting System 2000. A major complaint of this user was the excessive reload and redefinition time required for such a large data base, with an average quoted time of three days on full shift [not, according to the users, anywhere near the figure originally quoted by a vendor]. The user also noted that the design was geared for good response with fewer keys, but degraded the ad hoc query responsiveness. User skill was the primary weighting factor during selection, with staff skill level next.

WHAT IS THE MAXIMUM NUMBER OF USERS WHO COULD ACCESS YOUR DATA BASE AT ANY ONE GIVEN TIME ?

One	2 - 4	5 - 20	20-100	over 100 ✓

WHAT IS THE AVERAGE NUMBER OF RECORDS YOU MAINTAIN ON FILE IN THE DATA BASE ?

< 1000	1000-10,000	10,000-100,000 ✓	100,000-500,000	>½ million

WHAT IS THE MAXIMUM RETRIEVAL TIME YOU CAN ACCEPT AND STILL REMAIN AT FULL OPERATIONAL EFFICIENCY ?

< 3 sec. (query)	3-10 sec.	10 sec-1½ min. ✓	1½ min-30 min.	>30 min.

DO YOU MAINTAIN DATA IN THE DATA BASE THAT IS SECURITY-SENSITIVE TO PERSONNEL OUTSIDE THE FIRM ?

never	almost never	on occasion ✓	frequently	all of the time

DOES ANY DEPARTMENT MAINTAIN DATA WHICH MUST BE KEPT SECURE FROM OTHER PERSONNEL WITH ACCESS TO THE DATA BASE?

never ✓	almost never	on occasion	frequently	all of the time

WHAT WAS YOUR ACTUAL PURCHASE PRICE (✓) Purchased at
OR ANNUAL LEASE PRICE ? ( ) Leased for

less $1000.00	$1000 - $5000	$5000-$15,000	$15,000-$50,000	over $50,000

WHAT IS THE AVERAGE ANNUAL MAINTENANCE COST (DBMS ONLY, BUT INCLUDE IN-HOUSE COSTS IF IN VENDOR CONTRACT) ?

NA

less than $250	$250-$1000	$1000-$5000	$5000-$15,000	over $15,000

WHAT IS THE SKILL LEVEL REQUIRED BY THE REAL "END-USER" IN YOUR SPECIFIC APPLICATION ?

8th grade	HS diploma ✓	college degree	computer experience	degree in computers

WHAT IS THE SKILL LEVEL REQUIRED BY THE PERSONNEL WHO MAINTAIN OR SERVICE THE SYSTEM FOR YOU ?

HS diploma	college degree ✓	computer degree	computer science	advanced degree in computers

HOW FREQUENTLY DO YOU ADD ITEMS TO THE DATA BASE OR DELETE ITEMS FROM THE DATA BASE ?

monthly	weekly	< 10/day	< 500 a day	over 500/ day ✓

HOW FREQUENTLY DO YOU CHANGE THE GENERAL STRUCTURE OF THE DATA BASE ?

never	less than yearly	yearly ✓	twice yearly	quarterly or less

CAN YOUR OWN STAFF MODIFY THE SYSTEM TO ALLOW AN EXPANSION OF THE PRESENT APPLICATION ?

No, even with help	With help	Probably	For limited expansion	For any desired task ✓

WHAT IS THE AVERAGE RESPONSE TIME WHEN YOU REQUEST HELP FROM THE SERVICE REPRESENTATIVE ?

No rep available	return call w/i 2 days	Personal visit w/i a week ✓	visit w/i 2 days	visit from rep within ½ day

HOW IMPORTANT IS STANDARDIZATION TO YOUR APPLICATION ? (GOVERNMENT CONTRACTS; NETWORK CAPATABILITY; Etc.)

no impact	little impact	a consideration ✓	moderate impact	major impact

HOW MANY USERS OF THE DBMS YOU CHOSE DID YOUR FIRM CONTACT PRIOR TO SELECTING THAT SYSTEM ?

none	1 - 2	3 - 5	5 - 10	over 10 ✓

FACTORS

PLEASE RANK-ORDER THE SEVEN FACTORS ON THE RIGHT TO SHOW THE PROPORTIONAL WEIGHT ACCORDED EACH AREA AT THE TIME THE DBMS WAS SELECTED. (1= MOST IMPORTANT AREA; 7= LEAST IMPORTANT AREA)

6	COST	6
4	HARDWARE AVAILABLE	4
5	PEER RECOMMENDATION	5
3	SECURITY OPTIONS	3
1	USER SKILL-LEVEL REQUIREMENTS	1
2	STAFF SKILL-LEVEL REQUIREMENTS	2
7	STANDARDIZATION	7

PLEASE RANK-ORDER THE SAME SEVEN FACTORS, GIVEN THAT YOU ARE FAMILIAR WITH AT LEAST ONE DBMS AND HAVE BEEN HIRED TO ACT AS A CONSULTANT FOR A FIRM WHICH IS INTERESTED IN SELECTING A DBMS. (1= AREA WHICH SHOULD BE GIVEN MOST CONSIDERATION IN YOUR TYPE OF APPLICATION; 7= AREA OF LEAST SIGNIFICANCE FOR YOUR TYPE OF APPLICATION)

Figure 3.8 User G

USER H -- IDMS (See Figure 3.9)

A large university, with nearly a hundred entry sites and a record base of over 100,000. Required response times were given as 10 seconds to 1½ minutes. Exterior security was demanded all of the time, and interior security was needed on occasion.

The system was purchased for over $50,000, and there is an annual maintenance contract for slightly under the $5000 per year break point. End-user skill level was given as high school diploma, and college degrees were required of the systems personnel.

Insertions and deletions were conducted on a once-a-month basis, with reorganization being accomplished on an annual basis. The user is prohibited from modification of the data base, but felt that the staff was able to do any modification necessary.

Standardization was emphatically not a necessity, with more than 60% of the selection weight based on just the present hardware (CDC). The next most important area was user skill level, followed by security options. In a reselection process, the user would place security options at the top of the criteria list, since there is some dissatisfaction with the present system security.

WHAT IS THE MAXIMUM NUMBER OF USERS WHO COULD ACCESS YOUR DATA BASE AT ANY ONE GIVEN TIME ?

One	2 - 4	5 - 20	✓ 20-100	over 100

WHAT IS THE AVERAGE NUMBER OF RECORDS YOU MAINTAIN ON FILE IN THE DATA BASE ?

< 1000	1000-10,000	10,000-100,000	✓ 100,000-500,000	> ½ million

WHAT IS THE MAXIMUM RETRIEVAL TIME YOU CAN ACCEPT AND STILL REMAIN AT FULL OPERATIONAL EFFICIENCY ?

< 3 sec. (query)	3-10 sec.	✓ 10 sec-1½ min.	1½ min-30 min.	>30 min.

DO YOU MAINTAIN DATA IN THE DATA BASE THAT IS SECURITY-SENSITIVE TO PERSONNEL OUTSIDE THE FIRM ?

never	almost never	on occasion	frequently	✓ all of the time

DOES ANY DEPARTMENT MAINTAIN DATA WHICH MUST BE KEPT SECURE FROM OTHER PERSONNEL WITH ACCESS TO THE DATA BASE?

never	almost never	✓ on occasion	frequently	all of the time

WHAT WAS YOUR ACTUAL PURCHASE PRICE (✓) Purchased at
OR ANNUAL LEASE PRICE ? ( ) Leased for

less than $1000.00	$1000 - $5000	$5000 - $15,000	$15,000 - $50,000	✓ over $50,000

WHAT IS THE AVERAGE ANNUAL MAINTENANCE COST (ODMS ONLY, BUT INCLUDE IN-HOUSE COSTS IF IN VENDOR CONTRACT) ?

less than $250	$250-$1000	$1000-$5000	✓ $5000-$15,000	over $15,000

WHAT IS THE SKILL LEVEL REQUIRED BY THE REAL "END-USER" IN YOUR SPECIFIC APPLICATION ?

8th grade	✓ HS diploma	college degree	computer experience	degree in computers

WHAT IS THE SKILL LEVEL REQUIRED BY THE PERSONNEL WHO MAINTAIN OR SERVICE THE SYSTEM FOR YOU ?

HS diploma	✓ college degree	computer degree	degree computer science	advanced degree in computers

HOW FREQUENTLY DO YOU ADD ITEMS TO THE DATA BASE OR DELETE ITEMS FROM THE DATA BASE ?

✓ monthly	weekly	< 10/day	< 500 a day	over 500/ day

HOW FREQUENTLY DO YOU CHANGE THE GENERAL STRUCTURE OF THE DATA BASE ?

never	✓ less than yearly	yearly	twice yearly	quarterly or less

(NA)

CAN YOUR OWN STAFF MODIFY THE SYSTEM TO ALLOW AN EXPANSION OF THE PRESENT APPLICATION ?

No, even with help	With help	Probably	✓ For limited expansion	For any desired task

WHAT IS THE AVERAGE RESPONSE TIME WHEN YOU REQUEST HELP FROM THE SERVICE REPRESENTATIVE ?

No rep available	return call w/i 2 days	Personal visit w/i a week	✓ r/i 2 days	less than w/in ½ day

HOW IMPORTANT IS STANDARDIZATION TO YOUR APPLICATION ? (GOVERNMENT CONTRACTS; NETWORK CAPATABILITY; Etc.)

✓ no impact	little impact	a consideration	separated impact	major impact

HOW MANY USERS OF THE ODMS YOU CHOSE DID YOUR FIRM CONTACT PRIOR TO SELECTING THAT SYSTEM ?

none	1 - 2	3 - 5	5 - 10	✓ over 10

PLEASE RANK-ORDER THE SEVEN FACTORS ON THE RIGHT TO SHOW THE PROPORTIONAL WEIGHT ACCORDED EACH AREA AT THE TIME THE ODMS WAS SELECTED. (1= MOST IMPORTANT AREA; 7= LEAST IMPORTANT AREA)

FACTORS

4	COST	4
1	HARDWARE AVAILABLE	1
3	PEER RECOMMENDATION	2
5	SECURITY OPTIONS	5
2	USER SKILL-LEVEL REQUIREMENTS	3
6	STAFF SKILL-LEVEL REQUIREMENTS	6
7	STANDARDIZATION	7

PLEASE RANK-ORDER THE SAME SEVEN FACTORS, GIVEN THAT YOU ARE FAMILIAR WITH AT LEAST ONE ODMS AND HAVE BEEN HIRED TO ACT AS A CONSULTANT FOR A FIRM WHICH IS INTERESTED IN SELECTING A ODMS. (1= AREA WHICH SHOULD BE GIVEN MOST CONSIDERATION IN YOUR TYPE OF APPLICATION; 7= AREA OF LEAST SIGNIFICANCE FOR YOUR TYPE OF APPLICATION)

Figure 3.9 User H

USER I -- IDMS (See Figure 3.10)

A moderate-sized airline company, with approximately 40 entry sites and over 60,000 records on file. Query response times of less than 3 seconds were considered a mandatory requirement, and both exterior and interior locks for security of the system were considered a requirement.

The system is leased at $2,700 per month, with a 10% annual maintenance contract. Insertions and deletions are conducted at the rate of about six per day, and the system is reorganized annually.

The DBA did not believe that the staff had the skills to modify the system, and also stated that contact with the vendor was not applicable to that application (no specific reason was given, but the supposition is that the DBMS is part of an overall systems package, provided by either the hardware vendor or a third-party vendor).

Standardization was a moderate consideration, but only three users were contacted prior to the final selection of that DBMS. Selection criteria weights were led by staff skill level, followed by security options and then user skill level. Hardware was the least important factor, as the system is resident on a series 148 mainframe.

WHAT IS THE MAXIMUM NUMBER OF USERS WHO COULD ACCESS YOUR DATA BASE AT ANY ONE GIVEN TIME ?

One	2 - 4	5 - 20	✓ 20-100	over 100

WHAT IS THE AVERAGE NUMBER OF RECORDS YOU MAINTAIN ON FILE IN THE DATA BASE ?

< 1000	1000-10,000	✓ 10,000-100,000	100,000-500,000	>½ million

WHAT IS THE MAXIMUM RETRIEVAL TIME YOU CAN ACCEPT AND STILL REMAIN AT FULL OPERATIONAL EFFICIENCY ?

✓ < 3 sec. (query)	3-10 sec.	10 sec-1½ min.	1½ min-30 min.	>30 min.

DO YOU MAINTAIN DATA IN THE DATA BASE THAT IS SECURITY-SENSITIVE TO PERSONNEL OUTSIDE THE FIRM ?

never	almost never	on occasion	frequently	✓ all of the time

DOES ANY DEPARTMENT MAINTAIN DATA WHICH MUST BE KEPT SECURE FROM OTHER PERSONNEL WITH ACCESS TO THE DATA BASE?

never	almost never	on occasion	frequently	✓ all of the time

WHAT WAS YOUR ACTUAL PURCHASE PRICE [ ] Purchased at
OR ANNUAL LEASE PRICE ? [✓] Leased for

less $1000.00	$1000 - $5000	$5000-$15,000	✓ $15,000-$50,000	over $50,000

WHAT IS THE AVERAGE ANNUAL MAINTENANCE COST (DBMS ONLY, BUT INCLUDE IN-HOUSE COSTS IF IN VENDOR CONTRACT) ?

less than $250	$250-$1000	✓ $1000-$5000	$5000-$15,000	over $15,000

WHAT IS THE SKILL LEVEL REQUIRED BY THE REAL "END-USER" IN YOUR SPECIFIC APPLICATION ?

8th grade	✓ HS diploma	college degree	computer experience	degree in computers

WHAT IS THE SKILL LEVEL REQUIRED BY THE PERSONNEL WHO MAINTAIN OR SERVICE THE SYSTEM FOR YOU ?

HS diploma	✓ college degree	computer degree	computer science	advanced degree in computers

HOW FREQUENTLY DO YOU ADD ITEMS TO THE DATA BASE OR DELETE ITEMS FROM THE DATA BASE ?

monthly	weekly	✓ < 10/day	< 500 a day	over 500/ day

HOW FREQUENTLY DO YOU CHANGE THE GENERAL STRUCTURE OF THE DATA BASE ?

never	less than yearly	✓ yearly	twice yearly	quarterly or less

CAN YOUR OWN STAFF MODIFY THE SYSTEM TO ALLOW AN EXPANSION OF THE PRESENT APPLICATION ?

✓ No, even with help	With help	Probably	For limited expansion	For any desired task

WHAT IS THE AVERAGE RESPONSE TIME WHEN YOU REQUEST HELP FROM THE SERVICE REPRESENTATIVE ?

NA

No rep available	return call w/i 2 days	Personal visit from rep: w/i a week	w/i 2 days	within ½ day

HOW IMPORTANT IS STANDARDIZATION TO YOUR APPLICATION ? (GOVERNMENT CONTRACTS; NETWORK CAPATABILITY; Etc.)

no impact	little impact	a considerat-ion	✓ accurate im-pact	major impact

HOW MANY USERS OF THE DBMS YOU CHOSE DID YOUR FIRM CONTACT PRIOR TO SELECTING THAT SYSTEM ?

none	1 - 2	✓ 3 - 5	5 - 10	over 10

PLEASE RANK-ORDER THE SEVEN FACTORS ON THE RIGHT TO SHOW THE PROPORTIONAL WEIGHT ACCORDED EACH AREA AT THE TIME THE DBMS WAS SELECTED. [1= MOST IMPORTANT AREA; 7= LEAST IMPORTANT AREA]

	FACTORS	
7	- - - - COST - - - -	7
6	HARDWARE AVAILABLE	6
5	PEER RECOMMENDATION	5
2	- SECURITY OPTIONS -	2
4	- USER SKILL-LEVEL - REQUIREMENTS	4
1	STAFF SKILL-LEVEL REQUIREMENTS	1
3	- STANDARDIZATION -	3

PLEASE RANK-ORDER THE SAME SEVEN FACTORS, GIVEN THAT YOU ARE FAMILIAR WITH AT LEAST ONE DBMS AND HAVE BEEN HIRED TO ACT AS A CONSULTANT FOR A FIRM WHICH IS INTERESTED IN SELECTING A DBMS. [1= AREA WHICH SHOULD BE GIVEN MOST CONSIDERATION IN YOUR TYPE OF APPLICATION; 7= AREA OF LEAST SIGNIFICANCE FOR YOUR TYPE OF APPLICATION]

Figure 3.10 User I

USER J -- IDMS (See Figure 3.11)

A major stock exchange, with 300 entry sites and over a half million records -- all of which were cleared each day from the data base (clearly one of the more volatile record environments). Response time required is less than 3 seconds; external security is needed all of the time and internal security is needed at least on occasion.

The system was purchased at $98,000 (a little high), with a 10% annual maintenance contract. The end-user and staff skill levels were slightly lower than average, being 8th grade and high school diplomas, respectively.

There were 80,000 to 100,000 insertions and deletions per day (which does not correlate with the half million records, all cleared each day, unless this indicates a change to nearly 100,000 records inserted that day), and the system was reorganized on a less-than-yearly basis.

The staff could perform limited modification to the system, and standardization was of only moderate impact.

Of major interest is the vendor response area: the user demanded one-half day on-site vendor service, and a contract was awarded the vendor who met this demand.

Staff and entry skills required were the second and third most important criteria; hardware and cost were tied for the least important selection criteria.

WHAT IS THE MAXIMUM NUMBER OF USERS WHO COULD ACCESS YOUR DATA BASE AT ANY ONE GIVEN TIME ?

One	2 - 4	5 - 20	20-100	over 100 ✓

WHAT IS THE AVERAGE NUMBER OF RECORDS YOU MAINTAIN ON FILE IN THE DATA BASE ?

< 1000	1000-10,000	10,000-100,000	100,000-500,000	>½ million ✓

WHAT IS THE MAXIMUM RETRIEVAL TIME YOU CAN ACCEPT AND STILL REMAIN AT FULL OPERATIONAL EFFICIENCY ?

< 3 sec. (query) ✓	3-10 sec.	10 sec-1½ min.	1½ min-30 min.	>30 min.

DO YOU MAINTAIN DATA IN THE DATA BASE THAT IS SECURITY-SENSITIVE TO PERSONNEL OUTSIDE THE FIRM ?

never	almost never	on occasion	frequently	all of the time ✓

DOES ANY DEPARTMENT MAINTAIN DATA WHICH MUST BE KEPT SECURE FROM OTHER PERSONNEL WITH ACCESS TO THE DATA BASE?

never	almost never	on occasion ✓	frequently	all of the time

WHAT WAS YOUR ACTUAL PURCHASE PRICE (✓) Purchased at
OR ANNUAL LEASE PRICE ? ( ) Leased for

less $1000.00	$1000 - $5000	$5000 - $15,000	$15,000 - $50,000	over $50,000 ✓

WHAT IS THE AVERAGE ANNUAL MAINTENANCE COST (DBMS ONLY, BUT INCLUDE IN-HOUSE COSTS IF IN VENDOR CONTRACT) ?

less than $250	$250-$1000	$1000-$5000	$5000-$15,000 ✓	over $15,000

WHAT IS THE SKILL LEVEL REQUIRED BY THE REAL "END-USER" IN YOUR SPECIFIC APPLICATION ?

8th grade ✓	HS diploma	college degree	computer experience	degree in computers

WHAT IS THE SKILL LEVEL REQUIRED BY THE PERSONNEL WHO MAINTAIN OR SERVICE THE SYSTEM FOR YOU ?

HS diploma ✓	college degree	computer degree	degree computer science	advanced degree in computers

HOW FREQUENTLY DO YOU ADD ITEMS TO THE DATA BASE OR DELETE ITEMS FROM THE DATA BASE ?

monthly	weekly	< 10/day	< 500 a day	over 500/ day ✓

HOW FREQUENTLY DO YOU CHANGE THE GENERAL STRUCTURE OF THE DATA BASE ?

never	less than yearly ✓	yearly	twice yearly	quarterly or less

CAN YOUR OWN STAFF MODIFY THE SYSTEM TO ALLOW AN EXPANSION OF THE PRESENT APPLICATION ?

No, even with help	With help	Probably	For limited expansion ✓	For any desired task

WHAT IS THE AVERAGE RESPONSE TIME WHEN YOU REQUEST HELP FROM THE SERVICE REPRESENTATIVE ?

No rep available	return call w/i 2 days	Personal visit w/i a week	w/i 2 days	within ½ day ✓

HOW IMPORTANT IS STANDARDIZATION TO YOUR APPLICATION ? (GOVERNMENT CONTRACTS; NETWORK CAPATABILITY; Etc.)

no impact	little impact	a consideration	moderate impact ✓	major impact

HOW MANY USERS OF THE DBMS YOU CHOSE DID YOUR FIRM CONTACT PRIOR TO SELECTING THAT SYSTEM ?

none	1 - 2	3 - 5	5 - 10 ✓	over 10

PLEASE RANK-ORDER THE SEVEN FACTORS ON THE RIGHT TO SHOW THE PROPORTIONAL WEIGHT ACCORDED EACH AREA AT THE TIME THE DBMS WAS SELECTED. (1= MOST IMPORTANT AREA; 7= LEAST IMPORTANT AREA)

FACTORS	
7	COST 7
6	HARDWARE AVAILABLE 6
4	PEER RECOMMENDATION 4
1	SECURITY OPTIONS 1
3	USER SKILL-LEVEL REQUIREMENTS 3
2	STAFF SKILL-LEVEL REQUIREMENTS 2
5	STANDARDIZATION 5

PLEASE RANK-ORDER THE SAME SEVEN FACTORS, GIVEN THAT YOU ARE FAMILIAR WITH AT LEAST ONE DBMS AND HAVE BEEN HIRED TO ACT AS A CONSULTANT FOR A FIRM WHICH IS INTERESTED IN SELECTING A DBMS. (1= AREA WHICH SHOULD BE GIVEN MOST CONSIDERATION IN YOUR TYPE OF APPLICATION; 7= AREA OF LEAST SIGNIFICANCE FOR YOUR TYPE OF APPLICATION)

Figure 3.11 User J

USER K -- IDMS [See Figure 3.12]

A large eastern hospital, with over 100 entry sites and a record storage requirement of between 100 and 500 records [which computes to an occasional ratio of an entry site per record -- most unusual]. Response time was listed as less than three seconds maximum, and both internal and external security locks were full-time necessities.

The user would not discuss costs, stating that this was not the policy of the hospital.

End-users and staff skill levels were estimated to be high school diploma and college degree, respectively. An insertion and deletion rate of more than 500 a day[which also does not correlate well] was given, with a complete reorganization on a semi-annual basis.

Five users were contacted prior to the selection of IDMS: peer recommendation was listed as the major criteria for selection of that DBMS, followed by requirements at the staff level and security options. End-user skill is seen as one of the least important features during a process to select a new DBMS. Flexibility of design, recovery and rebuild capability, and customer support would be the major considerations for any future DBMS selection process.

WHAT IS THE MAXIMUM NUMBER OF USERS WHO COULD ACCESS YOUR DATA BASE AT ANY ONE GIVEN TIME ?

One	2 - 4	5 - 20	20-100	over 100 ✓

WHAT IS THE AVERAGE NUMBER OF RECORDS YOU MAINTAIN ON FILE IN THE DATA BASE ?

< 1000 ✓	1000-10,000	10,000-100,000	100,000-500,000	>½ million

WHAT IS THE MAXIMUM RETRIEVAL TIME YOU CAN ACCEPT AND STILL REMAIN AT FULL OPERATIONAL EFFICIENCY ?

< 3 sec. (query) ✓	3-10 sec.	10 sec-1½ min.	1½ min-30 min.	>30 min.

DO YOU MAINTAIN DATA IN THE DATA BASE THAT IS SECURITY-SENSITIVE TO PERSONNEL OUTSIDE THE FIRM ?

never	almost never	on occasion	frequently	all of the time ✓

DOES ANY DEPARTMENT MAINTAIN DATA WHICH MUST BE KEPT SECURE FROM OTHER PERSONNEL WITH ACCESS TO THE DATA BASE?

never	almost never	on occasion	frequently	all of the time ✓

WHAT WAS YOUR ACTUAL PURCHASE PRICE ( ) Purchased at
OR ANNUAL LEASE PRICE ? ( ) Leased For

less $1000.00	$1000 - $5000	$5000-$15,000	$15,000-$50,000	over $50,000

WHAT IS THE AVERAGE ANNUAL MAINTENANCE COST (DBMS ONLY, BUT INCLUDE IN-HOUSE COSTS IF IN VENDOR CONTRACT) ?

less than $250	$250-$1000	$1000-$5000	$5000-$15,000	over $15,000

WHAT IS THE SKILL LEVEL REQUIRED BY THE REAL "END-USER" IN YOUR SPECIFIC APPLICATION ?

8th grade	HS diploma ✓	college degree	computer experience	degree in computers

WHAT IS THE SKILL LEVEL REQUIRED BY THE PERSONNEL WHO MAINTAIN OR SERVICE THE SYSTEM FOR YOU ?

HS diploma	college degree ✓	computer degree	computer science degree	advanced degree in computers

HOW FREQUENTLY DO YOU ADD ITEMS TO THE DATA BASE OR DELETE ITEMS FROM THE DATA BASE ?

monthly	weekly	< 10/day	< 100 a day	over 500/ day ✓

HOW FREQUENTLY DO YOU CHANGE THE GENERAL STRUCTURE OF THE DATA BASE ?

never	less than yearly	yearly	twice yearly	quarterly or less ✓

CAN YOUR OWN STAFF MODIFY THE SYSTEM TO ALLOW AN EXPANSION OF THE PRESENT APPLICATION ?

No, even with help	With help	Probably	For limited expansion ✓	For any desired task

WHAT IS THE AVERAGE RESPONSE TIME WHEN YOU REQUEST HELP FROM THE SERVICE REPRESENTATIVE ?

No rep available	return call w/i 2 days	Personal visit from rep: w/i a week	w/i 2 days ✓	within ½ day

HOW IMPORTANT IS STANDARDIZATION TO YOUR APPLICATION ? (GOVERNMENT CONTRACTS; NETWORK CAPATABILITY; Etc.)

no impact ✓	little impact	a consideration	accurate impact	major impact

HOW MANY USERS OF THE DBMS YOU CHOSE DID YOUR FIRM CONTACT PRIOR TO SELECTING THAT SYSTEM ?

none	1 - 2	3 - 5 ✓	5 - 10	over 10

FACTORS

PLEASE RANK-ORDER THE SEVEN FACTORS ON THE RIGHT TO SHOW THE PROPORTIONAL WEIGHT ACCORDED EACH AREA AT THE TIME THE DBMS WAS SELECTED. (1= MOST IMPORTANT AREA; 7= LEAST IMPORTANT AREA)

4	- - - - COST - - - -	4
5	HARDWARE AVAILABLE	5
1	PEER RECOMMENDATION	1
3	- SECURITY OPTIONS -	3
6	- USER SKILL-LEVEL REQUIREMENTS -	6
2	STAFF SKILL-LEVEL REQUIREMENTS	2
7	- STANDARDIZATION -	7

PLEASE RANK-ORDER THE SAME SEVEN FACTORS, GIVEN THAT YOU ARE FAMILIAR WITH AT LEAST ONE DBMS AND HAVE BEEN HIRED TO ACT AS A CONSULTANT FOR A FIRM WHICH IS INTERESTED IN SELECTING A DBMS. (1= AREA WHICH SHOULD BE GIVEN MOST CONSIDERATION IN YOUR TYPE OF APPLICATION; 7= AREA OF LEAST SIGNIFICANCE FOR YOUR TYPE OF APPLICATION)

Figure 3.12 User K

3.4 Analysis of User Trends

A. Introduction

The user questionnaire was sent to a very select group: users of either IDMS, TOTAL or System 2000. The list of user names was solicited from each vendor, so it should be assumed that the vendor chose only those users who were satisfied with the product. The entire population of the user study is logically expected to be biased toward the product which they are presently using.

Davis noted that many of the established vendors of DBMS's tended to take very good care of the initial customers, but, as the popularity of the DBMS grew, the vendor was able to give a proportionately lesser quality of service to new customers purchasing the product [Dav, 75]. Since the majority of the users are "older" customers, it is likely that they received the best of treatment while attempting to adjust to the concept of a new DBMS.

Finally, most of the users contacted were large firms with a well-established processing center plus a greater amount of available cash with which to shop for a product than would be possessed by the "average" user. The majority of users expressed the feeling that cost was one of the least considered selection criteria; if cost is no object, the user can purchase expensive features which add to the power of the "stripped" DBMS available to the less wealthy average user.

B. User Selection Criteria Ranking

All of the users were asked to rank the selection criteria factors for the time of selection of the DBMS and for a subsequent reselection of a DBMS. The seven factors provided from which the users had to choose were as follows:

1. Peer Recommendation
2. Capability for Standardization
3. Cost
4. Security Features
5. Level of required skill for the End-user
6. Level of required skill for the Staff
7. Hardware Available

The limited number of users who replied to the user questionnaire, plus the biasing factors mentioned above, preclude any attempt at making statistical inferences from the sample. Several factors were evident from the user replies, however, and they are presented here as observations of a few present users of DBMS's in a modern commercial environment.

IDMS is the only CODASYL-type data base management system of the three; it could be expected that a user who was interested in standardization would be very interested in IDMS. In fact, the average reply of the users indicates

that the TOTAL users placed the security criteria higher on the ranking list than did the users of IDMS. One user of System 2000 stated that standardization was the number one criterion in the selection process; further questioning, however, revealed that the user was interested in compatibility with another subordinate firm which had CDC hardware -- not in the CODASYL-concept standardization. No System 2000 owner chose standardization as being anything other than the least important consideration in the selection of a DBMS; this would support the view of System 2000 being one of the less-standardized systems. The crossover between TOTAL and IDMS users is not so clearly explained. (See Figures 3.13 and 3.14)

System 2000 would normally be expected to be chosen by users with a low level of end-user computer experience; the company that needs a "user-friendly" DBMS is more likely to chose the powerful front-end query language of System 2000, which is one of the major selling points of that system. In fact, however, System 2000 was slightly below IDMS in the average ranking of the importance of user skill during the selection process. TOTAL was also considerably lower than IDMS in the number of users reporting user skill as a primary consideration at the time of DBMS purchase.

System 2000 owners did report the lowest concern for cost of the three different owner groups; this could have been expected, since System 2000, with just a few optional

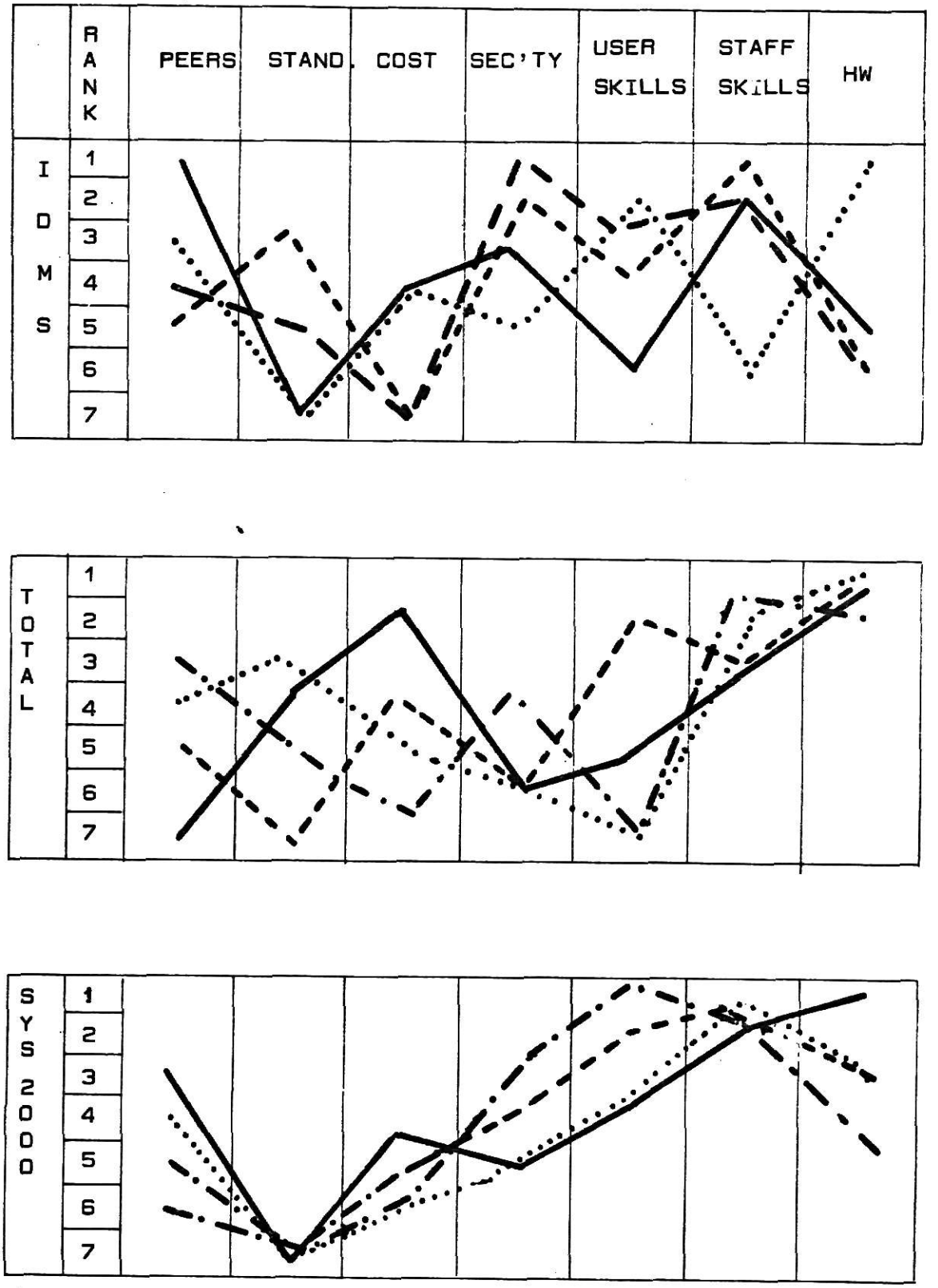

Figure 3.13 User Selection Criteria Ranking

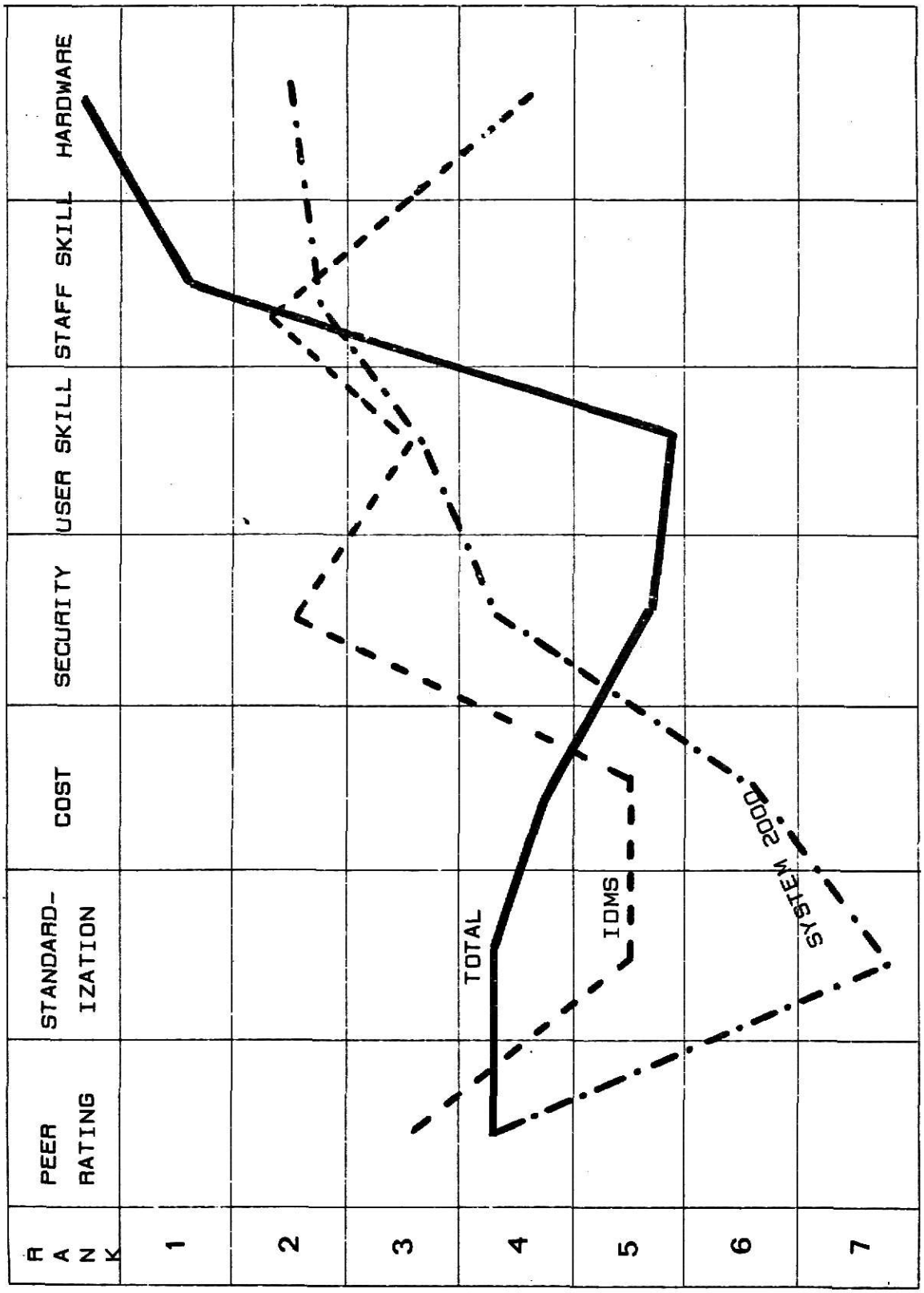

Figure 3.14 System "Average" Selection Criteria Ranking

features, is one of the more expensive DBMS's. TOTAL and System 2000 each exceeded the user concern with hardware of that reported for IDMS, but this was not unexpected since TOTAL operates on one of the most varied sets of hardware, and System 2000 runs on both CDC and Univac systems as well as IBM equipment.

The IDMS group chose security as one of the more important criteria in the selection process; the capability of lower-level structured locks would more easily satisfy the user who desired DBMS security, so this was also an expected response. The slight lead that IDMS shows in user concern with skill level demanded by the system is somewhat puzzling, since many users consider IDMS more complicated than TOTAL.

Only one firm selected peer evaluation of the product as the least important criteria, and they were using TOTAL, which already had the most widely spread and best known system available. A more thorough check of the user reply indicated that there were no peers using the TOTAL DBMS at the time of installation for that user, since it was one of the first installations by CINCOM. This indicates that, for lack of a better instrument to use in the selection process, other-user satisfaction is one of the most used devices for the entire selection process.

CHAPTER 4
CONCLUSIONS

The goal of this research was to design a model that would aid a relatively non-technical manager in the selection of a data base management system for a primarily single-machine, non-distributed data base environment. The proposed model appears to be a workable solution.

The model needs to be tested statistically in the field, and there should be some validation of the suggested weights and scores provided throughout the model. The model does not include many of the more recent developments in the area of data base technology: back-end processors; distributed data bases; DBMS's in the developmental stage; and RAP and other special interest areas not presently available to the general public.

The model, does, however, provide a two-phased method to handle the selection process: the "filter" to eliminate those DBMS's that are not applicable to a specific intended usage; and the "discriminator" to allow a quantification of information about potentially useful systems. This phasing of the model allows the manager to expend a minimal amount of time on the less acceptable DBMS's, while focusing on the characteristics of most concern in the intended area in which the model will be applied. Since there are no other previous models for reference, this model does provide the

point of departure for future work in the expansion of the selection model concept.

BIBLIOGRAPHY

Boylan, David T. "DBMS for Minis," Computer Decisions, January 1976, pp. 66, 69.

Boylan, David. "Minicomputer DBMS: Less than Meets the Eye," Computer Decisions, January 1977, pp. 50, 55.

Burns, Kevin J. "Keys to DBMS Security," Computer Decisions, January 1976, pp. 56, 65.

Cardenas, Alfonso F. Data Base Management Systems. Boston: Allyn and Bacon, Inc., 1979.

Chamberlin, Donald D. "Relational Data-Base Management Systems," Computing Surveys, March 1976, pp. 44, 66.

Chen, P. P. "The Entity-Relationship Model -- Toward a Unified View of Data," ACM Transactions on Database Systems, March 1976, pp. 9, 36.

Davis, B. Data Base Management Systems: User Experience in the U.S.A. Manchester, United Kingdom: NCC Publications, 1975.

"DBMS and Data Management Software," Computer Decisions, January 1978, p. 72.

Donovan, J. J. "Database System Approach to Management Decision Support," ACM Transactions on Database Systems, December 1976, pp. 344, 369.

Fry, James P. and Edgar H. Sibley. "Evolution of Data-Base Management Systems," Computing Surveys, March 1976, pp. 7, 42.

Gibbons, Fred M. "The Mini's Impact on Data Base Management Systems," Mini-Micro Systems, November 1976, pp. 26, 30.

Griffiths, P.P. and B.W. Wade. "Authorization Mechanism for a Relational Data Base System," ACM Transactions on Database Systems, September 1976, pp. 242, 255.

Hamilton, Dennis. "DBMS: A Question for Management," Interface Banking Industry, Fall 1978, pp. 6, 10.

Hunter, John J. "Decoding the CODASYL Database," Computer Decisions, January 1977, pp. 43, 49.

"It's Round Two for Software Patents," Datamation, July 1975, pp. 76, 78.

Katzan, Harry Jr. Computer Data Management and Data Base Technology. New York: Van Nostrand Reinhold Company, 1975.

Lochovsky, Frederick H. and Dionysios C. Tsichritzis. "User Performance Considerations in DBMS Selection," Proc. ACM SIGMOD Conf., Aug 1977, pp. 128,134.

Lozinskii, E. L. Databases: Improving Usability and Responsiveness, ed. Ben Shneiderman. New York: Academic Press, 1978.

Martin, James. Computer Data-Base Organization. 2d ed. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1977.

"Mass and Mini Storage," Computer Decisions, March 1977, pp. 48, 51.

Meadow, Charles T. Applied Data Management. New York: John Wiley & Sons, 1976.

Michaels, Ann S., Benjamin Mittman, and C. Robert Carlson. "A Comparison of the Relational and CODASYL Approaches to Data-Base Management," Computing Surveys, March 1976, pp. 126, 151.

Nijssen, G. M. ed. Architecture and Models in Data Base Management Systems. Amsterdam: North-Holland Publishing Company, 1977.

Nijssen, G. M. ed. Modelling in Data Base Management Systems. Amsterdam: North-Holland Publishing Company, 1976.

Prothro, Vivian C. Data Base Primer. New York: Petrocelli/Charter, 1976.

Ross, Ronald G. "Data Base Systems: Design, Implementation and Management, Part I," Computer World, May 1978.

Ross, Ronald G. "Data Base Systems: Design, Implementation and Management, Part II," Computer World, May 1978.

Scheuermann, P. "On the Design and Evaluation of Data Bases," Computer, February 1978, pp. 46, 55.

Schussel, George. The Information Systems Handbook, F. Warren McFarlan and Richard L. Nolan, eds. Homewood, Ill: Dow Jones-Irwin, Inc., 1975.

Schussel, George. "When Not to use a Data Base," Datamation, November 1975, pp. 82, 98.

Sibley, Edgar H. "The Development of Data-Base Technology," Computing Surveys, March 1976, pp. 1, 5.

Slonim, Jacob and others. "Technical Survey of DBMS Industry in 1977." Manhattan, Kansas: Computer Science Department, Kansas State University, 1978. [Mimeographed.]

Stocker, P.M. and P.A. Dearnley. "Self-organizing Data Management Systems," Computer Journal, May 1973, pp. 100, 105.

Sundgren, Bo. Theory of Data Bases. New York: Petrocelli/Charter, 1975.

Tanner, Donald J. "User Ratings of Software Packages," Datamation, December 1975, pp. 138, 154.

Taylor, Robert W. and Randall L. Frank. "CODASYL Data-Base Management Systems," Computing Surveys, March 1976, pp. 67, 103.

Tillinghast, James. "Get Users involved in DBMS Selection," Computer Decisions, January 1978, pp. 66, 71.

Tsichritzis, Dionysios C. and Frederick H. Lochovsky. "Hierarchical Data-Base Management: A Survey," Computing Surveys, March 1976, pp. 106, 123.

Tsichritzis, Dionysios C. and Frederick H. Lochovsky. Data Base Management Systems. New York: Academic Press, 1977.

"What is CODASYL?" Computer Decisions, November 1976, pp. 74, 78.

Wiederhold, Gio. Database Design. New York: McGraw-Hill Book Company, 1977.

Withington, Frederick G. "Beyond 1984: A Technology Forecast," Datamation, January 1975, pp. 54, 73.

Yao, S.B., K.S. Das and T.J. Teorey. "A Dynamic Database Reorganization Algorithm," ACM Transactions on Data-Systems, June 1976, pp. 159, 174.

A DBMS SELECTION MODEL

by

RICHARD WILLIAM FOX

B.A., University of Tampa, 1975

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Computer Science

KANSAS STATE UNIVERSITY
Manhattan, Kanasa
1979

ABSTRACT

This report describes a model for use in the selection of a Data Base Management System (DBMS) within a single-site, non-distributed environment. The model is aimed at the middle and higher-level manager who may not be fully-qualified in the more technical concepts of the field of computer science.

The model consists of two phases and a total of seven sections. Phase I is a rapid evaluation filter which aids in the elimination of those candidate DBMS's which are obviously unsuitable for the user. Phase II is a detailed discriminator which allows the user to tailor the selection criteria to the specific application for which the DBMS will be used.

The model consists of the following parts:

I. The Filter
   A twelve-question rating form to eliminate those DBMS's unsuitable for the user application.

II. The Discriminator
   A. User Requirements
      1. Identification of User Requirements Form
         A twelve-question format to establish the minimum tailoring requirements for the user.
      2. Requirements Consolidation Form
         A formatted evaluation sheet to consolidate user requirement effects on desired system architecture.
      3. Explanation Section
         A narrative section to explain the trade-offs between user requirements and user demand and system architecture.
      4. Vendor Capability Form
         A rating format to aid in the elimination of vendors which do not meet minimum user requirements.
   B. System Features
      1. Features Evaluation Form
         A suggested value scheme by which system features may be weighted by the user.
      2. System Comparison Form
         An evaluation sheet to compare vendor satisfaction of user-weighted feature demands.