7/25/68

DATA BANK PROGRAMS FOR URBAN PLANNING

A THESIS

Presented to

The Faculty of the Graduate Division

by

William Walter Hibbert, III

In Partial Fulfillment

of the Requirements for the Degree

Master of City Planning

Georgia Institute of Technology

June, 1970

DATA BANK PROGRAMS FOR URBAN PLANNING

Approved:

Chairman

Date approved by Chairman: 22 MAy 1970

ACKNOWLEDGMENTS

The author wishes to express his appreciation to all those who assisted in the preparation of this study. Special thanks are extended to Dr. A. J. Catanese and Professor M. G. Little of the Graduate City Planning Program of Georgia Institute of Technology. Their guidance and encouragement were invaluable for this work. Appreciation is also extended to Professor A. P. Jensen of the School of Information and Computer Science for his assistance.

Gratitude is also expressed to The Maryland-National Capital Park and Planning Commission of Silver Spring, Maryland, which granted the author a Professional Improvement Leave of Absence to complete this study.

A special appreciation is extended to my wife, Meg, who was very patient throughout the preparation of this study.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF ILLUSTRATIONS

SUMMARY

Planners need a method to more effectively and efficiently provide information for better decision making. A data bank program is such a method. The data bank movement is a relatively new concept throughout the United States, and only a few areas have operational programs.

The objectives of this study are to present a guide for the establishment and organization of a data bank program for urban planning and to recommend methods of implementing such a program. In developing this guide, a review was made of existing literature pertinent to the methodology of establishing and organizing a data bank program. In addition, interviews with persons experienced in data bank operation were held, and several data processing seminars were attended.

Several conclusions were reached which led to recommendations for establishing and organizing a data bank program for urban planning.

First, the data bank program should have the support of the chief policy makers and the appropriate organizational framework if it is to function properly. Program policies should be handled by committees, and the actual operation should be carried out by an experienced staff. Once the committees and staff are organized, a Data Bank Program Plan should be prepared to provide a systematic approach to the development and implementation of the program.

Next, data items for the data bank should be thoroughly screened.

Only essential data items should be included. As part of the screening process, data acquisition priorities should be established. Once priorities are established, the best data source agencies should be designated. Then data collection and maintenance procedural manuals should be prepared.

Finally, data release policies should be formulated to insure the confidentiality of the data in the data bank, and an educational and promotional campaign should be developed to demonstrate the usefulness of the data bank program.

CHAPTER I

INTRODUCTION

The burgeoning growth of urban areas has created a demand for
a sound, current, and continuous data source. This study is a guide
for organizing, developing, and implementing a data bank program to
satisfy this demand.

Businesses, research firms, and local governments have established
extensive record-keeping systems to make collected data available in
meaningful form. Each agency within a community has its own require-
ments and maintains its records in different ways for its own purposes.
In some communities, for example, the assessor keeps files relating to
real property so that he can prepare property tax bills; the board of
education keeps detailed records of attendance, academic achievement,
and test results for each pupil; and the police, fire, health, welfare,
and public works departments also maintain their exclusive records.

Duplication of effort and the tendency to make one-time data
collection surveys are characteristic of many governmental, business,
and research organizations. A study to determine the productivity of
research specialists showed that 30 to 85 per cent of research personnel
time is wasted in duplicated effort (1). In an attempt to reduce this
waste, several governmental agencies have recently developed data banks
to provide access to data and information for both technicians and policy
decision makers.

Definition of Terms

Data may be defined as the component parts from which information is derived. Data are the raw facts which are usually input items in the data bank; whereas information is the output or knowledge derived from the data. For example, the number of acres within a parcel of land would be a data item, and the average acreage derived from gross acreage and parcel data would be information.

The data bank is a central storage facility or in some cases a group of interconnected storage facilities for essential data which has been gathered from various sources. When needed, this data can be withdrawn in the form desired. A data bank can consist of file folders, punch cards, magnetic tapes, or similar materials.

The Use of the Data Bank in the Planning Process

Planning agencies are users of data and information. In special cases, such as in an attitude survey, planners personally collect original data. In the preparation of most plans, analyses, and other studies, however, the planning agency depends upon data obtained from the various record-keeping organizations. Often these organizations are unable to provide the required information. When accurate information is obtained, planning agencies serve as repositories for such data and information. Since this repository function tends to be very rudimentary and cumbersome, the planning agency usually initiates a data bank program in order to organize the collection, maintenance, and retrieval of data in a useful manner.

Planning agencies, when forced to handle the repository function,

must design data banks to serve the needs of other governmental and private organizations as well as their own. A data bank has the capabilities of (a) providing an analytical tool that will help local government officials evaluate and measure administrative management and service productivity; (b) providing information which businesses, research agencies, and local governments can use in making decisions, implementing plans, and operating programs; and (c) bridging the gap between planning policy and implementation so that planning studies result in action programs.

As the planner approaches the task of developing a data bank, he is often confronted with the difficulty of obtaining the necessary authority for establishing such a program. Insufficient funds and the lack of an overall concept of how to actually establish the program also often limit him. The guidelines presented herein are designed to aid the planner or anyone interested in organizing, developing, and implementing a data bank program.

Study Methodology

A procedure for establishing a data bank program based on a review of pertinent literature and interviews with knowledgeable officials has been established. This procedure is divided into three phases: (a) organization of the program; (b) development of the program; and (c) implementation of the program.

The first phase includes the formation of policy-oriented committees and a core staff to act as the catalyst for establishing the program. The second phase consists of establishing techniques

for developing the data bank program which include identifying

pertinent data items, determining the best sources of data, and

finalizing a list of essential data items for inclusion in the

data bank. The third phase covers the implementation require-

ments of the data bank program. Within this phase, the prior-

ities of data item inputs are established, the organizations res-

ponsible for data acquisition are designated, procedural manuals for

data collection and maintenance are prepared, policies for releasing

data and information to users are adopted, and a promotional campaign

is developed in order to educate the potential user as to the value

and immediacy of a data bank program. In discussing the steps for

establishing a data bank, possible alternatives will be presented,

positive and negative points for consideration will be included, and

a recommended method will be indicated.

CHAPTER II

ORGANIZING A DATA BANK PROGRAM

One of the first and most important steps that must be taken
in initiating a data bank program is the drafting of a policy state-
ment by the chief policy makers. Such a statement should reflect a
willing acceptance and support of the program and indicate the general
direction the program should pursue. The lack of such a policy
statement would result in a narrow concept of development and
subsequent use and make the resulting data bank program a tool
for the technician alone and not for the policy makers.

A policy statement can take the form of an executive order,
legislative resolution, letter of intent, verbal agreement, or some
locally accepted form of cooperation or intent. The preferable method,
particularly for a single jurisdiction, is the executive order or
legislative resolution because such action would definitively indicate
the solid endorsement of the data bank program by the chief policy
makers. This approach may not be simple for a metropolitan oriented
data bank program since an effective program requires the strong
cooperative and continuing leadership support of many governmental
agencies. In this case, an acceptable method may be a letter of
intent agreed to by several agencies which would indicate their
willingness to cooperate. The executive order or legislative reso-
lution executed by each participating organization, however, is still

the preferred form.

Official authorization for organizing a data bank program of any form should include policy statements on the following three subjects:

1. The endorsement of the data bank program by the chief executive or legislative body indicating its intent to exercise strong and positive leadership in the development of the program.

2. The formation of three basic committees--a Policy Committee, Data Identification Committee, and Data User Committee--and the delegation of the responsibility and authority for establishing the data bank program to these committees.

3. The formation of additional committees with specific responsibilities and authority when recommended by any of the above named committees.

Committee Structure

Lines of communication should be established and areas of responsibility should be delineated so that the individual committees, the core staff, the chief executive, and the legislative body are always aware of the authority invested in each group. There are numerous committee organizational structures, but the structure illustrated in Figure 1 is preferable. Within this structure, the committees are primarily advisory in function to both the chief executive or legislative body and core staff. They meet periodically and establish policies, but the core staff is actually the implementing and operational arm of the data bank program. The core staff performs
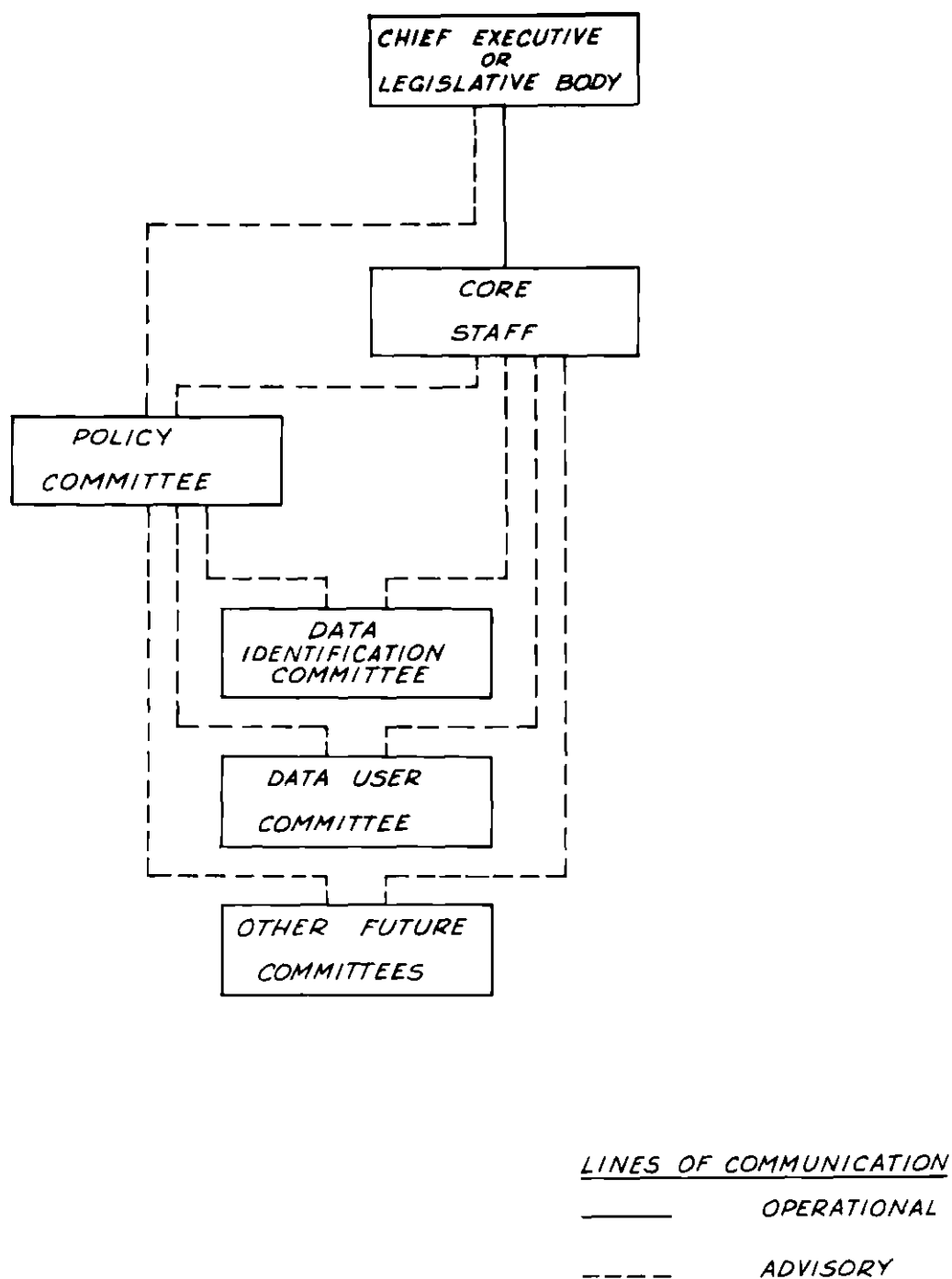
Figure 1.   Preferred Data Bank Program Organization.

the technical work based on guidance from the committees.

Policy Committee

The Policy Committee assumes overall responsibility for
developing a data bank program which would efficiently and econo-
mically meet the data and information needs of the users. It guides
the general work of the core staff operating the data bank program,
and when appropriate it reports to the chief executive or legislative
body. This committee is the focal point for bringing together the
governmental policy decision makers and the leaders of both the
business and research communities for the betterment of the community
as a whole.

As a rule, the Policy Committee is predominantly composed of
governmental leaders because the government is frequently the heaviest
contributor to and user of the data bank. The leadership from the
business and research communities is also significant because these
sectors of the community are both potential major contributors to
and users of the data bank.

The Policy Committee should be composed of elected officials,
appointed governmental officials, prominent businessmen, and research
leaders, who are currently involved in major policy decision making
for the community. The Policy Committee membership might include
members of the local legislative body, city or county managers,
planning commissioners, directors of planning, chief executives of
chambers of commerce, bankers, representatives of research institutions,
or other appropriate officials. The composition of the Policy Committee

is flexible since each community is structured differently.

There is no question of the impropriety of the private organization's having access to confidential governmental data, or vice versa. Being a member of the Policy Committee does not necessarily mean that a member organization would automatically have access to the data any more than would other data bank users. The establishment of data release policies will be discussed in Chapter IV.

The Policy Committee has as its primary function the making of policy decisions in regard to the data bank program. The committee is restricted to policy and planning functions and is not directly involved in technical functions.

In carrying out its task, the Policy Committee performs the following specific functions (see Figure 2):

1. <u>Establish a Set of Goals for the Data Bank Program</u>. The Policy Committee needs to state clearly what it wants and needs from the data bank. A possible goal would be to have the data bank function primarily as a tool for providing better information to community leaders to assist them in making better decisions. Goals and objectives will vary with each individual community, and the committee will have to select the most appropriate and highest priority goals and objectives for the local situation.

2. <u>Initiate a Program for Actively Coordinating all Aspects of the Data Bank Program with Government Officials and the Public</u>. As a prelude to any technical staff work, the Policy Committee initially informs all governmental officals about the proposed data bank program and solicits their cooperation, support, and
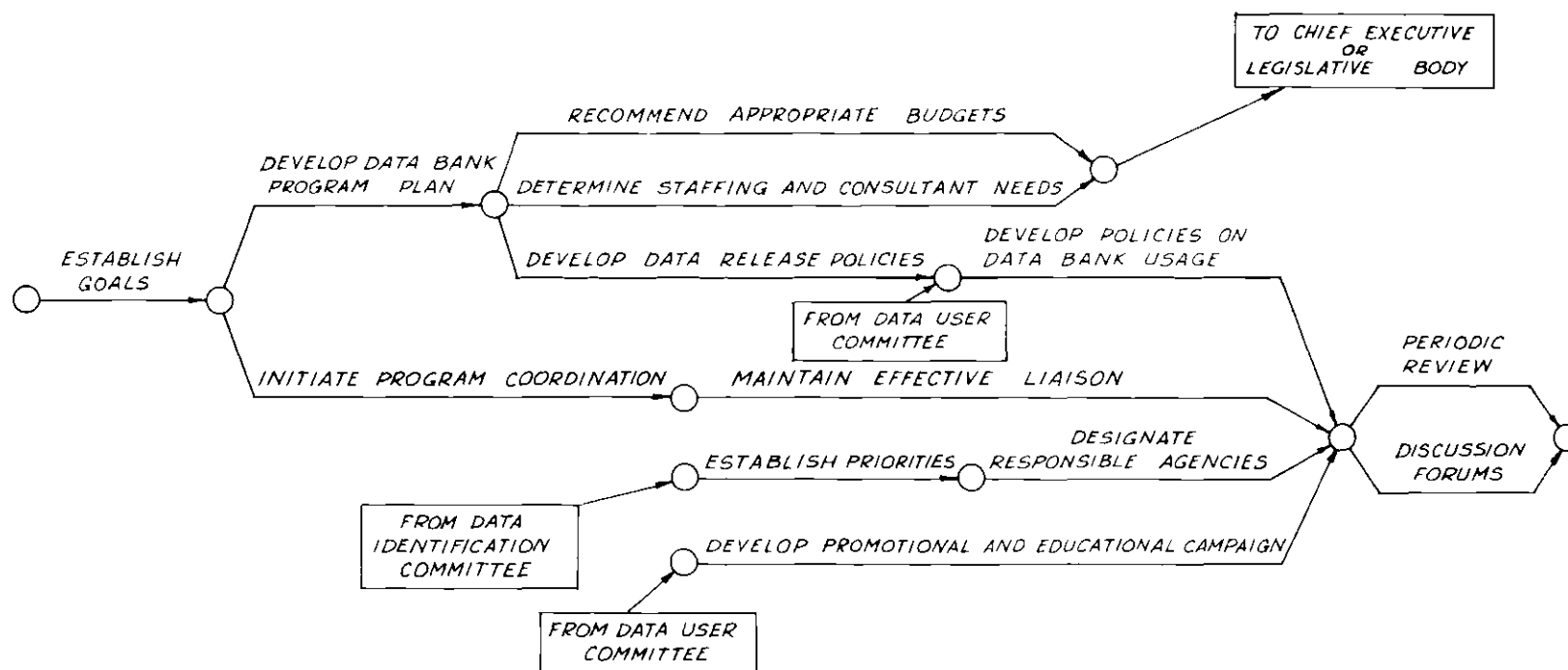
Figure 2. Schematic Flow Diagram of Policy Committee Functions.

participation. Later, a comprehensive program would be initiated with staff support to demonstrate the capabilities of the program and how it can assist various organizations.

3. <u>Develop a Data Bank Program Plan</u>. The Policy Committee prepares an orderly, step-by-step procedure for establishing the data bank program to reflect the needs of the community. This Data Bank Program is continuously reviewed and revised to reflect new requirements of the data bank program.

4. <u>Develop Policies on Disclosure Restrictions</u>. This part of the program involves the evaluation of all laws, regulations, statutes, constitutional guarantees, and existing administrative policies on the disclosure of confidential data. Data release policies are developed to govern the operation of the program.

5. <u>Assign Priorities in Stages for the Implementation of the Data Bank Program</u>. Relying upon recommendations from the Data Identification Committee, the Policy Committee assigns staging priorities of data collection. Thus duplication of effort is minimized, and an adequate time frame within budget limitations is established so that the program can be feasibly implemented.

6. <u>Designate the Responsible Agencies to Implement the Data Bank Program</u>. Following recommendations from the Data Identification Committee, specific responsibilities are assigned to various agencies for initial data acquisition and data maintenance. Their activities would then be coordinated with those of other agencies.

7. <u>Develop a Promotional and Educational Campaign to Promote the Data Bank Program</u>. The Data User Committee develops a promotional

campaign for educating the participating organizations, potential
users, and the general public as to the usefulness of the data bank
program and makes recommendations to the Policy Committee for action.
The Policy Committee adopts the appropriate policies and directs the
core staff to implement the promotional campaign.

8. <u>Develop Policies Governing the Use of the Data Bank</u>. A
procedure for using the data bank is developed by the Data User
Committee for recommendation to the Policy Committee. The Policy
Committee acts on the appropriate policies in order to provide the
core staff with the guidance in handling requests for information
retrieval.

9. <u>Determine the Need for a Data Bank Staff and Location of
the Staff</u>. The Policy Committee evaluates the alternatives for
staff organization and makes recommendations to the chief executive
or the legislative body as to the type, number, and physical location
of the staff personnel needed to operate the program.

10. <u>Determine the Need for Consultants</u>. The Policy Committee
also determines when a planning, systems design, or data management
consultant's service is needed. The committee then screens the
various consultants and recommends to the chief executive or the
legislative body the firm best qualified to perform the required
services.

11. <u>Recommend a Budget for the Data Bank Program</u>. The Policy
Committee develops a budget to operate the data bank program and
recommends it to the appropriate governmental agencies and business

organizations for approval. The allocated funds provide salaries for the core staff, fees for necessary consultant services, and equipment monies available for operating expenses so that the goals and objectives of the program are accomplished.

12. Periodically Review the Operation of the Data Bank Program. After the data bank program is in operation, the Policy Committee, along with the other committees, periodically reviews the program. From time to time, the Policy Committee suggests, guides, or directs either the core staff or one of the other committees to make improvements where necessary to increase the effectiveness and efficiency of the program.

13. Maintain an Effective Liaison with Community Policy Makers. As a continuing part of the comprehensive coordination program, the Policy Committee maintains close liaison with all elected governmental policy makers, business and research leaders, and appointed officials of governmental jurisdictions after the data bank program is operational. This liaison would keep all those involved informed of the activities of the program and would allow community policy makers opportunities to make suggestions for improvements.

14. Serve as a Forum for New Ideas and Techniques in Data Bank Development and Operation. New ideas and techniques are referred to the Policy Committee by anyone interested in improving the data bank program. The committee discusses fully the referrals and selects those innovations which would increase the effectiveness and efficiency of the program.

Data Identification Committee

The Data Identification Committee is created for the specific

purpose of clearly establishing the data requirements of the data

bank program and the most appropriate methods for meeting these needs.

This committee, in addition to working closely with the Data User

Committee, serves in an advisory capacity to both the Policy Committee

and the core staff (see Figure 1).

The membership of the Data Identification Committee should be

comprised of top staff personnel from the major participating govern-

mental agencies, businesses, and research firms. Their primary

qualification would be knowledge of their individual requirements

from the data bank program.

The importance of the work performed by the Data Identification

Committee cannot be over-emphasized. The success of the data bank

program relies heavily on the careful and explicit performance of

this committee. If the Data Identification Committee fails to

function at a high level of technical competence, then the data bank

program will not be strong.

The Data Identification Committee fills the primary technical

function of the data bank program. In fulfilling this function, the

committee develops the technical policies for organizing and implementing

the program (see Figure 3). More specifically, the Data Identification

Committee performs the following duties:

1. Determine the Data Needs of the Users. Through the use of

interviews and conferences, the Data Identification Committee canvasses

IDENTIFY ESSENTIAL
DATA ITEMS

IDENTIFY
DATA SOURCES   RECOMMEND PRIORITIES AND RESPONSIBLE AGENCIES

TO POLICY
COMMITTEE

PERIODIC
REVIEW

FROM DATA USER
COMMITTEE

DEVELOP STANDARDS AND CONTROLS   PREPARE
PROCEDURAL MANUALS

DISCUSSION
FORUMS
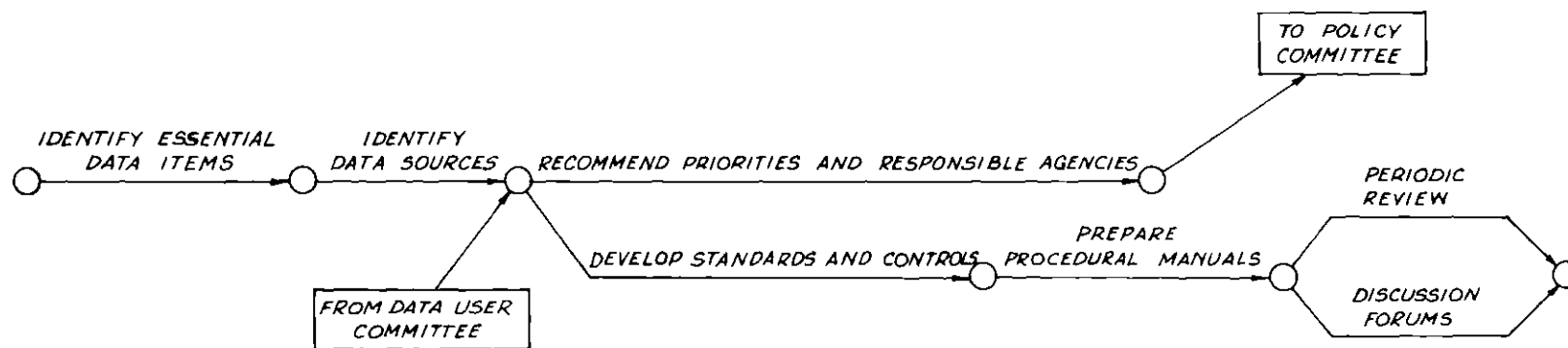
Figure 3.   Schematic Flow Diagram of Data Identification Committee Functions.

potential users and determines their data needs.

2. <u>Determine the Best Sources of Data</u>. As a part of the data requirement evaluation, the Data Identification Committee inventories the various agencies within government and private organizations in order to identify the best data sources. This investigation includes an analysis of the availability, reliability, and comparability of the data items. The result of this analysis is a master list of essential data items.

3. <u>Recommend Priorities for Data Input and Agency Responsibility for Data Acquisition</u>. The Data Identification Committee holds conferences and meetings with the participating organizations to develop a priority list for data collection and designate the appropriate agencies to be responsible for specific data items. This committee then makes a recommendation to the Policy Committee which guides the core staff in implementing the proposal.

4. <u>Develop Appropriate Standards and Controls</u>. The Data Identification Committee holds conferences and meetings and reviews literature on how to organize the storage file. The various methods of data base organization are discussed, the need for standard definitions and classification codes are considered, and the various techniques for editing the incoming data are analyzed. The ultimate objective is to develop standards and controls that insure the accuracy of retrieved information and produce a system for flexible data manipulation and retrieval.

5. <u>Establish Procedures for Data Acquisition</u>. In developing

data acquisition procedures, the Data Identification Committee
investigates the best methods for step-by-step explanations of
initial data collection and maintenance of the data bank on a
current basis. The end result of the process is the preparation
of data acquisition procedural manuals.

6. <u>Periodically Review the Data Requirements and the Main-
enance of the Program</u>. After the data bank program is in operation,
the Data Identification Committee periodically cooperates with the
Data User Committee in evaluating the existing data requirements
and maintenance procedures and in determining necessary improvements.
From time to time, the Data Identification Committee might make
recommendations to either the Policy Committee or the core staff
on certain improvements necessary for increasing the usefulness
and efficiency of the program.

7. <u>Serve as a Forum for Technological Innovations in the
Operation of the Program</u>. New techniques for data identification,
collection, and maintenance would be referred to the Data Identi-
fication Committee which would fully discuss new technology and make
recommendations where appropriate.

<u>Data User Committee</u>

The Data User Committee, which serves in an advisory capacity
to both the Policy Committee and the core staff, works closely with
the Data Identification Committee. Its membership should include top
level staff representatives from each governmental agency and private
organization who intend to use the data bank once it is established.

The primary purpose of the Data User Committee is to optimize

the use of the data bank by meeting the users' needs and encouraging additional usage of the program by governments, research institutions, and business organizations. Focusing on the needs of users for information is essential if a data bank program is to be of any value after it is operational.

The Data User Committee has a two-fold function, education and evaluation. A data bank program may be designed wisely and operating smoothly, but it will exist in a vacuum until the people understand it. A data bank program is not static because changing user needs require alterations in the operation of the program on a continuing basis. A continuing program of education and evaluation should be incorporated into a data bank program from the outset. In order to accomplish this task, the Data User Committee must initiate and carry out the following duties (see Figure 4):

1. Recommend a Promotional and Educational Campaign for the Program. Various public relations techniques are reviewed by the Data User Committee. These techniques form the basis for the development of training and orientation courses for key governmental, business, and research organization personnel on the concepts and potential uses of the data bank. Publicity programs for the news media are developed. The campaign recommended by the Data User Committee is referred to the Policy Committee for approval. The actual campaign is the responsibility of the core staff with appropriate assistance from the members of all the committees.

2. Recommend a Demonstration Project. As part of the public relations review technique, the Data User Committee works with the
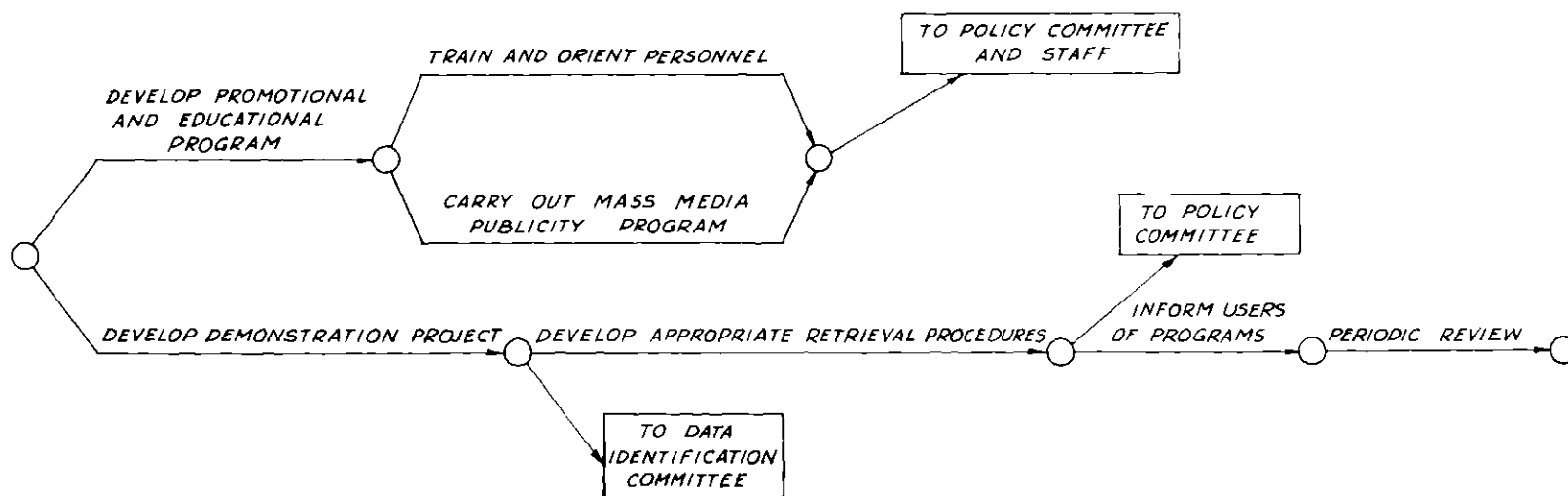
Figure 4.   Schematic Flow Diagram of Data User Committee Functions.

Data Identification Committee in selecting a demonstration project
for promoting the data bank program. To be effective, this demon-
stration project should be accomplished in the early stages of the
program operation.

3. <u>Develop Appropriate Retriveal Procedures</u>. Once the type
of equipment to be used in the program is selected, the Data User
Committee then prepares procedural manuals for data and information
retrieval. These manuals for data bank users explain the detailed
procedures for making a request.

4. <u>Keep Data Bank Users Informed of all Retrieval Programs</u>.
As part of the public relations campaign, the Data User Committee
keeps all potential users of the data bank program informed on the
various methods of retrieving data and information. This service
is performed in expanding the use of the data bank.

5. <u>Periodically Review the Data Requirements of the Program</u>.
Once the data bank program is in operation, the Data User Committee
periodically cooperates with the Data Identification Committee in
evaluating existing data requirements to determine if the program is
meeting the needs of the users. The Data User Committee may recommend
appropriate changes to either the other committees or the core staff.

<u>Staffing Requirements</u>

The efficient operation of data bank program committees depends
in large part on the size of the committee and the size of the supporting
staff. In the formation of these committees, the size of the committees
should be kept relatively small in order for them to function efficiently.

The amount of staff time that can be spent with the committees is another factor affecting the efficiency of the committees' performance.

A staff should be structured to meet the requirements of the data bank program peculiar to an individual community. Personnel may be assigned on a temporary basis from an organization, a new office may be created, or consultants may be contracted. The staff could be either a part of an existing agency or a separate agency unto itself.

Initially, the core staff should be temporary. An existing agency could be given the assignment of serving as an interim central staff during the embryonic stages of the data bank program development. Such an agency could be the office of data processing of the local jurisdiction. This temporary staff has the primary responsibility for developing the committees' agendas, preparing the meeting minutes, and handling limited technical staff assignments. This additional work should not interfere with the normal routine operations of the selected agency.

Long term data bank program projects could be seriously hindered or delayed if the temporary staff remains in a pre-existing agency. The staff would then become involved in daily production schedules, and the data bank program project might be submerged under normal operating priorities. To prevent this, a permanent staff should be hired to operate in a central office or an "information center." The permanent staff is then directly responsible to the chief executive or legislative body but receives guidance and direction from the three data bank committees (see Figure 1).

"In-house" resources are preferred over consultant services whenever possible. Few agencies, however, can afford to maintain a staff that has the overall experience and background necessary for developing a data bank program. Experienced consultants acting as supplementary staff often represent one of the most practical sources of assistance in designing a data bank program that will best meet the community needs at the lowest practical costs.

CHAPTER III

DEVELOPING A DATA BANK PROGRAM

After the organizational framework of the data bank program
is completed, the next step is to begin developing the program. The
first assignment in this phase is the preparation of a Data Bank
Program Plan to establish a time table. Once this plan is developed,
then the actual technical assignments of identifying data items
and locating best sources of data are made so that a master list of
essential items to be included in the data bank can be produced.
The phases of this program will be discussed in this chapter.

Data Bank Program Plan

The preparation of the Data Bank Program Plan is the initial
responsibility of the Policy Committee which is assisted by the Data
Identification Committee, the Data User Committee and any available
staff. Embodied in such a plan is the concept of systematically
solving the problems inherent in the process of establishing a data
bank program. These problems include identifying essential data and
information, keeping data current and accurate, maintaining confi-
dentiality of data, filing and making reliable data readily available
to its users, retrieving information quickly and efficiently, and
promoting the use of the data bank program.

The Data Bank Program Plan provides a time-frame and methodology
by which the committees and staff can meet the information needs of the

community. To do this, the plan provides a coordinated method of accomplishing the following objectives: (a) establishing a set of goals for the program; (b) indicating the data requirements and their potential sources; (c) developing standards and criteria for determining the reliability of data; (d) providing a systematic approach in data collection and maintenance; (e) developing imple-mentation stages and assigning priorities for data input; (f) de-signating the organizations initially responsible for collecting the data and maintaining the program in the future; (g) establishing policies on disclosure of data to users; and (h) determining a method for processing the data.

No single data bank program is universally applicable because the information requirements of each community differ. There are, however, basic guidelines which can be applied to the foregoing ob-jectives (b) through (g). Objective (a), establishing goals, depends entirely upon the local situation, and no general guidelines are applicable. Objective (h), determining a method of processing the data, is dependent upon successive technological advances so it can be determined only at a given point in time.

The remaining sections of this study deal with the general guidelines relating to objectives (b) through (g). Since the develop-ment of procedural manuals and the promotion of data bank usage are important elements in implementing the data bank program, they, too, will be discussed.

The Data Bank Program Plan can never be considered "completed"

in the static sense of the word. The plan should be periodically

reviewed, changed and enlarged to satisfy new requirements. Existing

methods of data handling should be revised in light of new advances

in technology so that the specific needs of present and new users

are met.

Data Identification

According to a study prepared at the George Washington University,

the most serious limitations of data banks are data congestion and

over-collection (2). All too frequently, there is an urge to obtain

data just because it is available. This tendency clutters the data

bank with excess data and increases the cost of developing the program.

Efforts should be made to limit data identification to primary essential

items.

This identification of data items to be included in the data

bank is probably the most crucial problem confronting the three

committees. In order for the Data Identification Committee to accom-

plish its responsibility of identifying data items, two major func-

tions are performed. First, the committee conducts interviews with

those public and private organizations that are potential contributors

to and users of the data bank program. Second, the committee analyzes

the results of the interviews and makes a preliminary identification

of essential data items based on established criteria (see Figure 5).

Preliminary Interview

An interview is defined as "a purposeful conversation planned

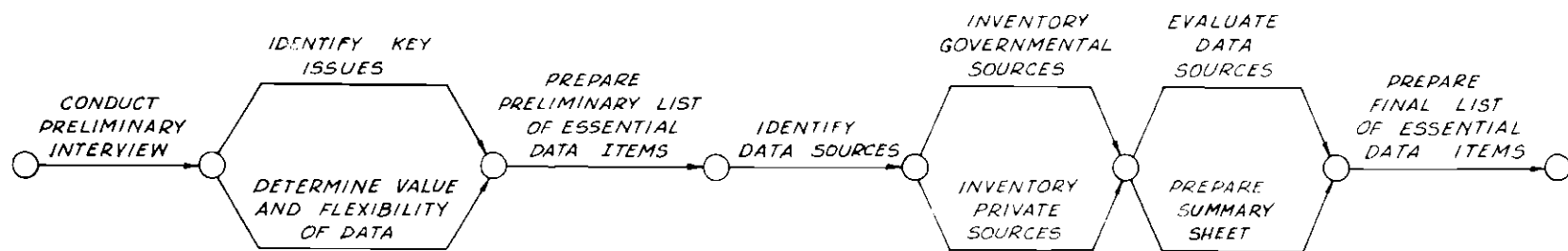and controlled to gain a specific objective" (3). The primary method

Figure 5. Schematic Flow Diagram for Identifying Essential Data Items and Sources.

of data identification is to interview officials responsible for
performing the various operational functions of their organizations.
The purpose of this interview is to determine what essential data
and information they need in their operations and what additional
general data they would like to have available.

Interview Form. One of the primary functions of the data bank
program is providing the information which would be helpful to decision
making or problem solving. It is necessary to determine, in general,
which decisions must be made and which problems must be solved. These
determinations can usually be made by reviewing the results of a well-
designed questionnaire.

If the questions to be asked are not planned before the interview
is conducted, the interviewee is likely to misunderstand the requests
and have difficulty in setting appropriate priorities (4). To avoid
this situation, a well-designed interview should evoke reliable answers
to the following basic questions:

1. To what major critical issues or problem areas in your or-
ganization could accurate and sound data be applied?

2. What specific types of data items does your organization
need?

3. In what specific functional areas will these data items
be used?

4. How often will your required data need to be updated?

5. In what physical format do you desire your data to be
retained?

6. Are there any disclosure restrictions on your data? If so,

what are they?

7. What are the priorities of your problem areas which require data items?

The questions above illustrate the importance of an interview form or set of questions in identifying the specific requirements for each critical issue or problem area. This technique allows for easier manipulation in the analysis and determination of preliminary essential data items.

Interview Techniques. Since the identification of data items is one of the most crucial points in developing a data bank program, the initial contact with the potential participants in the program should be well planned. Techniques for conducting an interview are of utmost importance, and the discussion of such techniques follows.

The least costly technique for measuring user requirements is the mailed questionnaire. The Systems Development Corporation has found, however, that "reliance on questionnaire techniques alone is insufficient for determining user requirements" (5). They feel that actual personal contact between the potential users and the data bank program personnel is needed.

The personal interview is probably the most productive form of securing the information needed by the Data Identification Committee. This is particularly true if the interviewer has acquired the trust and confidence of the interviewee. In order to promote mutual trust and confidence, the skilled interviewer can take certain preliminary steps known to be beneficial in establishing the necessary rapport:

1. Make an appointment with the person who is to be interviewed and indicate the purpose of the interview so that he will have sufficient time in which to think about the subject and to obtain any necessary information prior to the interview.

2. Have on hand a copy of the executive order, legislative resolution, or appropriate agreement requesting information and verifying the interviewer's authority for conducting the interview.

3. Define the objective of the interview and establish a list of questions to be asked prior to conducting the interview.

4. Become familiar with the personal background and work of the interviewee so that rapport can be established easily. It is helpful, of course, if the interviewer already knows the interviewee, but sometimes this is not possible.

5. Try to understand the interviewee's point of view about the proposed data bank program. An interview is conducted differently for a person who is in favor of the program than it is for an individual who is against its adoption. If the individual being interviewed is against the proposed data bank program, find out prior to the interview why he feels this way if possible. It will then be easier to counteract his attitude.

6. Maintain the proper attitude and questioning technique, and never argue with the interviewee. The skilled interviewer will ask thought-provoking questions, but at the same time he will maintain control of the interview. He will take careful notes on answers to questions, record all data during the interview, and watch for additional information to be disclosed outside the framework of the interview.

7. Avoid prolonged interviews by waiting until a later time to prepare an analysis of the data acquired during the interview.

Criteria for Identifying Essential Data Items

After completing the preliminary interviews, the Data Identification Committee should assemble the results obtained from the interviews. From these results several observations may be made. Extensive duplicate collection and storage of the same data item may exist. Data collected by one organization could be used by another organization if that organization knew the data were available and in a useable form. Data items may not be efficiently shared among organizations because of jurisdictional or procedural problems. Often data which would be useful to an organization may not be collected. Another organization could easily collect these data as part of its routine operations if the need were known (6).

The next step is that of analyzing the results. In assisting the Data Identification Committee in analyzing the results of the interviews, two processes are involved: the identification of key issues and the evaluation of the data.

Identification of Key Issues. Information is needed to detect and define problems such as traffic congestion, urban blight, crime, or pollution. Considerable effort should be devoted to identifying major issues having a significant effect on urban development in a specific area. This could be accomplished through a review of the minutes of local planning commissions, county commissions, city councils, and other selected operating agencies within the community. Other sources such as historical newspaper accounts, commission records of requests for

information, and interviews with public officials and agency heads could be consulted. Key issues may also be disclosed during the interviewing sessions. The Data Identification Committee can tabulate the answers obtained from the interviews and develop a list of key issues facing the community.

Since there will normally be a multiplicity of problems and issues to take into consideration, it is usually advisable to determine the priorities or urgencies involved. This procedure will also enable the data bank program to assimilate the most important data first, as it is usually extremely difficult to incorporate all of the required data into the program at its inception. The answers to the question in the interview--"What are the priorities of your problem areas which require data items?"--can be matched with the list of key issues and provide assistance in determining the general priorities.

An example of a key issue could relate to the transportation planning program where certain requirements must be met under the 1962 Federal-Aid Highway Act (7). The Bureau of Public Roads sets forth ten basic planning elements for which inventories and analyses are required in the establishment of a transportation planning process (8). If these requirements are not met, the community's highway funds are withheld.

Evaluation of Data Items. The value of each data item is deter-mined by its applicability in solving problems. Data and information may have value for one person but not for another. Their relevance would depend upon the potential action by an individual in a given environment and time.

Flexibility in the use of data increases its value. Data is worth more to the data bank program if it can be used for more than one purpose. For example, data on building conditions could have a multi-fold purpose. Such data could be used by the building inspector's office, fire prevention bureau, urban renewal department, health department, and planning department. Therefore, when the Data Identification Committee evaluates the question--"In what specific functional areas will these data items be used?"--particular attention should be given to the data items "common" to several data bank participants. Caution should be taken not to limit the data items to such common data but to extend beyond these data to include special data needs of at least one or two major data bank participants. The Metropolitan Data Center Project felt that if the data bank contained only common data, it would not satisfy any one organization's needs with complete adequacy and the data bank would eventually fall into disuse (9).

Preliminary Identification of Essential Data Items

After analyzing the interview results and using the aforementioned criteria for identifying essential data items, the Data Identification Committee prepares a preliminary composite list of the necessary data items. The development of this list should be based upon the determination of which data are required for the decisions which the data bank program is to support and not upon the basis of how many data items can be listed. Since it is assumed that each individual on the Data Identification Committee will likely strive to satisfy his own specific needs, the approach in determining the necessary data items would essentially

be one of negotiation and compromise. Data sources must be found

and evaluated before the final composite list of essential data

items can be made from this preliminary one.

Data Sources

Once the preliminary list of essential data items has been

prepared, then the Data Identification Committee should identify

the best sources for obtaining and maintaining the necessary data

items (see Figure 5). Since data can be obtained from many sources,

both obvious and obscure, a Data Source Inventory should be taken to

determine the most advantageous data sources to be used in the program.

Identification of Potential Data Sources

Before the Data Source Inventory is undertaken, each data

item listed on the preliminary list should be examined, and the

possible sources of each should be identified. Some data may be

assembled from presently existing records such as tax cards, motor

vehicle registrations, aerial photographs, and Bureau of the Census

information. In other cases, it will be necessary to institute field

surveys to provide the data not presently recorded. When data passes

into an office for any purpose, the method involved should be examined

as a possible input source for essential data items.

There are two basic types of data sources--primary and secondary.

Aerial photographs, property and topographic maps, any original written

records of specific observations or events such as traffic accidents or

crimes, field surveys, and interviews are examples of primary sources.

Written reports and studies are most often secondary sources. Most

reports and studies are made in response to a particular policy problem or as a part of a general study of the community. Since their data are usually summarized as part of an analysis, such reports and studies usually do not contain raw data, the prime ingredient of the data bank.

Data sources can be found in both public and private organizations. Both types of organizations keep records of events for the sole purpose of carrying out their defined functions.

Governmental Sources. Data is gathered and assembled at all levels of government--local, state, and federal. Each governmental level usually has specific data not found in any of the other levels. Local government tends to have more data concerning real property while both state and federal agencies have more data about individuals.

At the local level, investigation reports of traffic accidents and fires and inspection reports for property assessment and crime are data sources generated by direct observation of events or objects by governmental personnel. Citizens submit reports such as tax forms and birth, death, and marriage certificates, or they make applications connected with a government service or regulation such as building permits, business licenses, and rezoning requests. Specific items resulting from these sources might include: data on building conditions from inspection reports; data on new construction, alterations, and demolitions from building permits; and data on employment levels, sales, and floor space from business licenses.

State agencies can provide records which are potential sources of data. The employment security agency can furnish data on the number

of employees and the names and addresses of the employers. From the highway patrol, data can be obtained from accident reports and vehicle registration, and the highway department can provide data accumulated from origin and destination survey interviews and traffic counts.

Of all the federal agencies, the Bureau of the Census is the main source of data. Its published data is usually secondary because it is aggregated to a geographic unit. Basic raw data is usually not available. A major problem faced in obtaining raw data from the Bureau is that of disclosure restrictions which will be discussed later in this chapter. Census data includes such items as age, sex, race, housing conditions, number of households and retail businesses, and types of manufacturing establishments. Other potential federal sources might include the Bureau of Labor Statistics, the National Center for Health Statistics, and the Agricultural Reporting Service.

Private Sources. Private organizations are good sources of data that is not part of the operation of governmental agencies. Utility companies, for example, have records on establishment changes based on connections and disconnections. Real estate boards or multi-list services have data on commercial and residential construction and land sale figures. The local chamber of commerce and downtown businessmen's association are sources of data on industrial and commercial expansion. Banks provide gross data on receipts and deposits. When data is being sought from private organizations, two factors, reliability and availability of data, require careful consideration. These two factors will be discussed at length later in this chapter.

Data Source Inventory

When the potential data sources have been identified by the
Data Identification Committee, the next step is to take an inven-
tory to explore more fully the source and nature of the required
data items. Using the same interviewing techniques discussed pre-
viously, the Data Identification Committee or the core staff in-
ventories potential data sources by interviewing prospective data
contributors from each level of government, local associations and
organizations, and private companies.

The actual form for the Data Source Inventory can vary depending
on the individual requirements of the community. The basic format of
the inventory form, however, includes the name of the data item, the
unit of measurement, and the possible sources from which this data
item might be obtained. Each investigated source should list at
least the name of the organization; the definition of the data item
used by that source; the classification into which the data item is
subdivided; the file in which the data item is maintained, such as
real property assessment records or building permit records; the
storage facility used, such as paper document, punch card, or map, the
name of the document, such as rezoning or business license application;
the frequency with which this data item is up-dated in the organization;
and any relevant remarks or comments concerning the stability of the
source or the possibility that the organization might agree to share
in the maintenance of the data item. Only one form for each data item
should be prepared for each organization to guarantee easy manipulation
in the analysis of the data source. Figure 6 is an example of an

Figure 6.   An Example of a Data Source Inventory Form.

inventory form which was used by the Metropolitan Data Center Project
(10).

The knowledge derived from the Data Source Inventory assists
the Data Identification Committee in formulating the actual design,
organizing the data file, and developing procedures for collecting
and maintaining the data. It could also reveal a pattern of data
sharing which would be beneficial in developing the maintenance system.

Data Source Evaluation

When all potential data sources have been inventoried and
investigated, it is necessary to make an evaluation of the probable
usefulness and condition of the data obtained from the various sources.
Data may be presented in a form which is not easily usable for future
analysis. A case in point would be one in which the desired data
item is property acreage, but only property dimensions are given in
the data source. A decision, in this case, should be made at the
outset to convert to acreage measurements or to provide for the use
of dimensions. The willingness on the part of the source agency to
collect the raw data in acres rather than in dimensions should also
be considered.

As part of the evaluation of the Data Source Inventory, a
summary should be prepared for each key issue previously identified
which briefly lists the required data in the form desired, the
source of the data, and any major problems and suggestions related
to the potential data sources. In order to prepare each summary,
the Data Identification Committee should evaluate the results of the
Data Source Inventory in terms of (a) the availability or disclosure

of data; (b) the reliability of data and its sources; (c) the comparability of data; and (d) the updating of data.

Availability of Data. There are two major considerations in determining the actual availability of data items for the data bank program: physical access for initial collection and confidentiality of the data. In a few cases, however, these two problems may not arise. If the occurrence rate of the data item is very low, the data source media could possibly be eliminated. For example, if an address appears less than fifty per cent of the time on a particular source document, then that source media may not be useful to the program.

Physical access to the data item may be difficult when the source document is used constantly in an organization's daily operation. The data may not be free long enough to permit initial collection. A possible solution to this problem is the use of microfilm whereby all the source documents are captured on film in the operating office in a very short period of time and later transferred to the appropriate collection media.

Confidentiality problems usually occur when a proposed data bank program touches on personal data about individuals (11). Confidentiality of data is usually not an issue in a real property data bank because most of the records such as tax assessor files are open to the public. The creation of a person data bank, however, gets at the very heart of the confidentiality issue which, according to Albert Mindlin, Chief Statistician for the District of Columbia, is:

the use of personal information, obtained either voluntarily

> or compulsorily for a particular program at a particular
> time in a particular personal or social context--for a
> different program at a different time in a different
> personal or social context, usually without the knowledge
> by the person that the information is being so used (12).

Non-disclosure laws may also be a barrier to the availability of a desired data item. The Bureau of the Census has a disclosure rule wherein certain types of data are suppressed and not reported if they are available for less than a specified number of cases (13). This rule prevents the possibility of allocating known quantities from which other figures may be deduced specifically. For this reason, some data in the Census of Housing is not available for certain city blocks even though those blocks may be residential in character.

There are also secrecy laws at the state level of government. Historically, the creation of tax laws compelling an individual to disclose very personal data was in turn accompanied by the government's solemn pledge that such data would be treated in absolute secrecy. In the State of Georgia, the only persons outside the State Revenue Department who may properly have access to tax data are only those specifically named in the state statutes (14). Even departmental employees have access to only those data which are necessary in carrying out their official duties. The secrecy provisions in the state law, however, do not prevent the publication of statistics if they are arranged so that they do not reveal information on an individual taxpayer.

Most private organizations will not release data concerning their employees, customers, or operation. Some private organizations

are bound by regulations which do not allow disclosure of data.
Commercial banks are not free to release information on the state
of a depositor's account. Individual pieces of data held as confi-
dential, however, are sometimes published as summary reports and
distributed by private organizations.

Different organizations maintain separate data on individuals,
be it on matters of health, welfare, education, or employment. When
a proposal for inter-departmental sharing of such data is considered,
however, problems of "invasion of privacy" arise. Even though the
sharing of confidential records between separate governmental agencies
has been in existence for decades, one organization may be reluctant
to have its records made accessible to another because of laws and
regulations or conflicts of interest.

The most direct solution to the confidentiality problem of
obtaining access to essential data items for the data bank is to
inform the source organization of the regulations and policies
established by the Policy Committee on user access to data. These
policies are explained in the following chapter.

Reliability of Data. In order to evaluate the results of the
Data Source Inventory, the reliability of the original data source must
be known. One of the most important points to be investigated is the
criteria by which the organizations gather data. Many organizations
have legal sanction to collect data, but other organizations can only
request data. As a result, quality control is diverse, and varying
degrees of adequacy are possible. The permanency of the criteria for
collecting data and the extent and kinds of deviations are two crucial

points to be examined.

In collecting data for the Census of Population and Housing, the Bureau of the Census sometimes changes the definition of key terms so that statistics are more meaningful in the light of current developments. Thus Census statistics are often not strictly comparable from one census to the other and to treat them as being so would lead to error. The Bureau, in publishing material, however, includes a comprehensive report on how the definitions and criteria are to be interpreted and how they have changed since the previous census (15).

In the case of local governmental agencies and private organizations, a summarization of the standards and criteria used is often not available in written or in any other form. This situation would necessitate interviewing the organization's personnel about training and field methods and inspecting their instruction books in order to help determine the reliability of data from the standpoint of stability of criteria used.

When the quality of data differs from source to source, the most reliable data should naturally be used. Reliable and unreliable data should not be mixed because the resulting information would have to be considered unreliable to the degree of the least reliable data input. The fact that a governmental agency has collected the data does not guarantee that the reliability of that data is high. In some cases, private organizations may have more reliable data than governmental agencies because they require high quality data if they are to sell their services.

Comparability of Data. A data bank program depends heavily

upon data input from many organizations. In reviewing the potential

data sources, however, data comparability is an important factor

because various public and private organizations usually apply

different classification schemes to similar or identical data items.

The comparability problem often arises when there is a lack of

standardization in digital coding, work labeling, and defining data

items (16).

Digital codes and work labels are used to represent a defini-

tion of a data item. The problem arises when two or more organiza-

tions use different codes or words to represent an identically or

similarly defined data category. This disadvantage can be overcome

by converting items containing the codes to a compatible form. The

degree of comparability of codes or word labels, however, is of

relatively little importance because only a simple mechanical pro-

cedure is required in converting one set of codes or labels into

another if the definitions are the same (17).

Each individual organization views specific data from a

specialized angle and collects similar data with different techniques

to serve its own functional objectives. In the Metropolitan Data

Center Project, for example, it was found that the data item "Building

Condition" was defined differently by each of six governmental

departments. Data on "Building Conditions" could not be used inter-

changeably by the six different departments because their definitions

were substantially different (18).

When definitions from different sources are not equivalent,

the committee may have to select the best definition or develop a
new one which is mutually agreeable. If the data bank is to have
value for several organizations, data definitions should be clear.
This does not mean that each organization is required to use
precisely the same meaning, but it does imply that equivalences
such as digital codes or word labels could be worked out between
the various organizations to achieve compatibility. The resolution
of this task has a strong influence in the final choice of data
sources for collection and maintenance.

Updating of Data. A factor to be considered by the Data
Identification Committee when determining the best data source is
the desirable frequency of updating the data item. Updating the
data items may create several problems. Many adequate and well
presented bodies of data are not collected at frequent enough
intervals to allow adequate continuous usage. An example of this
would be the federal census which is taken every ten years. Some
bodies of data such as the results of a one-time historical survey
of off-street parking spaces are collected for a special purpose
at one point in time. Another problem is the time lag between the
recording of a data item and its subsequent entry into the data bank
program. The important consideration in this case is the matching
of the data source having the best updating frequency rate with the
data requirements of the various organizations.

Final Essential Data Item List

After the Data Identification Committee has prepared a summary
for each key issue, the committee then should use the summaries to

develop the final master list which identifies the essential data

items to be included in the data bank program. In addition, the

committee should also select the best sources for initially

collecting data and subsequently maintaining the data on a current

basis.

CHAPTER IV

IMPLEMENTING A DATA BANK PROGRAM

Once the essential data item list has been prepared, procedures should be developed for the initial and subsequent periodic collection of data. Prior to initial data collection, policies for releasing data and methods for promoting the use of the data bank program should be formulated. These data bank requirements will be discussed in this chapter.

Data Collection and Maintenance Systems

While it is relatively easy to design initial collection procedures for data, it is a challenge to design those for updating data. A major deficiency in past data collection procedures has been that the structure of the procedures used in collecting data for special study purposes has made using these procedures for future studies difficult. Standardized formats and spellings should be considered in reporting observations or events for file updating purposes as well as for initial collection purposes. The question of who should actually collect the data is also critical to the data bank program. A comprehensive plan for data acquisition and maintenance which covers all these and other salient points are diagramed in Figure 7.

Establishment of Priorities for Data Input

As part of the plan to acquire data, the Data Identification

DETERMINE DATA
BASE ORGANIZATION

ESTABLISH PRIORITIES
FOR DATA INPUT

DESIGNATE
RESPONSIBLE
AGENCIES

DEVELOP STANDARDS
AND CONTROLS

FORMALIZE STANDARD
DEFINITIONS AND CODES

PREPARE
PROCEDURAL MANUALS

TEST
PROCEDURAL
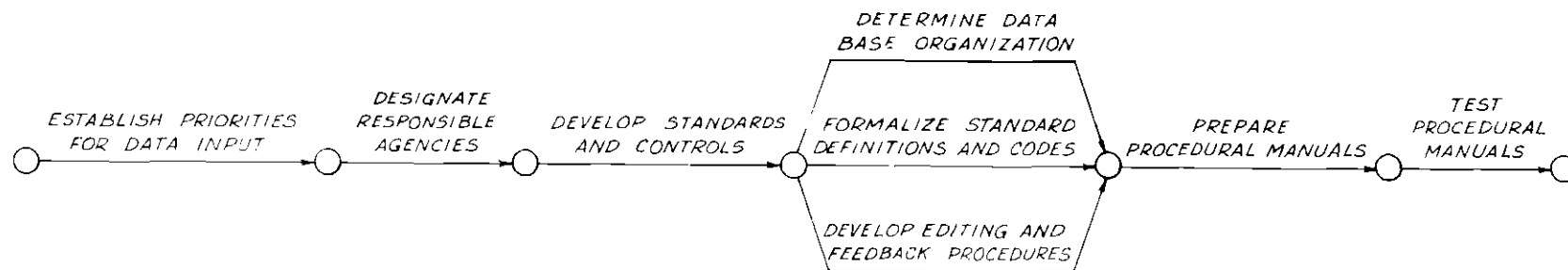MANUALS

DEVELOP EDITING AND
FEEDBACK PROCEDURES

Figure 7.   Schematic Flow Diagram for Data Collection and Maintenance Systems.

Committee should rank in order of importance each essential data

item being considered for the data bank program. These decisions

are based on the results of the preliminary interviews, with

particular reference to the key issues, and on the summaries

prepared as a result of the Data Source Inventory. These results

would give an indication of the urgency, level of accuracy, and the

frequency of collection needed for each data item.

In developing strategies for data input, it is important

that the data bank program be planned in phases rather than in a

single step. Edward Hearle, a data bank expert from the RAND

Corporation, feels the phases should have two purposes:

> First, each phase should be planned as an operating system
> satisfying in itself some real operating requirements. It
> should not simply be a point in the plan that is used to
> measure implementation progress, but should itself satisfy
> operating requirements. Second, each phase should be used
> explicitly to determine the operator's and manager's actual
> use of the information provided. The system can be thus
> modified appropriately based on the findings at each phase.
> Both purposes imply a period of operations at the conclusion
> of each phase before the next phase is undertaken (19).

A phased approach has the significant advantage of showing results

earlier so that converts to the new program are gained as each

interim stage is completed. The design of the program can also be

substantially improved as experience is gained in each new stage

of development.

In establishing priorities for data input, the Data Identifi-

cation Committee should refer the proposed priorities to the parti-

cipating organizations before the final decision is made. This gives

these organizations an opportunity to review and critique the proposed

priorities in relation to their own schedule of activities. If this is not done, the committee may be faced with the same serious problem which occurred in Santa Clara County, California (20). Santa Clara's committee did not refer the priorities for data input to the partici- pating organizations before it published the priority report. Since the Santa Clara Planning Department was listed in the third or last phase scheduled for 1969-71, and their critical data items were ready for entry into the data bank in 1966, the department had to request an earlier priority. The report had already been published, and extensive discussion and negotiation was necessary before an agreement was reached on changing the priority. This incident would not have occurred if the proposed priorities had been referred to the Planning Department prior to final adoption.

Designation of Responsible Agencies

The final master list of essential data items discussed in Chapter III includes the notation of the best source for the initial collection and subsequent updating of each data item. The Data Identification Committee should analyze this tabulation and recommend to the Policy Committee the assignment of specific data collection and maintenance responsibilities for each of the public and private participating organizations. Assigning data acquisition requirements to responsible agencies will minimize duplication of initial data collection, focus responsibility for keeping data continually current, and insure only one source for a specific data item.

In addition to designating the agencies responsible for data acquisition, a written agreement is recommended to insure operation

of the program. For governmental agencies, an executive order or a legislative resolution may be used while a letter of cooperation may be executed between the local government and other participating governmental agencies and private institutions. Little Rock, Arkansas, found that the cooperative effort of various organizations was an important part of the updating system. By using letters of cooperation, they had very little difficulty in gaining the cooperative assistance from other sources outside the local government (21).

Coordination among the various public and private organizations, however, is sometimes difficult to achieve. Complications may arise because there is a general lack of knowledge of what can be accomplished by the data bank program in relation to individual organizational needs and because present work loads of many organizations are too great to provide adequate file maintenance assistance. The first difficulty can be overcome through enlightening the reluctant organization on which products of the program can be of direct assistance to the organization in carrying out its legally assigned function. The second problem can be eased by not superimposing additional transactions upon the organization's regular office routines.

Initial Data Collection Source. Data should be captured at its source and entered directly into the data bank by the source agency. At the time of initial data acquisition, however, the input can be achieved by one of three methods: (a) by personnel of the data source agency using their own operational records, (b) by personnel of another organization using the source material from the source agency, or (c) by a special field survey.

The input of already accumulated records can best be implemented by the personnel of the data source agency. In some cases, however, the source agency will not be able to furnish personnel to extract the desired data from its records because it lacks funds or personnel. In cases of this sort, other organizations would be assigned the responsibility of collecting the data from the source agency. In both the Metropolitan Data Center Project and the Macon Data Bank some of the data source agencies were reluctant to extend themselves to provide the data, and personnel from the planning department had to visit the operating agency and obtain the data (22).

Another tool for securing the initial data for the data bank is the field survey. Such a survey serves two purposes: (a) to collect data which is not a part of the continuing functional operation; and (b) to supplement data when the source agency does not have complete records. The first purpose is illustrated by the home and roadside interviews which collect trip origin and destination data for transportation planning. These data would probably be kept up-to-date through periodic field surveys conducted by the initial collecting agency. A field survey could be used to supplement land use data found in the tax assessor's, and zoning and building inspector's offices. Tax records are usually incomplete in regard to precise land use categories. The land use records of the zoning and building inspector's office are of only recent vintage, and data prior to the adoption of zoning and building codes are not available. The agency responsible for this land use survey would not necessarily be the one responsible for maintenance.

In any of the field surveys mentioned, the responsible agency does not have to use its own personnel. Although planning departments usually conduct land use surveys, the land use survey for Alexandria, Virginia, was carried out by a team of ten firemen working during their off-duty hours. The firemen were chosen because of their knowledge of the city and their familiarity with maps and structures (23).

Data Maintenance Source. In the maintenance of data, the input of updated data should be the responsibility of the source agency. The input can result from any of the three methods described previously in the initial collection stage. In the case of maintaining the data item, however, only the source agency should provide the updated data. If data is shifted through as few hands as possible, data input is more accurate and the time lag between the capture of the data item and its subsequent entry into the data bank is shortened. The collection responsibility should be an integral part of the regular routine and not just an addition to the agency's normal daily operation. If data collection is not a part of the regular operation, the data will not be maintained properly and the usefulness of the data bank program will diminish (24).

Standards and Controls

Accessibility and flexibility are essential in making the data bank program a dynamic system which is capable of fulfilling the needs of its users. The data will be used more often if they are accessible and convenient. The data bank program should also be flexible enough to meet the needs for future data demands without discarding work

already accomplished. In order to initiate and maintain the data

bank program with the necessary accessibility and flexibility, the

Data Identification Committee should establish standards and controls

for data input. These should include the organization of an appro-

priate data base, the development of standard definitions and classi-

fication codes, and the development of editing and feedback controls.

Data Base Organization. The development of a data bank cannot

be accomplished without a sound, established framework upon which to

build. The type of framework or data base needed depends upon the

uses for which the data bank program is being developed, on the

volume and variety of data in the system, and on the storage media.

A data base can be organized with a simple or complex framework. The

two basic forms of a data base are person and geographic.

The basic unit of a person data base is the individual. This

type of data can be found in separate organizations concerned with

welfare, health, law enforcement, and education. The key to a person

data base is a common identifier, such as a social security number.

This type of data base has not been fully developed because of the

difficulties encountered in matching records and in maintaining

confidentiality. While it might prove to be contentious, a person

data base could be most important to a planning function.

A geographic data base can be divided into three types: (a)

parcel; (b) address face; and (c) grid coordinate. The parcel is

the basic legal entity upon which land and improvements on the land

are recorded. Address face is the recording of every street number

of a location whether it be a whole parcel, part of a parcel, or more

than one parcel. The grid coordinate system is the superimposing of two perpendicular sets of numbered parallel lines, equally spaced, on the land for the purpose of identifying points on the land.

The parcel data base is the only type that actually makes up the major portion of existing data banks (25). The basic reasons for this are: (a) most data items originate at the parcel level in building permits, property taxes, and licenses; (b) the parcel is most likely to have a common identification such as the lot and block number recorded in the files of separate agencies; and (c) aggregation is facilitated by using the smallest unit as a base.

The participating agencies of the Metropolitan Data Center Project found that the small size of the parcel made the summarization of data into census blocks and larger geographic units much more flexible (26). If larger geographic control units such as census blocks and tracts are used, these units are less likely to be compatible. In Tulsa, Oklahoma, the coordination of data for planning research purposes was difficult and expensive because the Tulsa Metropolitan Area Planning Commission's block numbering system did not correspond with the U. S. Census block numbering system (27). The Urban Renewal Department for the City of Spokane, Washington, found that the legal plat blocks and boundaries used by the city tax assessor did not correspond with the census blocks and tracts (28). When Macon, Georgia, was confronted with the problem of dividing the city into traffic zones for the Macon Area Transportation Study, the problem was partially resolved by revising the boundaries of the census tracts and blocks to coincide as much as possible with the traffic zones and legal

plat blocks. The multiplicity of overlapping larger geographic areas
such as school districts, traffic zones, utility districts, and police
precincts makes the parcel one of the most attractive geographic
control units for use in the data bank.

The use of the grid coordinate system can further aggregate
and manipulate data for more meaningful purposes than can the data
recorded by parcel unit. Grid coordinates can be used to identify a
point on land, to compute distances between points, and to compute
areal units. The grid coordinate system provides flexibility in data
handling which results in a much greater opportunity for effective
utilization of the data bank program.

In data base organization, the geographic control appears to
be most advantageous when the parcel unit and grid coordinate forms
are combined. This combination would allow for matching of data, more
confidentiality, and more flexibility for retrieving information. The
parcel unit should be the initial data base control to which grid
coordinates are later superimposed. If a grid system is developed at
the outset, incorporating all parcel data would be a major undertaking
which would delay the initial operation of the data bank program much
too long and possibly destroy the program.

Standard Definitions and Codes. In investigating potential data
sources there is a good chance of discovering incompatible definitions
of data items and related classification codes. The discussion in
Chapter III revealed the need for standardization of data item definitions
and classification codes for data collection. As one of the requirements
for developing a data collection and maintenance system, the Data

Identification Committee needs to formalize uniform definitions and classification codes. These are usually developed in combination, but they can be developed separately if the complexity of the storage and retrieval media so demands.

Attempts have been made to standardize certain data items throughout the country. As a result of the Federal-Aid Highway Act of 1962, which requires a comprehensive, cooperative, and continuing transportation planning process, the Bureau of Public Roads has been developing standard classifications for both land use (29) and highway functions (30) to be used in transportation studies.

Editing and Feedback. All data entering the data bank program should be edited to insure reasonable accuracy. Data accuracy should be continuously questioned and constantly tested in reviewing "all appropriate blanks" for completeness and in detecting errors in reporting. All data should be correctly supplied or exceptions explicitly noted. For example, correct parcel numbers must be assigned to each data item in a parcel oriented data bank, and extreme care must be taken to insure that the correct number is attached to each appropriate data item.

The major problem of editing is time. The Metropolitan Data Center Project discovered that the editing of data for even small test areas was a huge job, and in most instances it took far more time than anticipated. The project report also advised that editing procedures be organized prior to data collection, and that computer methods be utilized to the maximum (31).

As a part of the editing process, inconsistencies and errors are

noted and returned to the responsible agency for correcting. The

responsible agency corrects the error and returns the correction

to the editor. A notation of the correction is then made so that

the error will not be made in the future. The ability to feedback

to the original data source is one reason for using a uniform

classification code. Another reason for returning to the original

source is to satisfy a data user's need for additional data not

included in the bank.

Preparation of Procedural Manuals

As a final part of establishing the collection and maintenance

system, procedural manuals should be prepared prior to actual data

acquisition. These manuals enable the data collector to understand

the data bank program by describing the data elements and providing

a step-by-step explanation of all data collection processes, problems,

and solutions. A manual simplifies the data acquisition process and

reduces the number of decisions to be made by the data collector while

he is gathering data. As a result, the person gathering data is guided

in his collection efforts to minimize any time lag and probability of

errors. Identical work tasks are performed in identical ways at all

times. Thus the use of standardized forms, instructions, and controls

is permitted.

Procedures for data acquisition will vary in each community.

The procedure used in a particular data bank program will be dependent

upon the organizational structure of the local public and private

organizations and the type of equipment used. Each community will

have to develop its own data collection and maintenance procedures

to fit its own situation.

Initial Collection Procedures. Stringent requirements are needed in the preparation of a procedural manual for the initial data collection process. These requirements are normally illustrated in detail by means of a descriptive outline of the methodology to be used in the collection process. The outline serves as a guide in completing standardized forms which are required in the pre-survey work, the actual field or office survey, and the post-survey work.

Data collection is essentially a manual process, and recording by hand will probably remain the predominant form of original data collection. The initial data source might, however, use the same or similar automatic equipment as that proposed for the data bank program. In this case, the initial acquired data could be transferred into the bank automatically by gang-punching cards or by duplicating paper or magnetic tape.

One of the most important elements of the procedural manual is the data summary sheets. An example of such a sheet used by the City of Downey, California, is illustrated in Figure 8 (32). The desired data items are transferred from the original source document to the data summary sheet. These sheets can be used for editing for accuracy, for reference and cross-reference purposes, and for transferring data to automatic machines.

The procedural manual is most important because the step-by-step procedures developed for the initial data collection will usually be the guide to future data maintenance. The initial data sources de-signated as data maintenance sources can often use the same procedural

Figure 8.   An Example of a Data Summary Sheet.

manual for both types of data acquisition. In some instances, however, the initial data source is not the data maintenance source. In such cases, the two types of procedural manuals would probably differ.

Updating Procedures. In order to properly maintain the data bank, careful reporting is required. Updating can be a big problem because of the numerous transactions required to keep the data current. The magnitude of updating is illustrated by one community in which approximately 2,000 changes per month in the land use file had to be kept to the same detail as the original survey or inventory (33). Thus the need for a data maintenance manual to keep the program current is crucial.

As discussed in the previous section, the data maintenance manual would be similar to or the same as the procedural manual used for the initial collection process. It would have the same descriptive procedural outline and might include the data summary sheet. The emphasis in this manual, however, would be on basic procedures for entering a new data item, deleting an existing data item, and re-placing an existing data item with a new data item.

Two alternative maintenance processes have to be considered in determining the methodology for updating data. One alternative is to periodically re-survey the entire data field, and the other alternative is to set up a continuous method of adding, deleting, or changing the original recorded data. Both methods are in current use. The Bureau of the Census replaces the entire data file on population and housing every ten years by taking a new census. Most local data banks maintain continuous updating through their daily operations. Any methodology

chosen for the data maintenance system should be devised in collabora-
tion with the data sources in order to maintain the continual flow of
data into the program.

Testing the Procedures. Before the data bank program actually
is in operation, the data collection and maintenance procedures should
be tested for feasibility and completeness. Testing is particularly
important in "de-bugging" the step-by-step procedures to be sure that
the system will function properly. A faulty procedural manual could
affect both the reliability and availability of the data and the cost
of operating the program.

<div align="center">Data Release Policies</div>

Disclosure and potential invasion of privacy can present serious
problems in releasing data. In Chapter III the Data Identification
Committee evaluated the various potential data sources, and the
problem arose concerning confidentiality and sharing of data. Whether
or not an agency's private file should be entered in the data bank
program is a serious issue faced by the Policy Committee. A decision
must be made as to how the data bank files are to be protected from
indiscriminate disclosure.

Problems of Releasing Data

The threat of potential invasion of privacy by a data bank
program has been expressed by those involved in developing data
banks and by other persons and groups concerned with data disclosure.
Both houses of Congress have had subcommittee hearings on invasion
of privacy related to the proposed Federal Statistical Data Center

and its use of the computer (34). Some agencies are reluctant to have

their separate files accessible to other agencies or persons not

presently sharing such data with them.

The apprehension experienced by these agencies leads to

several serious problems. As discussed in Chapter III, the major

contention supporting non-disclosure of data is the invasion of an

individual's privacy. This invasion is associated with person files

such as those kept by health and welfare agencies, school boards,

and police departments. A person data bank records data obtained

voluntarily or compulsorily on an individual, and the capabilities

of the data bank program allow the manipulation of such data to

result in a complete dossier on the individual. In contrast, a

geographic bank usually records data on real property. The problem

of confidentiality is eliminated here because real property is public

information.

Another problem is the potential of accidentally or inten-

tionally misusing data. Robert Colwell of the U. S. Department of

Housing and Urban Development feels that

> the maintenance of public confidence rests more on faith
> in the integrity of those intrusted with data than it
> does on logic. A single indiscretion or an innocent
> error might destroy the confidence built up over many
> years . . .(35)

This is a sensitive area in which moral, legal, and ethical questions

are involved. Colwell further insists that "the structure of data

files be carefully examined to anticipate whether control of the

file by persons who were not benevolent in their motives could be

harmful or capricious." (36). Development of procedures, careful

personnel selection, legal penalties, and usage monitoring are courses

of action that can be taken to minimize the chances of data misuse.

For example, the penalty for divulging information regarding income

taxes in the State of Georgia is a fine not to exceed $1,000,

imprisonment not exceeding one year, or both. If the offender is

an employee of the State, he is dismissed from office and cannot

hold any public office in the State for the subsequent five years (37).

The type of equipment used in the data bank program is also

a factor to be considered in regard to confidentiality. Electronic

computers provide a surer guarantee of confidentiality than do manual

data bank programs. Only a few controls such as locked file cabinets

can be maintained in a manual system, but a computer can be programmed

to give information to only persons with proper credentials such as

storage access code keys and computerized procedural controls.

Pre-release of data and information is another concern of

planners. For years, planners have been apprehensive about releasing

data and information before it has been checked for accuracy. A

recent study indicates that there is a reluctance to release even

non-confidential data in many area transportation studies for this

reason (38).

The overhead costs and procedural methods in retrieving the

information from the data bank also present a data release problem.

Expenses for salaries and overhead should be accounted for, and the

procedures used by the staff personnel to answer a request should be

examined. Those requesting data should be prepared to reimburse the

data bank program for overhead expenses; however, this is often not

required. In the formative years of the Washington, D. C., data bank, all information was initially offered free (39). Procedural methods for releasing data, however, should be as efficient as possible to eliminate any unnecessary personnel time. The problems faced by the Policy Committee in releasing data indicate a serious need for protecting data in the data bank program and for providing a procedure for retrieval.

Development of Data Release Policies

From the discussion in the previous section, it is evident that the Policy Committee should provide for security measures in the data bank program and adopt appropriate procedures and policies for releasing data. Specific regulations should be established to control access to the data bank. The Bureau of the Census operates under strict laws and regulations on disclosure, and it has been quite successful in upholding formal policies and procedures for data release. Data release policies on the local scale, however, have not been developed properly with respect to non-disclosure. The few local agencies that have adopted data release policies have concerned themselves primarily with retrieval procedures and costs of service (40).

In order for the data bank program to be accepted by both data source agencies and information users, data release policies should be discussed by the Policy Committee at the outset and should be formalized concurrently with the developing stages of the program. These policies will assist the data bank promoters in convincing data source agencies that their confidential data will be protected and private citizens that there will be no invasion of privacy. They can also guide the

Data Identification Committee in determining the appropriate data base organization.

It is suggested that data release policies incorporate the following:

1. Each person inquiring for information should record his name and address, and the source agency should be checked to determine security clearance of the individual.

2. An individual should be allowed the right to read his own file and challenge its validity. It should not be necessary, however, for the data bank program to obtain authorization from an individual in order to enter his person data in the data bank.

3. Legislation should be enacted to make unauthorized access to data a civil or criminal offense punishable by law.

4. Purging confidential data should be considered. A five or seven year statute of limitations might be imposed on all data items except for those required for the continuing execution of an agency's legal function.

5. The data release policies should be publicized so that confidence may be maintained in both the public and private sectors of the community.

These ideas are merely suggestions which may be considered in establishing data release policies. The local administrative and political conditions would influence the precise policies needed for releasing data in a community.

Promotion of Data Bank Users

A data bank program is of little value unless the community

truly knows how to use the program once it is established and operating.
A number of communities which have made major efforts in developing a
data bank program have found that it has received only minimum use (41).
Once the program is operational, the potential users often are not
adequately briefed on how the program works and on what its potential
may be for them. The Data User Committee, therefore, should develop
a promotional package that will inform potential users of the advantages
and capabilities of the data bank program.

The task confronting the Data User Committee is that of convincing
the potential data source agencies of the program's value so that they
will reciprocate by furnishing source data. Once this has been accomplished,
other potential users should be recruited and informed of the values of
the program if it is to be a worthwhile project for the entire community.
A special project should be developed and implemented as soon as the
program is operational if the community is to be convinced of its
usefulness. A vital part of the assignment is that of building a good
public image so that the entire community will support the data bank
program (see Figure 9).

Convincing Data Source Agencies

The first step in promoting the usage of the data bank is
directed towards the data source agencies. These agencies should
be told what they will derive from the program if they furnish source
data. If the most desired data are to be obtained for the data bank,
a well organized program to promote the participation of agencies
which offer the best data sources might be necessary.

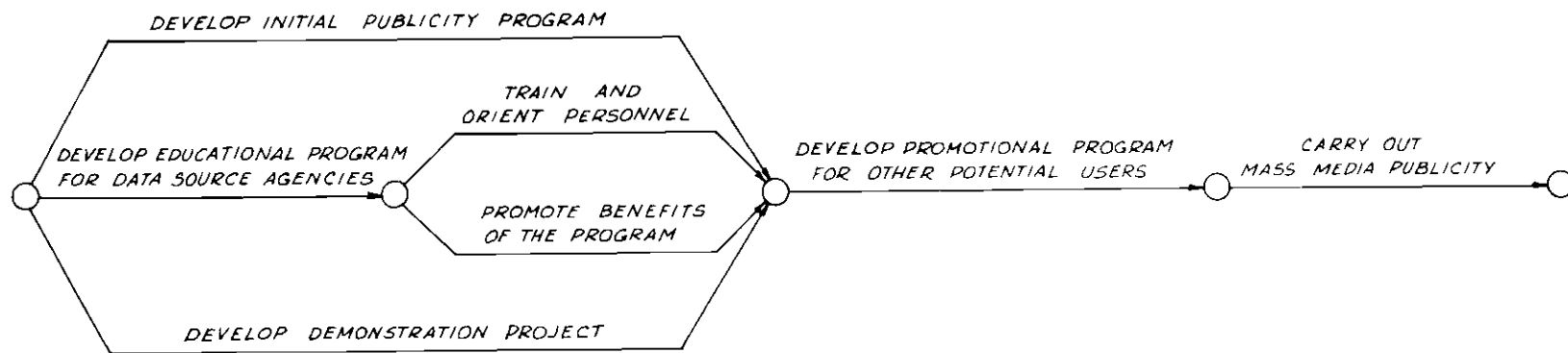Opposition to Change. Attitudes about a data bank program

Figure 9.   Schematic Flow Diagram for Promoting Data Bank Usage.

may range from hostility to enthusiasism. Initially, most government

employees will not be enthusiastic about change. According to Dennis

Price, who has established several programs in the State of New York,

the major problem is the "recumbent incumbent" who is by nature

resistant to change (42). There is also the threat that a part of

his bureaucratic empire may be lost.

In contrast, the City of Alexandria, Virginia, experienced only

enthusiasism for their data bank program. According to John Parker,

former Budget and Research Officer for the City of Alexandria, the

department heads took a positive attitude because

> we didn't threaten to make them do anything, and we
> didn't criticize them for the mistakes they were obviously
> making . . . and we have not taken away any of their
> manual files; they have voluntarily given up several of
> them as soon as they found they could get the information
> in a better way (43).

Promoting the Benefits of the Program. Only by informing the

potential user of the merits of the data bank program can any opposition

be overcome. Promoting such a program is not like selling insurance or

automobiles. Such tangibles have proven their worth over the years and

have built-in selling qualities which can actually be seen before pur-

chasing. A data bank program is an innovation, and the promotion of

such a program is quite a different matter.

When there is opposition to the data bank program, the potential

data source agencies might have to be shown what they can accomplish

with the program. These agencies may not be fully aware of the benefits

they could gain from the processed data. The best method of convincing

a potential data source agency is to demonstrate different ways its

existing data can be used if it becomes part of the program. In the

initial stages of promoting the program, the emphasis should be on

the benefits and not on the detailed techniques because the overall

concept of the program is the most important factor.

When contact is made with a potential source agency, that agency

should be made aware of the following: (a) the objectives of the data

bank program and the methods proposed for accomplishing the objectives;

(b) its specific contribution toward the establishment of the program;

and (c) the ways in which the data bank program might assist the agency's

operations by making data available.

Orientation of Personnel. Once the data source agency is con-

vinced that it would reap benefits from its input to the program, all

personnel involved should be briefed on the program. If possible,

procedural manuals should be distributed. After the personnel have

had time to read the manuals, another briefing session should be held

to answer any questions that might have arisen.

These briefings should emphasize the importance of continuous

updating because the fundamental problem most data banks have is keeping

the bank adequately maintained. The City Manager's Office in Alexandria,

Virginia, reports that improper maintenance of files leads to a decline

in data bank usage (44). Therefore, if the program is to remain useful,

the personnel responsible for updating should be thoroughly indoctrinated

to perform the data collection task with the utmost efficiency.

Convincing Other Potential Users

Demonstrating the values of the data bank program to additional

users is very similar to the techniques used in convincing the data

source agency, but the arguments are different. The data source

agency is providing an input and potentially reaping an output while

other agencies and individuals are only interested in the output

portion of the program. The similarities between their needs are

emphasized, and the flexibility and usefulness of the program are

thus demonstrated. Both types of users are needed in supporting

the program so that the entire community benefits from the funds

expended in operating the system and utilization of the retrieved

information. If many users participate in the data bank program,

more information can be circulated throughout the community to assist

in policy decisions which affect the area.

A method of stimulating the community to use the data bank is

to initially offer the retrieved information free and then gradually

begin charging for paper and related materials. If a computer is

used, charges for computer time and related materials such as punch

cards or reels of tape may be made. A final cost might then be set

which would include all of the foregoing plus staff time exceeding

normal daily operations. The data bank in Washington, D. C., used

this technique. Additional users became increasingly aware of the

value of the program and its potential and began to use the services

of the program (45).

Various agencies and individuals who are potential users of

the data bank include chambers of commerce, real estate firms, lawyers,

private planning consultants, market analysts, banks, utility companies,

radio and television stations, newspaper reporters, and other organiza-

tions which are not data sources. These are only a few examples since

the number of potential users is virtually infinite.

The user's needs are matched only by the imaginative and creative ideas of those who can see beyond the current display of information and visualize how the data bank can satisfy specific needs by various data manipulations. When this imagination is at work and the potential users discover what the data bank program can do for them, requests will probably be made for processed data which are not elements of the program. Such requests should be recorded because the users will then be generating future priorities for expanding the data bank program.

Demonstration Project

One of the best methods of promoting the program is to show concrete results through a demonstration project. Such a project should be considered in determining the staging priorities for implementing the data bank program. The project should be one which can be completed within the shortest period of time after the program is operational.

For example, the demonstration project for Macon, Georgia, was the retrieval of manipulated traffic accident data on a monthly basis to assist the traffic engineer in making decisions on operational traffic flow improvements. This part of the data bank was implemented within three months. While serving the needs of the Traffic Engineering Department, it demonstrated the usefulness of the data bank program in the early stages of development.

The Public Image

A public information program should be developed to keep the

community informed about the data bank program. Initially, publicity can be achieved through a combination of orientation meetings, a few addresses before luncheon groups, local conferences, and situations which provide appropriate opportunities. No widespread publicity through the mass media should be promoted initially "because of the belief that ambitious plans loudly announced in advance often create impossible expectations . . ." (46). In Washington, D. C., the operation of the data bank was not publicly announced through the press until the program had been organized and in operation for a year and a half (47). The Urban Data Center in Cincinnati, Ohio, also deliberately avoided publicity through any local press during the first year of operation. The Cincinnati action resulted in no significant opposition to the program and involved key persons within the community power structure. Political risk and the necessity of public debate were eliminated (48).

Only after the data bank program has been tested for operational adequacy should any mass publicity be released. When this is done, the data bank program should be heralded as a service which will increase the efficiency of the agency by improving the present flow of data. News releases should also justify the expenditure of public funds for the organization and operation of the program and inform the public of the data release policies to allay any misapprehensions about extravagant costs and invasion of privacy. The publicity should not dwell on technicalities but on the broad values of the data bank program.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

The need for better data and information has motivated the

search for new ways to file and manipulate data. Behind this need

lies the compelling urgency for better decision making about man and

his environment. The data bank program promises better immediate

decisions about the environment and community development. It may

also enhance professional research capabilities which may lead to a

better understanding of ecological relationships and to the procreation

of new planning tools and techniques.

The objective of this thesis is to develop a guide for organizing,

developing, and implementing a data bank program and to point out

pitfalls to be avoided in the process. Specific conclusions have

been drawn as a result of this study. Based upon these conclusions, a

set of recommendations is presented which can guide a community in

organizing, developing, and implementing a data bank program. Since

the purpose of this study is primarily to lead the reader through a

process of establishing a data bank without discussing actual data

management and use of equipment, the recommendations set forth reflect

only the step-by-step process of organizational management required in

initiating a data bank program (see Figure 10).

1. In order that a data bank program be successful, the chief

policy makers should be solidly behind the program. The support of the

RECOMMEND BUDGETING, STAFFING AND
CONSULTANT NEEDS

IDENTIFY ESSENTIAL
DATA ITEMS AND
SOURCES

ESTABLISH DATA
ACQUISITION PRIORITIES
AND RESPONSIBILITIES

PERIODIC
REVIEW

OBTAIN SUPPORT
OF CHIEF
POLICY MAKERS

ESTABLISH
ORGANIZATIONAL
FRAMEWORK

DEVELOP
DATA BANK
PROGRAM PLAN

DEVELOP
DEMONSTRATION PROJECT

DEVELOP
STANDARDS
AND CONTROLS

PREPARE
PROCEDURAL
MANUALS

DISCUSSION
FORUMS

DEVELOP DATA RELEASE POLICIES
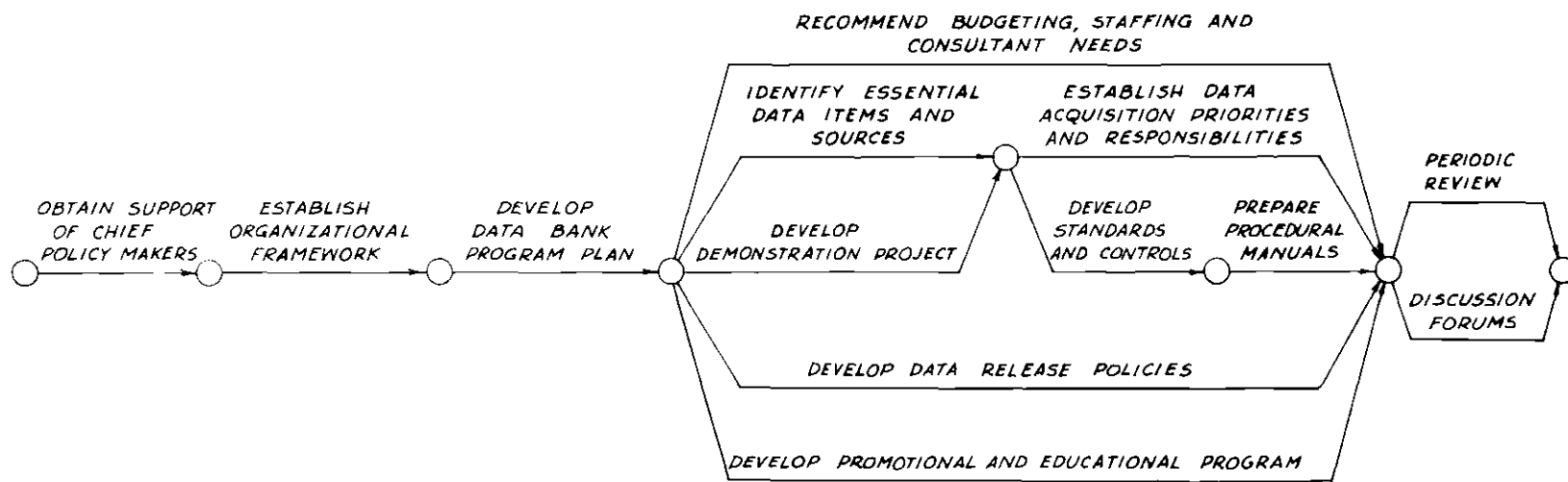
DEVELOP PROMOTIONAL AND EDUCATIONAL PROGRAM

Figure 10. Schematic Flow Diagram on Establishing a Data Bank Program.

chief policy makers should be obtained through an executive order or

a legislative resolution before the data bank program is initiated.

This will indicate the general direction the program should pursue.

Without this support, the data bank program will not be an effective

tool for research, planning, and policy making.

2. There should be an organizational framework which includes

at least three committees--the Policy Committee, Data Identification

Committee, and Data User Committee--and a core staff. The committees

set policy and direct the staff in its task of developing and implementing

the data bank program. The core staff should be a separate department

unto itself and not a part of any other organization. The core staff

is essential to the success of the program because the detailed work

required to accomplish the program objectives is beyond the practical

involvement of the committees.

3. A Data Bank Program Plan should be developed to provide the

committees and the core staff with a systematic methodology for develop-

ing and implementing the program. Since the plan is continuous, it

should be periodically reviewed, changed, and enlarged to meet the new

requirements of technology and user needs.

4. A crucial step in developing a data bank program is the

identification of essential data items. Through carefully selected

interviewing techniques, the Data Identification Committee should

determine the key issues, the data items required to solve these

issues, and the priorities of these issues. Through negotiation and

compromise, the committee can arrive at a preliminary list of essential

data items by identifying the key issues and evaluating the value and

flexibility of the data.

5. Either governmental or private organizations can be excellent sources for obtaining and maintaining data. In order to determine the best data source for a specific item, a Data Source Inventory should be taken. Later the results of the inventory will assist the Data Identification Committee in formulating the actual design and organization of the data files and in developing procedures for collecting and maintaining the data. After such an inventory, the committee can prepare a summary for each previously identified key issue. Such a summary consists of a description of the required data in the form desired, the source of the data, and any major problems and suggestions related to the potential data sources.

The criteria used in preparing a summary include the availability, reliability, comparability, and updating of data. The problems related to availability are providing physical access for initial collection and maintaining confidentiality of data. Reliability questions are focused on the procedures of data capture. Unreliable data should never be mixed with reliable data because the information retrieved would then be unreliable. Standardized definitions and classification codes are the problems in matching data, and the Data Identification Committee should search for the data source with the best frequency rate so that the frequency needs of each organization will be met. Finally, the summaries should be evaluated and a master list of essential data items developed.

6. The actual acquisition of data items is a major process. First, the Data Identification Committee should establish priorities

for data input based on the key issues and on the summaries prepared as a result of the Data Source Inventory. It is important that the process be developed in stages and not implemented all at once. Implementation in stages assists the program designers in improving the system because experience is gained in each new stage of development.

Second, the best data source agencies should be responsible for data collection and should agree in writing as to their specific responsibilities. This designation would minimize duplication of data acquisition by insuring only one initial collection and updating source for a specific data item.

Accessibility and flexibility are vital requirements of the data bank program. These requirements can be met by designing a data base which combines the parcel unit and grid coordinates, by standardizing data item definitions and classification codes, and by providing feedback controls which are particularly helpful in the editing process.

Finally, in order for data acquisition to be successful, manuals which explain explicitly the step-by-step procedures of all collection processes, problems, and solutions should be prepared. The manuals for both the initial collection and updating process may or may not be the same, but they could be if the initial data collection source and data maintenance source for a single data item are the same. Before the manuals are used, they should be tested for reliability.

7. Specific data release policies should be formalized con-currently with the developing stages of the data bank program. If this step is taken, the common fears about invasion of privacy,

pre-releasing of data and information, and undesirable procedures

of retrieving information may be alleviated.

Disclosure of data presents serious problems. The difficulty

is centered around the invasion of privacy concept regarding data

about individuals. Person data problems exist at both the data collec-

tion and data release stages. Solutions to the problems include the

use of a geographic data base rather than a person identifier and the

adoption of formalized data release policies. A geographic data base

has the capabilities of aggregating essential data items without

revealing data on specific individuals. Formalized data release policies

provide a tool for convincing data source organizations that the con-

fidentiality of their data will be retained and for allaying the private

citizen's fear of privacy invasion.

Another problem associated with data release is deciding which

procedures should be used in retrieving the information for data users.

These procedures need to be as efficient as possible to eliminate any

unnecessary expense.

8. Promotional efforts which describe the benefits of the data

bank program should be directed toward those organizations that may be

potential data sources and users. Convincing the potential data source

agencies of data value has the first priority because more accurate data

are sorely needed. These agencies should be shown what benefits they

will derive from the program in return for providing the source data.

The other potential users are convinced by using the same tactics

because their support is needed if the entire community is to benefit

from the utilization of the retrieved information and the funds

expended in operating the program.

A demonstration project can show concrete results and prove the usefulness of the data bank program in a tangible way. Such a project should be completed within the shortest period of time after the program is operational.

In developing the public image, mass media publicity should be delayed until the data bank program is operational and tested. Premature publicity may do the program more harm than good. Orientation meetings, speeches, and conferences are acceptable publicity outlets until the data bank program is ready for widespread operation.

The procedure described in this study is only one of several which can be used in establishing a data bank program. If this procedure is followed, an agency establishing a data bank program should have no difficulty in creating a beneficial planning tool. The decision makers then would have information readily available to guide them in choosing alternate courses of action.

BIBLIOGRAPHY

LITERATURE CITED

(1) Irving Hirsch, William Milwitt and William J. Oakes, "Increasing the Productivity of Scientists," Harvard Business Review, Vol. 36, 1958, p. 69.

(2) Robert D. Campbell and Hugh L. LeBlanc, An Information System for Urban Planning, Washington, D. C.: George Washington University, June 1962, p. 5.

(3) U. S. Army Adjutant General School, The Collection and Charting of Documents, Memorandum 96-3, Fort Benjamin Harrison, Ind.: U. S. Army Adjutant General School, September 1967, p. 3.

(4) Joan E. Jacoby, Seminar at The Maryland-National Capital Park and Planning Commission, Silver Spring, Maryland, March 22, 1969.

(5) U. S. Department of Housing and Urban Development, Urban and Regional Information Systems: Support for Planning in Metropolitan Areas, Part 1, Washington, D. C.: U. S. Government Printing Office, October 1968, p. 28.

(6) Edward F. R. Hearl and Raymond J. Mason, A Data Processing System for State and Local Governments, Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1963, p. 29.

(7) Federal-Aid Highway Act, Sec. 9, 23 U.S.C. sec. 134(1962).

(8) U. S. Department of Commerce, Bureau of Public Roads, "Urban Transportation Planning," Instructional Memorandum 50-2-63(1), Washington, D. C.: Bureau of Public Roads, September 1963, pp. 1-10.

(9) Metropolitan Data Center Project, Tulsa, Okla.: Metropolitan Data Center Project, February 1966, p. 23.

(10) Metropolitan Data Center Project, Interim Report, Tulsa, Okla.: Metropolitan Data Center Project, October 1963, p. 33.

(11) Albert Mindlin, "Confidentiality and Local Data Systems," Papers Presented at the Conference on Small Area Statistics, Census Tract Papers, Series GE-40 No. 4, Washington, D. C.: U. S. Government Printing Office, December 1967, pp. 15-24.

(12) Ibid, p. 21.

(13) U. S. Department of Commerce, Bureau of the Census, U. S. Census of Housing: 1960, Vol. 3, City Blocks, Series HC (3), No. 406, Washington, D. C.: U. S. Government Printing Office, 1961, p. viii.

(14) Acts of 1937-38 and 1945, Code of Georgia, Annotated, Title 92, sec. 92-8414 (1969).

(15) U. S. Department of Commerce, Bureau of the Census, pp. viii-xi.

(16) U. S. Department of Housing and Urban Development, pp. 45-49.

(17) Ibid, p. 48.

(18) Metropolitan Data Center Project, p. 21.

(19) Edward F. R. Hearle, "Alternate Approaches to Data System Design for State and Local Systems," Proceedings of the Conference on EDP System for State and Local Governments, New York, N. Y.: Systems Development Corporation, 1964, p. 113.

(20) Robert A. Clark, "LOGIC: The Santa Clara County Government Information System and Its Relationship to the Planning Department," Threshold of Planning Information Systems, Chicago, Ill.: American Society of Planning Officials, 1967, pp. 73-74.

(21) Frank W. Brutt, The Land Inventory Data Bank and Electronic Data Processing in Planning, Unpublished Masters Thesis, Michigan State University, 1964, pp. 42-43.

(22) Metropolitan Data Center Project, p. 9.

(23) John K. Parker, Budget and Research Officer, Interview at City Hall, Alexandria, Virginia, April 23, 1965.

(24) Jacoby, Seminar.

(25) Mindlin, p. 18.

(26) Metropolitan Data Center Project, p. 19.

(27) Tulsa Metropolitan Area Planning Commission, Land Use: A Procedural Manual for Collection, Storage, and Retrieval of Data, Tulsa, Okla.: Tulsa Metropolitan Area Planning Commission, March 1965, p. 25.

(28) Norma Arlene Olsonoski, A Data Automation Process for Urban Renewal Analysis, Unpublished Masters Thesis, University of Washington, 1962, pp. 71-72.

(29) U. S. Department of Commerce, Bureau of Public Roads; and Housing and Home Finance Agency, Standard Land Use Coding Manual, Washington, D. C.: U. S. Government Printing Office, January 1965.

(30) U. S. Department of Transportation, Bureau of Public Roads, _1968 National Highway Functional Classification Study Manual_, Washington, D. C.: U. S. Government Printing Office, April 1969.

(31) _Metropolitan Data Center Project_, p. 30.

(32) City Planning Department, _Electronic Data Processing Procedures Utilized in the Downey Planning Program Including Mapping and Land Use Techniques_, Downey, Calif.: City Planning Department, April 1965, pp. 24-25.

(33) Brutt, p. 91.

(34) Mindlin, p. 15-16.

(35) Robert C. Colwell, "Compatibility of Urban Land Unit Data," _Papers Presented to the Census Tract Conference_, Washington, D. C.: American Statistical Association, August 1966, pp. 19-23.

(36) _Ibid_, p. 22.

(37) _Acts of 1931_, _Code of Georgia_, _Annotated_, Title 92, sec. 92-9914 (1969).

(38) U. S. Department of Housing and Urban Development, p. 71.

(39) Joan E. Jacoby, "How to Build a Data Bank with Other People's Money," paper delivered at the Conference on The Large-Scale Public EDP System: Its Problems and Prospects, April 1966, p. 5. (mimeo.)

(40) U. S. Department of Housing and Urban Development, p. 78.

(41) U. S. Department of Housing and Urban Development, Part 4, pp. 1-100.

(42) Dennis G. Price, "Automation in State and Local Governments," _Datamation_, Vol. 13, March 1967, pp. 22-25.

(43) Erwin Goldman, "A Report on the Joint University of Georgia/SDC Conference on Application of Electronic Data Processing: State and Local Government," _SDC Magazine_, Vol. 8, May 1965, p. 9, quoting John K. Parker, "General Purpose Information Systems," paper read at the Conference.

(44) U. S. Department of Housing and Urban Development, Part 4, p. 24.

(45) Jacoby, "How to Build a Data Bank," p. 5.

(46) Fred J. Lundberg, "Urban Information Systems and Data Banks: Better Prospects with an Environmental Model," <u>Threshold of Planning Information Systems</u>, p. 67.

(47) Jacoby, "How to Build a Data Bank," p. 5.

(48) Lundberg, p. 66.

OTHER REFERENCES

Brounstein, Sidney H. "Some Concepts and Techniques for Constructing and Using a Geographically Oriented Urban Data Base." Paper read at the Annual Symposium of the Association of Computing Machinery, New York, N. Y., November 1967. (mimeo.)

Candeub, Fleissig, and Associates. Interim Report No. 1: Data Collection Manual. Woodbridge, N. J.: City Department of Planning and Development, December 1966.

Department of City Planning and Urban Renewal. Land Use Survey Manual. Alexandria, Va.: Department of City Planning and Urban Renewal, undated.

Heany, Donald F. Development of Information Systems. New York, N. Y.: Ronald Press Co., 1968.

Hearle, Edward F. R. "Electronic Data Processing in Planning: A Framework of Alternatives." Planning 1965, Chicago, Ill.: American Society of Planning Officials, November 1965, pp. 301-305.

Kornfeld, Leo L. "The Consultant's Contribution." Proceedings of the Conference on EDP Systems for State and Local Governments, New York, N. Y.: Systems Development Corp., 1964, pp. 135-136.

Lamoureaux, Jeanette. Applications of Automatic Data Processing Methods to Urban Planning Studies. Unpublished Masters Thesis, University of North Carolina, 1961.

Martino, R. L. MIS-Methodology. Wayne, Pa.: Management Development Institute, 1969.

McGinty, Richard T. "Description of the Approach of One Planning Commission." Planning Research 1966. Washington, D. C.: American Institute of Planners, 1966, pp. 81-86.

Morstein and Associates. The Conceptual Framework of a DVRPC Information Services Plan. Philadelphia, Pa.: Delaware Valley Regional Planning Commission, January 1968.

Orange-Seminole-Osceola Planning Commission. Metropolitan Data Bank. Orlando, Fla.: Orange-Seminole-Osceola Planning Commission, April 1968.

Parker, John K. "Guidelines for Establishing a Metropolitan Databank." Paper read at the Conference on a Databank for Metropolitan Milwaukee, Wisconsin, November 1965. (mimeo)

Parker, John K. "The Comprehensive Municipal Data System of Alexandria, Virginia." Papers Presented at the Census Tract Conference on September 10, 1965 at Philadelphia, Pa. Washington, D. C.: U. S. Department of Commerce, Bureau of the Census, 1965, pp. 14-17.

Report on the Management Study of Automatic Data Processing. Albert B. Tieburg, Chairman. Sacramento, Calif.: State of California, November 30, 1964.

Von Plinsky, Alexander H., Vice President of Southern Data Processing Service, Augusta, Georgia, private interviews on September 25 and October 3, 1967.

Williams, W. K. "Computers in Urban Planning." Paper read at the Annual Symposium of the Association of Computing Machinery, New York, N. Y., November 1967. (mimeo)

Wise, Fred H., Director, Center for Management Systems Development and Analysis, University of Georgia, Athens, Georgia, personal letter, August 23, 1967.