DOCUMENT RESUME

ED 092 141 IR 000 701

TITLE	Regional Information System: Feasibility Report. Final Report.
INSTITUTION	Lower Pioneer Valley Regional Planning Commission, West Springfield, Mass.
SPONS AGENCY	Department of Housing and Urban Development, Washington, D.C.
PUB DATE	Mar 73
NOTE	48p.
EDRS PRICE	MF-$0.75 HC-$1.85 PLUS POSTAGE
DESCRIPTORS	*Computer Oriented Programs; *Computer Science; Data Processing; Feasibility Studies; *Information Systems; *Regional Planning

ABSTRACT

A study was made of the feasibility of establishing a computerized regional information system. Existing systems were reviewed, and the location of the system was discussed. The system's structure was described. A cost benefit analysis looked into capital costs and operating costs. A cooperative system was found to be feasible, based on utilization of existing computer facilities and support from municipalities and agencies participating in the system. Participating agencies would collect and maintain data and support processing; output would be available to participating agencies. (Author)

BIBLIOGRAPHIC DATA SHEET	1. Report No. LPV-CRIS-1	2.	3. Recipient's Accession No.

4. Title and Subtitle

Regional Information System: Feasibility Report

5. Report Date: March 1973

6.

7. Author(s): Lower Pioneer Valley Regional Planning Commission

8. Performing Organization Rept. No. LPV-CRIS-1

9. Performing Organization Name and Address

Lower Pioneer Valley Regional Planning Commission
1499 Memorial Avenue
West Springfield, Massachusetts 01089

10. Project/Task Work Unit No. OPD-72-10.04

11. Contract/Grant No. CPA MA 01 06 1019(A)

12. Sponsoring Organization Name and Address

Department of Housing and Urban Development
Region I Area Office
Boston, Massachusetts 02114

13. Type of Report & Period Covered: Final

14.

15. Supplementary Notes

16. Abstracts

Study of feasibility of establishing a computerized regional information system. A cooperative system was found to be feasible, based on utilization of existing computer facilities and support from municipalities and agencies participating in the system. Participating agencies would collect and maintain data and support processing; output would be available to participating agencies.

17. Key Words and Document Analysis. 17a. Descriptors

Regional Planning
Information Systems
Computerization

17b. Identifiers/Open-Ended Terms

17c. COSATI Field/Group

18. Availability Statement	19. Security Class (This Report)	21. No. of Pages
Release unlimited. No original copies available.	UNCLASSIFIED	47
	20. Security Class (This Page) UNCLASSIFIED	22. Price $3.00

NTIS-35 (REV. 3-72)

THIS FORM MAY BE REPRODUCED

USCOMM-DC 14952-P72

REGIONAL INFORMATION SYSTEM:

FEASIBILITY REPORT

March 1973

Lower Pioneer Valley Regional Planning Commission
1499 Memorial Avenue
West Springfield, Massachusetts 01089

The preparation of this report was financed in part through a comprehensive planning grant from the Department of Housing and Urban Development.

LOWER PIONEER VALLEY REGIONAL PLANNING COMMISSION

OFFICERS

Theodore S. Bacon, Jr. - Chairman
Norman F. Rauscher - Vice Chairman
Edward A. Koziol - Secretary
Edward C. Peck, Jr. - Treasurer
Howard T. O'Day - Assistant Treasurer

MUNICIPALITY	COMMISSIONER	MUNICIPALITY	COMMISSIONER
Agawam	J. Burns	Ludlow	E. Falkiewicz
Amherst	T. S. Bacon, Jr.	Middlefield	R. S. Bellows
Belchertown	D. S. Cole	Monson	R. A. Auclair
Chester	G. Young	Montgomery	J. G. McKay
Chesterfield	Rev. B. H. Harvey	Northampton	F. P. Ryan
Chicopee	F. Zazyczny	Palmer	R. Marier
Cummington	D. Forgea	Pelham	R. T. Jones
E. Longmeadow	R. W. Clarke	Russell	H. Mason
Easthampton	J. LaFontain	South Hadley	F. Charron
Goshen	F. S. Dresser	Southampton	B. Coombs
Granby	N. C. White, Jr.	Southwick	R. Ellershaw
Granville	H. Chapman	Springfield	M. O. Saunders
Hadley	M. Kostek	Tolland	C. W. Clark
Hampden	C. F. Libby	Ware	A. Mucha
Hatfield	W. H. Burke, III	Westfield	A. J. Paquette
Holland	E. E. Parron	Westhampton	P. Swanson
Holyoke	W. R. Kane	W. Springfield	R. McCarroll
Huntington	A. J. Bartos	Wilbraham	N. F. Rauscher
Longmeadow	M. L. Bowers	Williamsburg	Martin E. Peterson
		Worthington	C. S. Packard

STAFF

K. M. Munnich	Planning Director
D. M. Gossland	Assistant Director
A. M. Berke	Senior Planner
M. H. Khan	Senior Planner
*R. Moore	Senior Planner
G. J. Boyle	Planner
S. Hayes	Planner
B. A. Kwetz	Planner
C. Hankins	Planning Assistant
A. Taft	Administrative Assistant
E. L. Sumowski	Secretary
F. E. Eades	Secretary

*Principally responsible for this report.

PREFACE

This report is a preliminary report. It is a preliminary because it is the first step towards establishing a rationalized regional information system. It is preliminary because it does not set forth in detail the manner of establishing such a system, nor does it present a detailed plan for such a system.

This report explores the feasibility of a computerized regional data bank as a central part of a regional information system. But, more than this, it investigates information requirements for regional planning and how these requirements can be met by existing data sources. It also outlines the overall problem of building a regional system out of a multitude of independent participants and users.

It is hoped that this study can be carried forward, within the limits that funding will allow, toward the establishment of a comprehensive regional information system which will be of substantial benefit to the communities in the region.

TABLE OF CONTENTS

I. BACKGROUND

The Lower Pioneer Valley Regional Planning Commission provides in its Overall Program Design for the study of a regional information system to assist its regional planning efforts. This report is the product of the appropriate sub-category in the LPVRPC's Annual Work Program for 1972-1973. The program was funded by Commission funds and matching funds of a comprehensive planning assistance grant from the U.S. Department of Housing and Urban Development.

OBJECTIVE AND WORK ELEMENTS

As set forth in the 1972-1975 Overall Program Design, "Sub-Category 20.04--Information System" has as its objective "to examine the feasibility of a computerized regional data bank in terms of cost, users, and regional and local benefits."

Work elements for the first year (1972-73) are:

1. Research existing literature on regional computerized information systems.
2. Examine possible locations, applications and initial capital costs for this region.
3. Determine potential users and the type and degree of support.
4. Determine the range and detail of information required and possible in terms of using existing local, state and federal records and developing standardized input and output formats.
5. Estimate staffing and operating costs.

This objective and those work elements set the basic framework for the study. In the conduct of the study some additional elements had to be considered, but it is felt the

basic work called for has been as adequately investigated as possible, within the constraints of time and funding.

STUDY LIMITATIONS

There are limitations to what can be accomplished in a study of this type, least of which are those set by the basic work element framework. There are theoretical limitations inherent in the nature of information systems and the requirements of regional planning. There are practical limitations dictated by the nature of the end product being studied and by the situation in which the study is being conducted.

At the same time, feasibility itself can be viewed as theoretical feasibility or as practical feasibility. Theoretical feasibility deals with the availability of equipment and the ability to utilize this equipment to handle information in the way required. Practical feasibility deals with having the wherewithal in money and the necessary authorization and cooperation required to establish and operate any proposed system.

From a theoretical point of view, the machinery and technical ability to utilize such for the processing of information exist to an adequate degree for the formation of a regional information system to support regional planning. This has been demonstrated by other regional planning organizations establishing information systems.

It can be assumed, then, that it is theoretically possible

to establish our contemplated system. It might even be assumed that it is possible, benefiting from the experiences of others, that we can create a better system than any yet established. The practical problem of doing so remains, and this will be most affected by available funding and support for the effort.

To begin with, this initial study is limited in time, money and manpower. The study's funding is $4,000 for 4½ man-months of work. This compares with Fairfax County, Virginia, where a two-year project of design and pilot test of their Urban Development Information System cost $265,000. This was supplemented by $470,000 for extending the field-tested System to the entire county. Fairfax County at that time (1970) had a population of 455,000 and an area of 404 square miles. The Lower Pioneer Valley region's two counties had a 1970 population of 583,031 and an area of 1,182 square miles. A more elaborate information system in Los Angeles County cost many millions of dollars.

It is obvious, then, that if a comprehensive regional information system is to be developed for the Lower Pioneer Valley, where such funding is not available, the establishment of such a system will need to be carried out over a protracted period of time. It will also have to be developed incrementally--a step at a time. And it will need to receive support and cooperation from a number of other agencies and units of government.

INFORMATION SYSTEMS

At this point it is probably appropriate to describe what we mean by "information system." What we are aiming at is a computerized information system, and this is good because understanding a computerized information system greatly assists an understanding of information systems in general.

All information systems are concerned with data processing. But, what is data? Data can be said to be items of information. More specifically, data can be defined as crude statement of fact. Information, on the other hand, is data which has meaning, which becomes useful. "The sea is rough," is data. When it means you can't sail a small boat on it, it becomes information.

An information system puts data together in a way which helps doing work. We gather data on temperature, time, rainfall, etc. to understand weather so we can know when to plant and when to harvest. A planning information system helps planners to do the work of planning. A regional information system, such as we are investigating, helps to do regional planning. It collects, analyzes, processes, projects, etc., data necessary for the making of a plan for regional development.

Because much of this processing of data is routine and menial, it can be done by machines. Even some decisions required in the processing of the data can be predetermined according to certain criteria, and they can be automated.

Because machines can do this work quickly and accurately, great amounts of data can be processed in a short time and at low cost. Machine processed, or computerized, data handling makes information systems more efficient and more economical. It also frees humans for more creative work--work that machines cannot do. Used in the proper way, computer handling of information can make planning or any other task more productive.

There are many types of computerized information systems. They vary in terms of their content, some containing bibliographic data, some geographic data, financial data, census or statistical data.

In terms of structure, they vary internally, according to programs and formats. Programs (the directions to the computer as to how it is to manipulate the data) can control mathematical manipulations, transactional, iterative, or merely storing and retrieval moves. Externally they vary in hardware, including type of computer (digital or analog), input/output equipment (graphic plotter, CRT display, teleprinter, COM, etc.), and mode of operation (time-share, batch).

They also vary in terms of use or application. Some systems are used merely for inventory control. Others are used for directing production machines, for financial accounting, monitoring, forecasting, or simulation. They are used for maintaining records, assisting decision-making, or for actually directing or accomplishing work.

EXISTING LITERATURE REVIEW

Existing literature relating to computerized information systems is very extensive. Articles and books have been published on everything from the mathematical aspects of information theory to computer programming to the use of specific hardware in specific manufacturing machine operations. Bibliographic citations of more than 3,000 separate writings have been perused for possible use in the study. Of these, only those directly bearing on regional information systems were looked into; and of these, only those that appeared the most informative were read and analyzed in detail.

Three experiences in developing a regional information system stood out as worthy of detailed analysis and consideration. These included the Los Angeles urban information system experience, reported in a series of publications by the Los Angeles Community Analysis Bureau in May of 1970. A second experience was that of the Metropolitan Washington Council of Governments. This was reported in a publication, Metropolitan Planning Data from Local Governments, published in 1970. This system, however, was based largely on the report, An Information System for Urban Planning, by Campbell and Leblanc, published in 1962. The third experience was that of Fairfax County, Virginia. The final report, Urban Development Information System (UDIS), came out in March 1972.

Two other works were important in considering the structure of information systems, in themselves, and the use of such

systems by small communities. One of these was the discussion included in the Bibliography on Urban and Regional Information Systems: Focus on Geographic Perspectives, by Wellar and Graff, published in September of 1972. The other was the Feasibility Study: The Possible Application of Computerized Methods to the Decision-Making Process of Rural New England Towns, by Public Decisions Research, published in February of 1971.

EXISTING SYSTEMS

There are a number of existing computerized regional information systems in the world: from the Philippines to Sweden; from KCRIS in Kansas City to GIST in New York City. These include in-house management systems, census systems, land records systems, and in some instances even simulation models.

A survey was made of the twelve regional planning agencies in Massachusetts, and only one reported working on a regional information system. This was the Southeast Regional Planning and Economic Development District (SRPEDD). Their system, however, is primarily an industrial site location file, used to induce new industries to locate in the region. As such, its use at the present time is limited.

It appears, then, that this will be the first attempt in Massachusetts at establishing a computerized general information system for regional planning purposes.

There are several municipal systems in existence within the Lower Pioneer Valley region. Most of these, however, are

only partial or are for special non-planning purposes. They may nevertheless, be used as sub-systems of the proposed regional system. The City of Springfield, for instance, has a well developed land use and housing data bank which could be an important input to the regional system.

Other systems of importance to the creation of an information system in this region include the Census Bureau's DIME File (Dual Independent Map Encoding) for the Springfield-Chicopee-Holyoke Standard Metropolitan Statistical Area. This file identifies census tracts and other data by street address location. It is presently being corrected, updated, and extended into newly urbanized areas by the Springfield Planning Department. Its outstanding significance is that it permits the computerized retrieval, processing, and analysis of data within specific geographic areas bounded by streets, railroads, streams, or other geographic features.

Also included must be the DPW file on road inventories. A pilot study is now being conducted by DPW in Agawam and Longmeadow on merging the DPW road inventory file, the DIME file, a maintenance inventory file, an accident record file, and others to create a Consolidated Information Data File for highway planning.

A third system, for future integration, is the proposed DCA system, covering primarily housing data. Census data, and forthcoming population projections, natural resource inventories, and other state supported data systems should eventually be

included in the regional information system.

Probably most important to the input side of a regional system are what might be termed the municipal sub-systems. These are existing data files maintained by municipalities in the Lower Pioneer Valley region. They include assessor records of real properties, water use billing, and other files used in municipal administration. Many of these are already computerized. Their use for input to a regional data base is essential to the function of a regional information system.

HARDWARE IN THE SYSTEM

Probably the best way to see how a computerized information system works is to look at the hardware that comprises it. Any handling of information is limited by the machines that do the handling. The manner in which information is put into a system, the way it is manipulated, and the form in which it is given to a user can only be that allowed by the machinery of the system, the hardware.

Central to the system is the central processing unit (CPU), the electronic computer itself. This carries on the work of manipulating the data: putting in and taking out, combining, separating, replacing, sorting, arranging. A "non-computer" version of this is the card sorter, which sorts, collects, counts, and tabulates machine record cards (tabulating cards).

Included with the central processor is an internal memory, that part that electronically records the data during processing. It usually is a ferrite core system, which records data

as differing magnetic states in a number of doughnut shaped pieces of iron alloy, arrayed in a wire matrix. A new form is thin film, a system of thin films of magnetizable material, arranged in a wire matrix on a plastic sheet. New forms of memories are being developed continually, but these two are the most used at this time.

The central processing unit and its internal memory comprise the basic computer. But in order for it to work, it must have a means of receiving information. It needs, then, input and output machinery.

Equipment for inputting information to a computer are of many types. Which type is used will depend in great measure on the way in which information is collected and recorded prior to being sent to the computer.

The most simple type is the machine record card, or tabulating card, with its punched holes which a card reader can sense, extracting the information for sending to the computer. Similar to the card is the punched paper tape, whose holes also form a code which a tape reader can read for input to the computer.

Electronic technology has developed magnetic tape, which can record information as well as sound, and this is read by a tape reader unit. In a similar way, a strip of magnetized material is placed on a card and read by a special magnetic card reader. Magnetic strips can be on tab size cards, on ledger cards, and even on letter size bills or statements.

Magnetic tape can be on open reels or in the newer compact cassettes.

The magnetic tape principle is also used in the magnetic disk units and in the magnetic drum units. These, however, are used more as supplemental memory units for the central processing unit. They either enlarge the memory capacity of the computer or hold data for intermediate steps in computation.

Input to the computer can be made directly by keyboard. Most common of these is the electric typewriter used as a teletype. Another way is through optical character recognition equipment. This is equipment which can read special style lettering directly for input to the computer.

Among the more sophisticated and exotic ways is the light pen used in conjunction with the cathode ray tube (CRT) display terminal. The light pen senses the light points in the display and can be used to instruct the computer to add, delete, or change the display in a certain way. A similar technique is the Rand tablet, which is a form of plate made up of an array of wires. A special pencil or pointer can be used with the plate in the way the light pen is used with the CRT terminal.

Other equipment which handles spatial location of data is the video scanner, basically a television camera, and the XY digitizer. The video scanner can convert scans to electronic currents conveying data. The XY digitizer converts

data locations, such as on a map, to numerical coordinates.

Output equipment tends to be the same as input equipment, used for receiving instead of sending. Thus, we have the teletypewriter, the CRT, and pen plotters--either on a drum or on a flat bed. A high speed printer is usually used instead of the teletypewriter. The magnetic tape units are also used to record output from the computer, as are, in an intermediate mode, the magnetic disk and the magnetic drum units. Card and tape punching units can also be used, though they usually get their information from magnetic tape units which, in turn, have gotten data more directly from the computer. A special output device is the COM (computer output microfilming) which prints out data from a CRT directly to microfilm.

Rounding out the computer system are the control units. These control input and output to the computer so as to achieve more economical operation. They have permitted time-sharing of computers by permitting interruption of computations for input or output. They include tape control units, direct control consoles, program interrupts, and data channels or memory-snatch units. As such, they are more a part of the central processing than input-output devices.

The design of a computerized information system must consider not only what equipment is available for handling information but also in what form the information is, which is to be input to the computer. In the same way, the acquisition of equipment will depend on the form of information which is

to be input and the type of display or output of information being sought. Manipulation of information intermediate between input and output is a problem of detailed design of the system. The more important part of the design, overall design, must deal with input and output, for the value of the system lies outside the system--in the use of the finally output information.

II. LOCATION OF SYSTEM

Possible location of the proposed computerized regional information system must be in one of three types of institutions if it is to be anywhere other than at the LPVRPC. These are governmental, educational, or commercial. These institutions now have, or soon will have, the computer hardware necessary for the system, and they can make that hardware available for the processing of regional information system data.

GOVERNMENTAL

Potential locations for regional information system data processing by governmental institutions can be found on three levels: federal, state, and municipal.

On the Federal Level

The most promising facility is that of the Environmental Protection Agency, Office of Water Programs (OWP). This office has a computer facility in McLean, Virginia, a few miles west of Washington, D.C., which is being used by state and local water quality control agencies throughout the country.

The system is called STORET (for STOrage and RETrieval), and it is available to additional users. The only cost is for rental of a typewriter terminal which is linked to STORET by telephone. Cost of computer time, data storage, and processing is paid by EPA. Experts from the OWP are available to teach each user how to feed his information into the

computer and to order the processing he wants. Programs and instructions are already prepared by.OWP. The user merely selects the instructions that fit his needs and the type of printout he wants: tabular, graphic, or even digital plots on maps.

STORET's memory contains data on water quality from sampling stations throughout the country over periods as long as 20 years. These can be retrieved in statistical form--means, extreme values, deviations, etc., as desired--for any locations, watersheds, or individual streams. Federal and regional water quality standards are stored, and constantly updated, for comparison with sampling data. Records of municipal and industrial waste discharges and fish kills are stored along with management information on pollution abatement needs, costs, and implementation schedules.

All these data are constantly being updated as OWP and cooperating states and localities use STORET and add new information. The data are protected from loss, and if a user wishes to restrict his data from the common store, he may do so. And all of this is available with one typewriter terminal and a telephone connection.

On the State Level

The Department of Community Affairs, in their attempts at establishing a state housing data system, has recommended utilization of the computer facility belonging to the Department of Public Works. The envisioned system has evolved to

a "Land Use Information System for Planning and Development." It is hoped that numerous state agencies can contribute to the data bank for the system, including Public Works, Natural Resources, Labor and Industries, etc. It is hoped that municipalities and regional planning agencies can also contribute to the system, which would provide for sharing of data among participants. Either a terminal for each participant would have to be provided for time-share use of the system--the terminals connected to the DPW computer or some other under State arrangement--or data would have to be conveyed to and from the central computer facility by courier for batch processing.

On the Municipal Level

Municipalities within the region are the most accessible governmental unit having computer facilities.

The City of Springfield has the largest municipal facility: a UNIVAC 9400 with 98K core memory capacity. It also has extra disk and tape memory units, card reading and punching units, and a high speed printer. Plans call for future CRT terminals and a teletype terminal in the City Planning Department. Besides numerous city accounts, the system also handles accounts for the nearby Towns of Ludlow, East Longmeadow, and Wilbraham.

The second largest municipal facility is that of Chicopee. This facility includes an NCR Century 100 central processing unit, a 2-disk memory unit, high speed printer, card reader,

and two key punches. The system includes accounts for payroll, water billing, real estate tax billing, and assessor's lists.

No other municipality in the region has a municipal computer facility, but several use other public computers within the municipality or have data processing done by nearby educational or commercial facilities.

The City of Westfield uses a computer facility of the School Department to do some municipal account processing. The City of Holyoke has some processing done by commercial services and some by Holyoke Community College. Agawam has a computer in the School Department which is used by some municipal departments. Amherst has used the facility at Amherst College for some data processing, as well as that of the University of Massachusetts. Northampton's School Department has and uses a computer, though the City still uses business machines. West Springfield uses data processing services provided by a commercial firm.

EDUCATIONAL

Educational institutions represent another possible location for the computerized regional information system.

Probably the most feasible facility, and certainly the largest educational facility, is that at the University of Massachusetts Computer Center in Amherst. This center, which serves educational and research purposes, houses two Control Data Computers, models 3600 and 3800, each with 32K capacity

in their central processing units. Supporting these are 8 tape drives, 3 disk units, 3 drum units, 2 card readers, 1 card punch, and 2 high speed printers. One small computer, a GTIS-1000, multiplexes messages for time-sharing, and a PDP-8 computer monitors the main line system. For terminals there are 9 CRT's, 102 Teletype typewriters, 44 Datel selectric typewriters, 3 Carterphone 1030 typewriters, 1 Exuport portable typewriter terminal, 1 Techtronics graphics terminal (using a CRT for two-way graphic communication), 1 CalComp drum plotter, and 1 Hewlitt-Packard flatbed plotter.

Holyoke Community College has a Honeywell Series 200 central processing unit with 32K memory. This facility has 2 disk drives, 1 card reader/punch, 1 high speed printer, 3 key punches, and 1 card sorter. The school plans to acquire new hardware which would provide greater capacity as well as time-share capability.

Amherst College has a computer which has been used on occasion for batch processing by the town. Agawam, Northampton, and Westfield School Departments have computers which do data processing for the schools and in some cases for the municipal administration. Other schools, including commercial educational institutions, have computer facilities, but their small size and limited accessibility preclude their use for a regional information system.

Outside the Lower Pioneer Valley region, Worcester Polytechnic Institute's computation center has a large facility

which has been used by the LPVRPC for the processing of census summary tapes. It is possible to make another arrangement with this institution for processing data for the regional information system.

COMMERCIAL

There are four firms in Springfield and four in West Springfield which provide data processing services as commercial enterprises. In addition to these, there are commercial data processing services in Hartford, Connecticut, and its suburban communities. Services are also available from firms in Boston and its suburbs and from other towns and cities in Massachusetts outside the Lower Pioneer Valley region.

Though it might not prove economically or technically feasible (since most of these firms are organized for financial accounting services) to utilize any of these for the computerized regional information system location, their services might form an important input to the system since many do processing of data for municipalities in the region. Some arrangement could be made to eliminate duplication of such data processing or data bank construction in the creation of the computerized regional information system.

III. SYSTEM STRUCTURE

The structure of the proposed computerized regional information system must be based on the use to which the system is to be put. For the purpose of this study, the use of the system will be to assist the regional planning process. Applications and design of the system, therefore, must satisfy the needs of regional planning above all others.

APPLICATIONS

Applications of the information system will be determined, for the most part, by the data available and by the manipulations of that data which the computer can do. The demands of the regional planning process will, at the same time, prescribe certain applications as basic if the system is to have utility.

Without attempting to enumerate all of the fields of interest to regional planning, we must admit that the regional planning process has need of data on natural resources, land use, housing, population, public utilities and facilities, and transportation. This data permits basic planning of water and sewer systems, outdoor recreation and open space, land use and development, and the preservation, conservation, and utilization of natural resources. Regional planning needs other data to do a comprehensive job of planning. Financial data of the various governemts in the region would be needed for capital improvement programming, and other data for other planning functions.

The basic manipulations the system could do would include recording, storage or retrieval, analysis or synthesis, monitoring, mathematical calculations, and manipulations as intermediate steps in a work process. The basic applications of the system would include inventory control, record maintenance, financial or other transactional accounting, forecasting or projection, and simulation. These applications would be instrumental in maintaining and updating data, directing or accomplishing routine work, and assisting decision-making in the formulation of plans and programs.

Applications of this system in regional planning would include population forecasts; land use mapping; development monitoring and reviews; functional analysis of transportation, water, sewer, and solid waste systems; housing inventories; and housing needs forecasts. There are many other possible applications of a regional information system, but consideration of these should be sufficient for a system feasibility study.

REQUIREMENTS AND CONSTRAINTS

An information system consists of people, computer equipment (hardware), data processing programs (software), a collection of data (data base), and adopted procedures for processing the data. To function it must collect, store, and periodically update data, and make that data available to users on request or on schedule.

The Regional System

The purposes of a regional information system are, first, to facilitate the operational requirements of planning and, second, to permit the exchange of information among governmental units and with the public. This type of system should also provide for the aggregation and analysis of information on a regional basis, or for any defined geographic area within the region--an added benefit to system participants and users of such information. The outcome of it all should be a superior foundation of information on which to base decisions affecting regional planning and development.

The municipality must be considered the basic building block of a regional information system. Much information of importance to regional planning, and most information of a statistical nature, originates at the municipal level. Even data gathered by other levels of government, such as the federal census, are reported in municipal units. At the same time, regional planning is implemented primarily by municipal decisions and actions. Since the origin of data and the final destination of planning use of such data is most often at the municipal level, an essential by-product of the regional system must be satisfaction of municipal requirements.

A comprehensively designed regional information system which has the municipality as the basic building block must encompass all major functions performed by any municipality. It must also be designed to accommodate new functions or

variations in functions which might be performed by different municipalities.

An information system based on municipalities as building blocks must include each municipality as a major sub-system. Municipal government functions must also be included as sub-systems. These municipal government functions exist as organizational entities within the municipal government and as information sub-systems. Information sub-systems, however, must be considered independent of any municipal organizational structure. They must be defined in terms of information handling: input, storage, processing, and output.

The components of the subsystems are points in the system at which input and output of information occur and where the information sub-system structure and the municipal organizational structure coincide. This generally is at the operational level.

The regional information system, then, must also be operationally grounded. That is, it must be supported by the operations of municipal agencies, and these agencies must contribute to the maintenance of the data base for the system.

The regional system might also be designed to support, to a degree, municipal operations. While providing for subsystem input of data required by the needs of regional planning, the system could also provide for output of information useful to the municipal function sub-systems and their operational components.

Computerized information systems are relatively new in the world. There is a particular lack of experience with comprehensive information systems which support regional planning. Where municipalities have established information systems, they have done so with goals limited to the computerization of specific administrative operations, such as payroll and accounting, personnel records, school administration, tax records, etc. These applications of computer techniques to handling information could be managed with a minimum of documentation, required little if any organizational change, and avoided the complexities of geocoding or other useful but difficult manipulations of data.

There is a wide gap between what could be done with current information technology and what is done in municipalities or at other government levels. While the benefits to be gained from utilizing such technology are readily apparent to those that understand the situation, adoption of such techniques by the various levels of government is quite another problem. Indeed, a regional information system could profit greatly from the utilization of computer techniques at the municipal level. The municipalities in the region should therefore be educated to the benefits of both the handling of municipal information by computers and the participation of municipalities in a regional system, even if this must be a long-term and gradual process.

The Computer Process

Though we are considering here the feasibility of a computerized regional information system, and though a computer is essential to such a system, the formation of a regional information system is primarily an intellectual problem of handling information for the regional planning process. The planning process should be emphasized and the computer should be de-emphasized to its proper position as merely a tool of the system.

Nevertheless, the computer has its own requirements for participation in the system. To use the computer process, the following must occur:

1. The data required must be put into a form the computer can use.
2. Someone must tell the computer what to do.
3. The computer must accept the data and manipulate them according to the instructions.
4. The results of the computer's manipulations must then be put into a form people can use.

The computerized regional information system must be provided, then, with appropriate and adequate core equipment and peripheral devices. This includes a central processing unit, accessory memory units, and input and output devices--all matched for most efficient use. This equipment will comprise the basic hardware of the system.

The hardware of the system must have the following abilities:

1. to accept and output data in a variety of forms

2. to handle large volumes of data

3. to operate on the individual values in a data set

4. to manipulate and alter data set structures

5. to restrict the ability of users to use or alter data set structures

6. to display data being manipulated, by alpha/numeric writing or by graphic display

7. to retrieve and analyze data

8. to provide data base reference service and documentation

9. to provide process management and control

In order to use the hardware to accomplish the objectives of the system, appropriate programming (software) must be accomplished. A conventional, higher order language should be used. The programming should emphasize automation, be self-documenting, and be generalized enough to meet the broad needs of the system while minimizing the programming knowledge required of the user. Adequate manuals and technical training of staff should be included in the system.

The development of a computerized information system must, for practical reasons, consider a special dimension of system structure--that of sophistication. There are many levels of sophistication in systems and sub-systems, and a system can be developed incrementally, moving from a low level to a higher level of sophistication as the system is built.

At the lowest level of sophistication are simple data processing techniques. Software for this is generally limited

to data and files management, cross-tabulation, and report generation. The mode employed is usually batch-processing of data.

A second level of sophistication is control of operations including scheduling, allocation, and monitoring. Required here is rapid response capability, and the mode is on-line, real-time. Time-share computer configuration is called for.

A third level of sophistication is planning and administrative support. Software must permit PERT and/or critical path analysis, planning-programming-budgeting systems, and statistical analysis.

A final level of sophistication is the policy-making and management level. Software required here is for exception-reporting techniques, automatic alerting systems, on-line cross-tabulation, simulation, and automated decision reporting.

Development of the system in an incremental manner would permit short-term benefits through concentration on important sub-systems. Improvements to the system could also occur through experimenting with sub-systems or with new levels of sophistication without the need to interfere with or alter those parts of the system already established.

Data Handling

The nature of data and the need for its economical and efficient handling impose other requirements on the structure of the regional information system. As to data itself,

whatever is collected must first have some direct bearing on regional planning operations. It is better for it also to have direct or indirect bearing on municipal operations. The data should be readily available. It is more economical if the data are maintained and updated by the originating agency. The data should be retainable as histories so that some understanding of change in data over time can be had. The data should be in or translatable to some common geographic or areal unit, and the smaller the unit the better. The data should be capable of being aggregated, and this should include aggregation by geographic units.

Data discipline.--Certain discipline and restraint should be practiced in the collection and handling of data. The system should contain no data for which there is no known use, and only as much of a particular class of data as is needed should be put into the system. The way data are to be used should determine the way they are to be put into the system, i.e., visual, verbal, numerical, analog, etc. The frequency with which data are needed should determine their accessibility in the system. New data should be put into the system as often as is necessary.

Data acquisition.--The actual acquisition of data for the system must be preceded by the development of a definite plan for acquiring, organizing, and updating data. The development of the plan should begin with an inventory of existing sources and the data they can furnish. The data should be

reviewed as to machine-readability and the data formats as related to the proposed system formats. In many cases, the data acquisition plan will contain agreements with different entities on the furnishing of data and the sharing of that data with third parties.

Data base management.--One of the larger fields of activity with respect to sustaining the information system is that of data base management. This includes data documentation, assuring data compatibility, control of the data base, and continuing acquisition and updating of data.

Data documentation is providing data about data. It is especially necessary for a system which receives data from a number of independent participants. The documentation should cover the exact name of each datum, a discussion of its meaning, its source, the frequency with which it is reported, the range and kind of values associated with it, the meaning of codes associated with it, the medium, place, and identity of its storage, and any limitations on its use. Cross-reference should be made to records and files in which the data appear. Finally, the data documentation should be kept in such a manner as to be intelligible to users other than those who are involved in maintaining the data base.

The data base should contain a large number of commonly defined, readily comparable elements of data. In many cases this can be done by standardization of data formats. In cases where data formats cannot be standardized without disruption

of data sources or data flows, translation of data identifiers will be required.

Assuring data compatibility involves setting up data identification schemes of a number of types. Data can be identified by various names, codes, and classifications, which would include or be coupled to data descriptors. Data identifiers may need to be translatable where different identifiers are used by different entities to identify the same data. Limitations on data use could be included in the list of possible descriptors and identifiers.

Data release is an important aspect of the regional system where much of the data will be derived from participating but independent entities. This is especially true where sensitive data are involved. A data access control plan will have to be an important part of the data management scheme.

The data access control plan should include explicit rules governing the release of data or aggregations of data. These will generally be the result of agreements with those entities providing data to the system on what can or cannot be done with their data. In these agreements and in the data access control plan, authorized users should be identified, what data may be released to each, when and how often it may be released, and in what state or aggregation.

The plan, however, should also allow for providing restricted data at special request. This could be by a board or official which acts upon such requests while providing an

auditable record of such actions. Though the data access control plan is primarily necessary to protect sensitive data from improper use, it must also facilitate the release of data to legitimate users. It must be remembered that the system itself has as one of its purposes the provision of information to support better decisions affecting regional planning and development.

The data access control plan should contain software design as one of its means of control. This would limit the need of human administration of data access in routine situations. The software programming should at the same time include self-documentation with respect to requests for and release of controlled data.

Data maintenance is another aspect of the data base management. An important consideration is that data collected by any means has only transitory value. This may be expressed as half-life. For instance, the half-life of socio-economic data collected by the U.S. census is estimated as about two years. This means that in two years half of the data is no longer of value; in another two years half of the remaining data no longer has value; and so on. The half-life of all such data is becoming shorter in our dynamic society. Continual collecting and updating of data is, consequently, essential to continued utility of the information system.

Geocoding.--Geocoding gives a dimension to data of special importance to regional planning. This is a system of using

data descriptor codes related to maps or other geographic files. The assignment of such codes permits, first, the geographic location of the data. Geocoding also facilitates the aggregation of data into any one of a number of geographic groupings or jurisdictions. Finally, it permits spatial analysis of the data and graphic display in map or other form.

Geocoding should provide for unique location identification, such as coordinate values, political jurisdiction, street address, block front, grid squares, etc. The processing part of the information system should allow for translation of one such identification code to another. It should also allow for identification of any one code structure's relation to another, and it should permit distance and area calculations and correlations.

Interfacing.--Data handling for a regional system must also include provisions for interfacing or linking with other systems. Many of the inputs to the regional system will be data from municipal or other systems in machine-readable form. The regional system should, conversely, be able to output data to these participating systems. The regional system should also be open-ended to permit expansion through the addition of direct links to other existing or newly created information systems. Since the multiple sources of data will require translation of data formats and data identifiers for the flow of data in and out of the system, interfacing should be a significant part of the software design of the regional

system, whatever level of sophistication is attained.

System Documentation

In designing the information system and its documentation, a guide should be the statement, "if it can be misunderstood, it will be misunderstood." In this sense, documentation should include not only instructions but also explanations of instructions. The system should be made explicit in every detail. Local circumstances requiring deviation from the general structure should be identified. The system documentation should be written for comprehension by those not directly involved in the system operation.

System documentation should include the following:

1. Catalogs--inventorying all data sets, data files, programs, and other system documents.
2. Dictionaries--defining data names, codes, files, etc.
3. Thesaurus--for translation of data names, values, areal units, etc.
4. Directories--indicating where material is physically stored.
5. Lists--of data sets, inventories, etc.
6. Chronologies--of data for time analysis of inventories.
7. Bibliographies--inventorying reports and publications pertinent to the system.
8. Programs--for operating the system.
9. Manuals--on how to operate the system and its equipment.
10. Data acquisition, management, and control plans.
11. Agreements with participants of the system on provision and use of data.

12. A comprehensive description of the system and its documentation.

DATA AVAILABILITY

The collection of data to form the data base for a computerized regional information system, for practical reasons, should begin with data already available in machine-readable form. There is, of course, a great deal of data available for regional planning. The amount and type in machine-readable form, however, is rather limited.

The more detailed data is available at the level of the lowest geographic unit, usually the municipality. Investigations into this have shown that the most usable data for regional planning information are data on land use from the assessor's records and data from water use billing. Financial records which were processed by computer were found too sensitive for input directly into a regional system and consequently were not investigated in detail at this time. Where such data are not maintained by the municipality itself, commercial services or other service entities which process the data might be able to provide them as input to the regional system--subject to agreements.

Other important data at the municipal level are available from the U.S. census summary tapes. Printouts of these tapes are kept at the LPVRPC offices, and the original tapes can be made available for further processing. These tapes contain data for census tracts and enumeration districts,

most of which are smaller than municipalities.

The Bureau of the Census also can provide numerous programs for processing these tapes and programs for connecting these data with geographic data. The Bureau created the DIME system (Dual Independent Map Encoding) for geocoding and the MMS (Metropolitan Map Series), which covers the Springfield-Chicopee-Holyoke metropolitan area, as a geographic base. The DIME system, which formerly covered only the urbanized area, is now being extended to cover the entire standard metropolitan statistical area.

The LPVRPC also has a printout of the Massachusetts Department of Public Works' road inventories for the entire planning district. DPW is working on tying these inventories to the DIME system, using Agawam and Longmeadow as a study area. This will make available a great amount of data from DPW which can be correlated with census data.

Other federal agencies can also provide both data and computer programs for processing data. These include the Department of Housing and Urban Development (HUD), the Department of Transportation (DOT), and the Environmental Protection Agency (EPA). The Commonwealth of Massachusetts can also provide data from a number of its agencies for inclusion in the regional information system. At the same time, data from the system can be exchanged with these agencies.

IV. COSTS AND BENEFITS

COSTS

There are two kinds of costs which must be considered in determining feasibility of a computerized regional information system. These are capital costs and operating costs. From a purely practical point of view, both types of costs are too great with respect to the present Commission budget for development of an information system. From a theoretical point of view, however, the system might still be feasible.

Capital costs include the hardware of the system (the central processing unit and peripheral devices of the computer equipment), software (the programs and documentation of the system), and space to house the system. From an early point in this study it was clear that hardware and space costs and much of the software costs were beyond the foreseeable in-house capabilities of the regional planning commission.

The objective of this element of the Commission's work program has not therefore aimed at the feasibility of acquiring a computer facility. Rather, the study examined the feasibility of utilizing computerized data processing in the handling of regional planning information. It was felt that applying computerized information handling technology to the work of the Commission would increase productivity and improve the quality of planning recommendations.

Cooperative utilization of computer hardware facilities and computerized data files existing within the region for

regional planning work was found to be a feasible manner of establishing a regional information system. It was concluded that a greater return would be achieved by directing the future use of Commission funds in this area towards the development of such a cooperative activity, particularly if existing facilities are not now fully utilized.

Operating costs should include time and material costs in the use of the computer and its peripheral equipment, staff salaries and overhead, and special staff training and consultant services. Normally, operating costs of an information system would also include the costs of gathering data and making data available to the system through converting it to machine-readable form. However, the gathering of data and its processing, where it is being done by entities other than the Commission, represents a cost which could be avoided by the regional system if selected data were to be collected and prepared for input by each participating entity. Avoidance of duplication in this way reduces considerably the operating costs which would normally form a large part of the regional system's cost. Maintenance of the basic data by the participating entity also need not be duplicated in the system's operating cost. In practice, then, much of the operating costs can be borne by participating entities.

The use of existing computer facilities, the sharing of data, the maintenance of the data base by the participating entities originating that data, and the avoidance of cost

duplication wherever possible make the difference between the task of establishing a prohibitively costly, independent regional information system and an at least theoretically feasible, cooperative system. A cooperatively built system is the most economical way to create a regional information system, and probably the only way, considering the source and cost of data for the system.

USER BENEFITS

Of course the principal beneficiary of the regional information system will be the Lower Pioneer Valley Regional Planning Commission. The system should provide the Commission with basic data and information for use in the regional planning process. Examining only the most promising elements, we can easily see the regional information system could support the recording and review of land use and development trends in the region. It also could support water quality management planning, transportation planning, capital improvement programming, and required reviews of development projects involving state or federal assistance.

Participants in the system would also receive substantial benefits from the regional collection and correlation of data. Municipalities and areawide agencies would gain from regional aggregation and analysis of data. Such information would be instrumental in municipal and areawide planning efforts, permitting a view of specific problems in relation to regional patterns and processes. Information from the system could

also be used in municipal review of projected public works, proposed zoning changes, and applications for building permits. Municipal capital improvement programming could also benefit from greater use of regional data and closer correlation with regional development trends.

State and federal agencies would also benefit from participation in the regional information system. Data provided by the regional system would be an important consideration in the design and management of their programs. Such data could be an aid to the allocation of funding assistance to lower levels of government and could indicate to the higher levels the local needs for assistance. These agencies could also use the regional data in their review of development proposals.

SUPPORT

Support of a cooperative regional information system would be dependent to a great extent on the implementing arrangement for the establishment of the system. This arrangement would, because of the number of participants, be a collection of agreements, each differing according to the degree and type of cooperation obtainable from the participant.

One of the more important considerations in any participation agreement would be the independence of the parties. The municipalities and the state and federal agencies have their own sources of funds and restrictions on spending such funds. Their own operating needs will affect the manner and

the extent of their participation and any financial or other support they can give to the system. Another consideration in any agreement would be the sensitivity of data to be provided for the system and the restrictions to be put on the use of such data.

The LPVRPC would be the central party in the formulation of these participation agreements. Its functions should be well defined and should include the following:

1. Serving as a coordinating mechanism with respect to development and operation of the regional information system.

2. Recommending, encouraging, and coordinating fiscal and material support for the information system from participating agencies and municipalities, supplemented by resources from other sources.

3. Studying and recommending standardization of data elements and processing procedures for general use in the information system.

4. Devising and recommending the best arrangements for participation by federal, state, and local entities in the information system.

Financial support for the information system would have to come from two general sources--from participants and from the LPVRPC. Participants can contribute financially to the support of the system by supporting their own data processing by computer methods. They can thereby assist in lowering system operating costs by obviating duplication of work and accompanying costs. Participants can also support the system financially through supporting the processing of data from which they benefit directly. This can be through user fees for specific requests or through group support of data

exchanges or data aggregations and/or analysis.

The LPVRPC can financially support the system from its general operating funds where the system assists its normal operations. This is where computer economies will show in increased productivity. The Commission can also obtain, on occasion, grants from its own supporting agencies for system development or special studies utilizing the system. Some of the grants could come from HUD, DOT, or EPA, among others. Support could also be forthcoming from the state where the regional system could support state use of regional data. Other sources of financial or material support of the system could become evident as the system is developed.

V. CONCLUSIONS AND RECOMMENDATIONS

Significant conclusions from this study are few and can be simply stated.

1. The creation and operation of a computerized regional information system for the Lower Pioneer Valley Regional Planning District, with the Commission as coordinating authority and with the active participation and support of other agencies, is feasible.

2. Acquisition of new hardware for a computer facility is too costly for the present or anticipated budget of the Commission.

3. Cooperative construction of a data base for the system and cooperative support of operation of the system are necessary.

4. Multiple agreements with participants are a necessary arrangement for establishing and operating the system.

5. There now exist sufficient data for initiating data base construction.

6. Use of computer techniques could result in economies derived from greater productivity, lower costs, and higher quality in planning or administration.

Guided by the above, the following recommendations for establishment of a computerized regional information system are made:

1. The system should be developed incrementally: bit by bit, sub-system by sub-system.

2. A beginning should be made by integrating existing data files on land use, roads, water use, and census data.

3. The system should first support land use planning, transportation planning, and water quality management planning.

4. System development should be continued in coverage and degree of sophistication.