EPA-450/3-76-011
August 1975

CDHS EXECUTIVE SUMMARY REPORT

U.S. ENVIRONMENTAL PROTECTION AGENCY
Office of Air and Waste Management
Office of Air Quality Planning and Standards
Research Triangle Park, North Carolina 27711

EPA-450/3-76-011

CDHS EXECUTIVE SUMMARY REPORT

by

International Business Machines Corporation
18100 Frederick Pike
Gaithersburg, Maryland 20760

Contract No. 68-02-2054

EPA Project Officer: Lloyd Hedgepeth

Prepared for

ENVIRONMENTAL PROTECTION AGENCY
Office of Air and Waste Management
Office of Air Quality Planning and Standards
Research Triangle Park, N.C. 27711

August 1975

This report is issued by the Environmental Protection Agency to report technical data of interest to a limited number of readers. Copies are available free of charge to Federal employees, current contractors and grantees, and nonprofit organizations - as supplies permit - from the Air Pollution Technical Information Center, Environmental Protection Agency, Research Triangle Park, North Carolina 27711; or, for a fee, from the National Technical Information Service, 5285 Port Royal Road, Springfield, Virginia 22161.

This report was furnished to the Environmental Protection Agency by International Business Machines Corporation, Gaithersburg, Maryland 20760, in fulfillment of Contract No. 68-02-2054. The contents of this report are reproduced herein as received from International Business Machines Corporation. The opinions, findings, and conclusions expressed are those of the author and not necessarily those of the Environmental Protection Agency. Mention of company or product names is not to be considered as an endorsement by the Environmental Protection Agency.

Publication No. EPA-450/3-76-011

NOTICE

This report discusses the following three subsystems of the Comprehensive Data Handling System (CDHS):

Air Quality Data Handling Subsystem-II (AQDHS-II)

Emissions Inventory/Permits and Registration Subsystem (EIS/P&R)

Enforcement Management Subsystem (EMS)

While the EMS was originally included as a subsystem of the CDHS, installation, maintenance and other matters dealing with EMS are handled by the Division of Stationary Source Enforcement (DSSE). Any inquiries regarding the EMS should be directed to:

Office of Enforcement and General Counsel
Office of General Enforcement
Waterside Mall, 3220L
Washington, D. C. 20460

All inquiries regarding AQDHS-II or EIS/P&R should be directed to the Office of Air Quality Planning and Standards.

1.0 THE COMPREHENSIVE DATA HANDLING SYSTEM - AN OVERVIEW

This section provides a conceptual description of the Comprehensive Data Handling System (CDHS) to establish its role in air pollution control and to give the user an insight into the overall application of the system. This foundation will allow the user to incoporate CDHS subsystems into his normal routine with greater appreciation of the advantages and benefits that can be obtained.

1.1 INTRODUCTION

Air Quality Management is today one of the most important of the public disciplines. It assumes this importance because of its relative newness and its pervasive influence on our society.

While the social importance has certainly caused some helpful fiscal backing, it has also created severe challenges for Air Quality Managers. Working under severe time constraints with limited resources the state Air Quality Manager has had to:

a. explore this new area of concern,

b. establish procedures for enforcing air pollution control,

c. prepare and implement long term policies which will lead to improvement of the air quality,

and always in the harsh spotlight of public scrutiny.

The exercise of air pollution control is highly dependent upon the collection, processing and analysis of large amounts of data. This data, which describes both the sources of pollutants - the emission inventory and the resultant

ambient concentrations - the air quality file, is constantly changing and growing in amount. Proper maintenance and processing imposes special resource requirements on the state and local agency.

The Environmental Protection Agency has recognized the work load imposed on state and local agencies as the result of data collection, processing and reporting functions arising from existing Federal regulations. EPA has responded by creating a data handling system patterned to satisfy both state and Federal needs for collecting, processing, and reporting air quality and emissions data. The need to manage resources and schedules is also recognized as an inherent part of this system. The EPA solution is the Comprehensive Data Handling System (CDHS) which is comprised of:

The Emission Inventory/Permits and Registration Subsystem (EIS/P&R)
The Air Quality Data Handling Subsystem (AQDHS II)
The Enforcement Management Subsystem (EMS)

The CDHS has been designed to support state and local agencies by providing an automated means for managing their emissions and air quality data in support of the many functions of air pollution control. A significant feature of the system is the capability to automatically respond to the Federal reporting requirement for both air quality and emissions data. Figure 1.1-1, CDHS System Summary, provides a ready reference relating air pollution control functions to CDHS support capabilities. The manner in which CDHS support capabilities are applied at the state and local level are purposely left to the descretion of the Air Pollution Control Manager so that he can have full control of his operation while exploiting the advantages of automation.

EPA's approach to assisting both state and local governments in air quality management has been to regard air pollution control as an overall system in and of itself. This system is complex with many internal interrelationships, with distinct feedback loops with many external influences. The system depends on the collection, evaluation and reporting of a large amount of data. Figure 1.1-1 illustrates these features and highlights the data relationships between

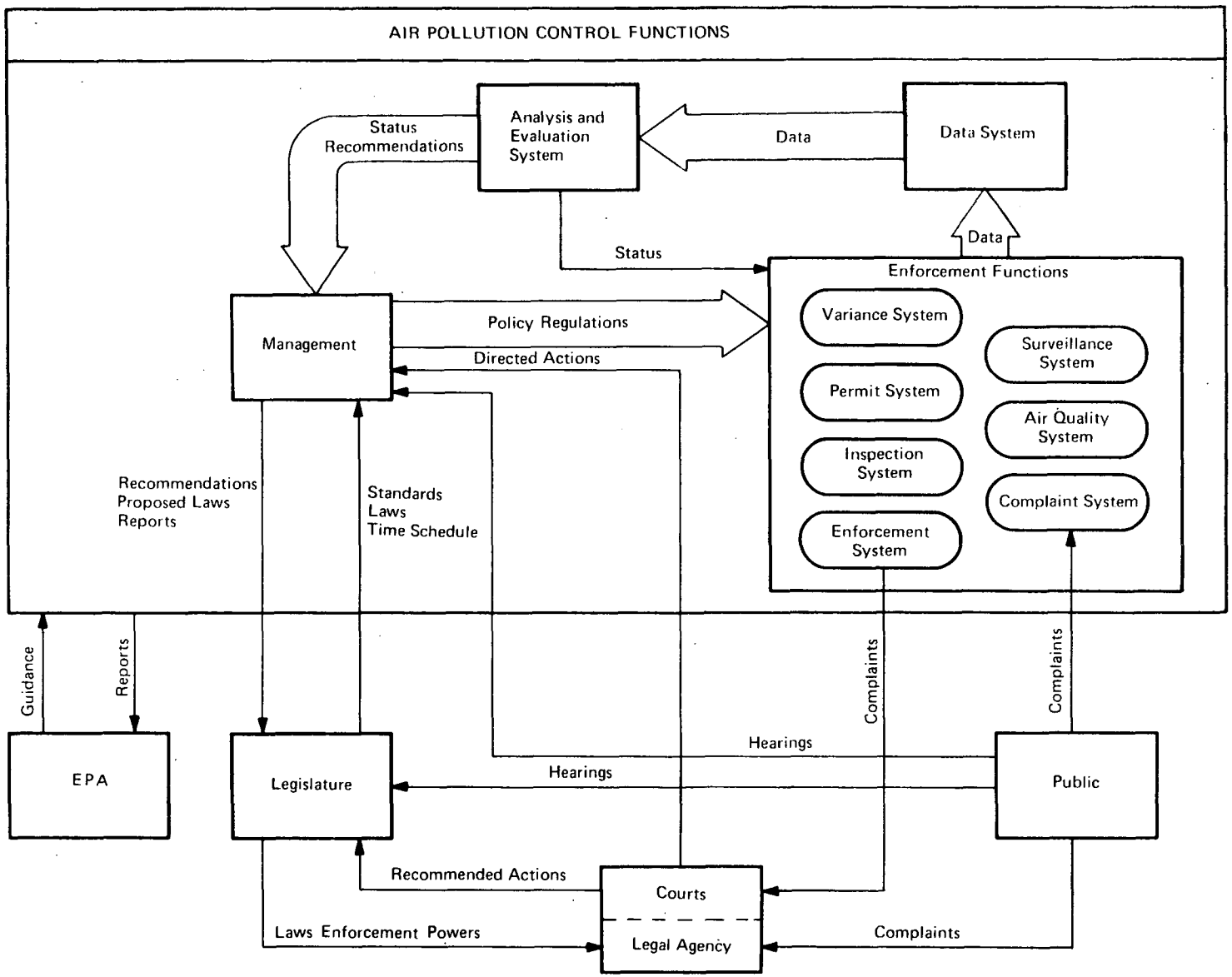

Figure 1.1-1. An Air Pollution Control System

management, analysis and the enforcement functions. The source enforcement functions generate data describing emission sources, air quality and related information. This data, in turn, is analyzed, evaluated and reduced to management reports measuring the success of control activities. Management translates the findings into policies and, where necessary, regulations.

The nature of the end product in air quality management suggests that measurements are needed both at specific locations and at less specific or global locations. This monitoring must be of sufficient resolution so as to adequately describe the air quality found; it must also be measured with sufficient rapidity to be of more than historical value. With these considerations in mind, the system structure has been adapted for CDHS to address the many interfaces and feedbacks of the air pollution system as well as the pure data handling functions.

1.2 COMPREHENSIVE DATA HANDLING SYSTEM (CDHS)

Management of air quality requires constant surveillance of air quality and careful recording of all significant pollution sources with respect to:

a. Location by name and address
b. Categorization by emission type and quantity
c. Identification with processes
d. Relation to other facility attributes.

Measurements are required to trace progress of particular emission control attempts and of the general state of air quality. Involved long range plans are required to change and control specific plant processes. Detailed analysis is required to predict future air pollution conditions and to determine the contributory effect of any particular emission. These considerations must, in

urn, be related to Air Pollution Control Implementation Plans to ensure that nforcement actions are achieving desired goals. These and the many unmentioned asks in air quality control add up to a massive data handling requirement.

he CDHS was designed to meet these specific requirements by responding to the omplex relationships of air pollution control as depicted in Figure 1.2-1. It rovides computerized data storage and retrieval subsystems to serve in rocessing the details of point and area emission inventories as well as high olumes of repetitive air quality measurements. Files and procedures necessary o manage the long range scheduling inherent in bringing emission levels under ontrol are provided as part of the system.

All subsystems have been developed with state and local needs in mind. Consequently, two system features are significantly important. All systems contain the capability for agencies to locally devise qualification criteria for accessing the data bases. This means that each agency can selectively retrieve data records (or portions thereof) according to their special needs. All systems can accommodate a wide variety of data (e.g., multiple sources, pollutants and air quality) which are selected locally. All subsystems enable users to select records from the data base and process the answers using subsystem programs to produce several classes of reports, including detail and summary reports. Provisions are made to simplify the addition of programs to extend the type, content, and formats of reports that can be generated from the data bases. In addition, the extended version of EIS permits an agency to record, store, and access information whose content and format are locally selected. This capability gives the agency a means for handling special information beyond the normal scope or of different content from generally applicable air pollution control data.

The subsystems are capable of accepting machine readable information formatted according to the CDHS defined or the Federal (EPA) formats for emission and air quality data. The internal files are designed to accept multiple formats

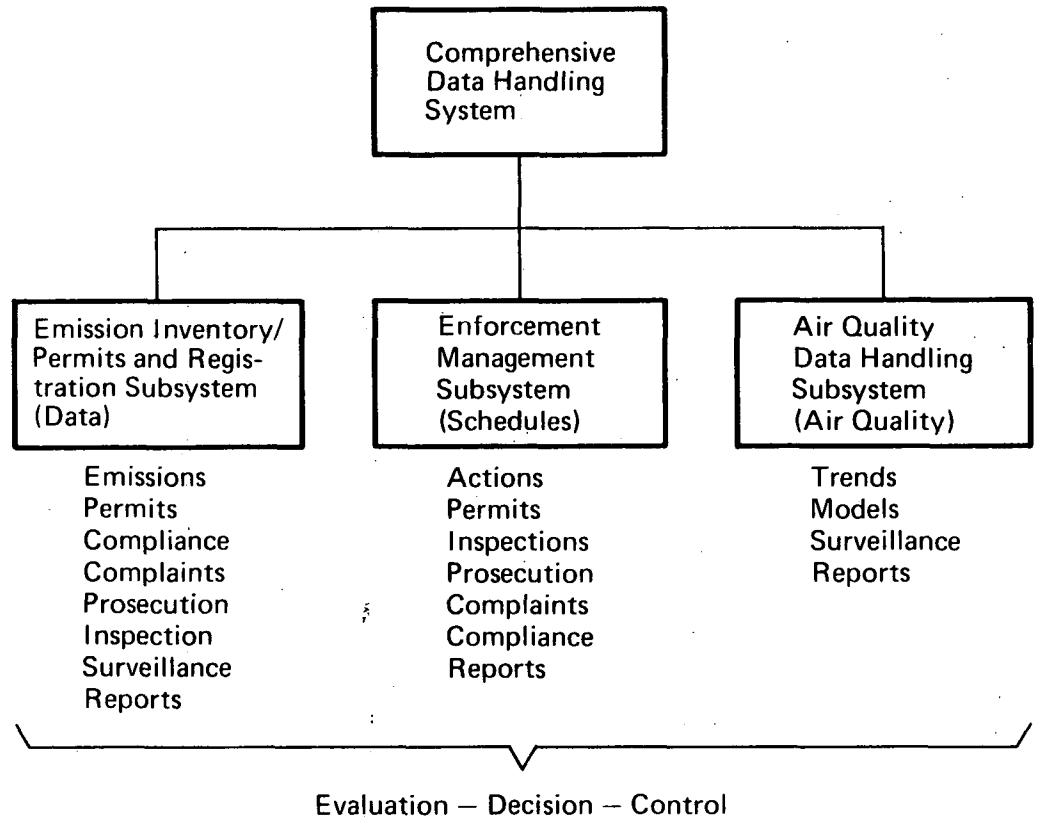

Figure 1.2-1. CDHS Design Overview

including those defined by the system, National Emissions Data System (NEDS), the Storage and Retrieval of Aerometric Data (SAROAD) System or the local agency. Full compatibility with Federal requirements is, therefore, achieved while allowing agencies to handle data in local formats.

The retrieval language format is common to all subsystems, thus providing a uniform user interface. In all cases, this facility takes simple but powerful logical expressions of the type :

IF a condition is true, THEN select the record,

e.g., IF STACK ='100'

and converts them into computer program statements which test the data base. The language allows for compound logical queries which give the user the ability to selectively retrieve from one to many records. This capability lets the user rapidly respond to many different types of questions related to air pollution control. If the test is successful, a copy of the qualifying record is added to an output file called an answer set from which appropriate reports are generated.

CDHS serves an Air Quality Control Agency by:

a. Allowing better utilization of professional personnel by providing organized data handling systems.

b. Assisting management by giving day-to-day visibility of the status of operations such as air quality monitoring, source monitoring, permits and registrations, inspections, complaints, surveillance, compliance, schedules, and legal actions.

c. Providing a mechanism for appropriately formatting the basic data for inputs to Federal reports.

CDHS, or any of its subsystems, is available without charge to state and local agencies through the appropriate EPA Regional Office. EPA encourages the use of CDHS by state and local agencies and conducts a project to provide installation assistance and detailed system orientation training when requested by state and local agencies. System maintenance and further development is scheduled for at least the next several years.

CDHS will store, update, retrieve and maintain a history of information to support the following agency functions:

a. Emission Inventory Data Base Maintenance
b. Air Quality Data Base Maintenance
c. Scheduled Activity Monitoring
d. Permit System Operation and Enforcement
e. Compliance System Operation and Enforcement
f. Inspection System Operation
g. Complaint System Operation.

CDHS supports these activities by a combination of independent subsystems: the Emissions Inventory/Permits and Registration Subsystem (EIS/P&R), the Enforcement Management Subsystem (EMS), and the Air Quality Data Handling Subsystem (AQDHS-II). A conceptualization of the system is shown in Figure 1.2-2.

The Emissions Inventory Subsystem, with its extension to handle Permits and Registrations, (EIS/P&R), provides a means for monitoring point and area source engineering and emissions data. It has special capabilities for recording permit data and can handle narrative information such as rules and regulations. EIS/P&R also provides means for handling special data of local importance while maintaining full compatibility (for reporting purposes) with Federal requirements. EIS/P&R is especially significant since it can be used to support such agency activities as permit control, source inventory, legal actions, and the monitoring and recording of enforcement and inspection activities.

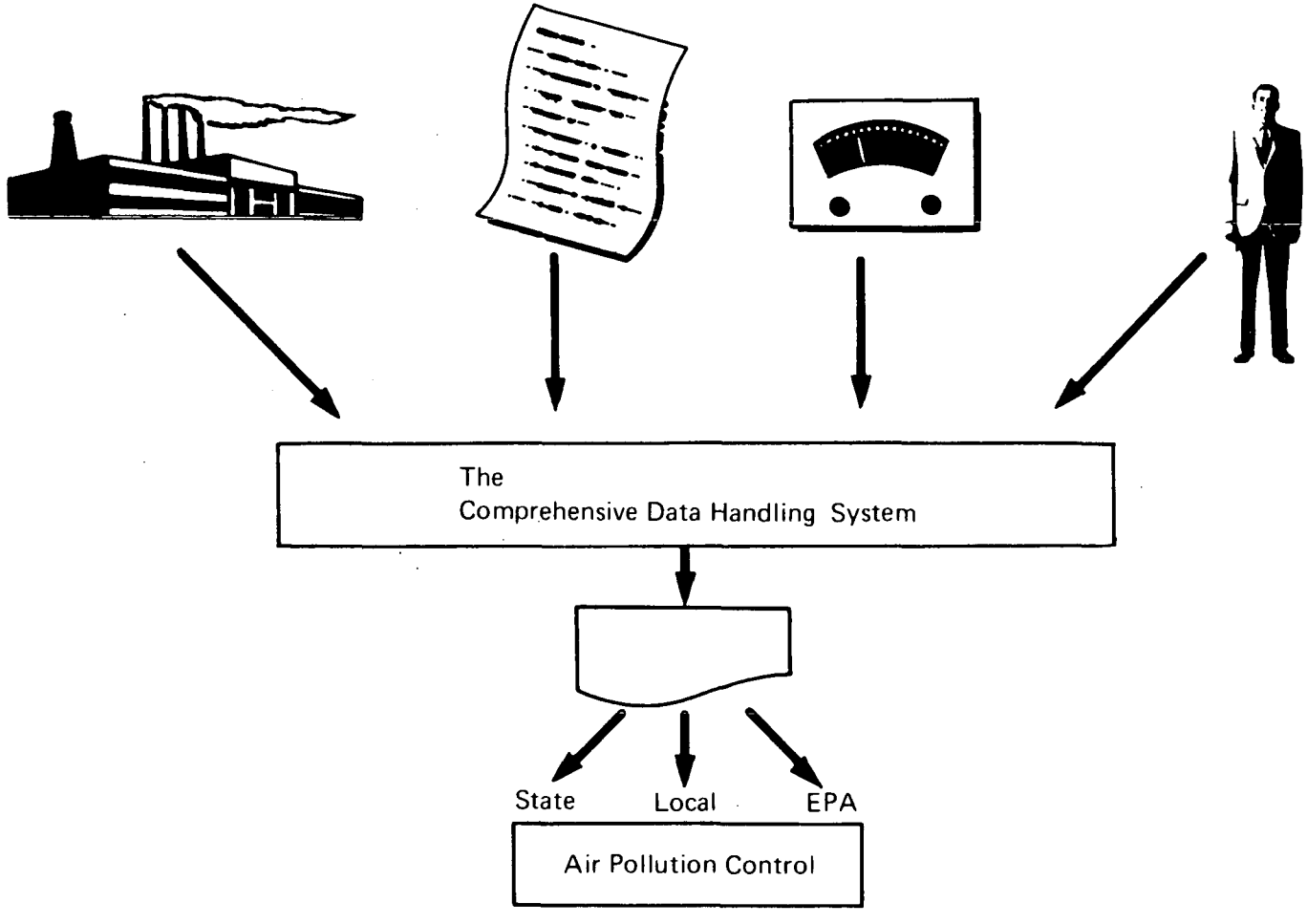

Figure 1.2-2. CDHS Concept

The Enforcement Management Subsystem (EMS) provides the means for recording and monitoring the many scheduled activities needed for air pollution control. EMS assists management by reporting scheduled actions (and providing individual Action Cards) as well as generating form letters to organize activities such as permit applications, meetings with industry, and enforcement notices for which appropriate standard letters can be locally devised.

The Air Quality Data Handling Subsystem (AQDHS-II), as its name implies, records and processes air quality data. In addition it provides a means for monitoring special site information such as sensor descriptions. AQDHS-II contains a full set of statistical processing routines designed to reduce air quality data into a meaningful form in presentation. The reports generated by AQDHS-II assist agencies in monitoring the progress and effectiveness of air pollution control. AQDHS-II simplifies reporting requirements by automatically generating appropriate Federal air quality reports.

The CDHS subsystems are straightforward and simple by design. They derive their power from the ease of execution of three basic capabilities: data storage, data retrieval, and report generation. The general flow of information through the system is shown in Figure 1.2-2.

The greatest utility of CDHS comes from its capability to selectively retrieve and display in various formats the contents of one or more data elements in a source record. Using this capability users can answer such requests as, "Show me all data on firms whose principal product is steel." or "Show me all data for plants who are within a square bounded by four UTM coordinates and whose stacks are less than 100 feet and emit SO2." or "How many facilities have stacks taller than 100 feet?" or "What is the distribution of emissions in Montgomery County by major categories of pollutant sources?" using EIS/P&R. AQDHS-II supplies information supporting answers to such questions as: "Are there any centers of high pollution concentration in Western City?," "Has the particulate standard been exceeded at the City Hall site this year?," "Is there any periodicity to the concentration of hydrocarbon at the corner of 5th Avenue and Vine?," "Has the

sulfur oxide concentrations in Carroll County decreased in the last two years?" or "What are the monthly mean carbon monoxide concentrations along Interstate 70S?" EMS on the other hand helps agencies keep track of scheduled events and supplies answers to such questions as: "What inspectors will be out next week and where will they go?," "Do we have any court cases scheduled in September?," "Are there any sources who are delinquent in requesting permits?," or "Do we have any special surveillance activities next month?"

The fact that separate computer program facilities are used to extract data from files and to list the data makes the system generally amenable to local extension with specialized programs. These locally developed programs resolve special local problems and analyze particular atmospheric conditions.

1.3 EMISSION INVENTORY/PERMIT AND REGISTRATION SUBSYSTEM (EIS/P&R)

The Emission Inventory/Permit and Registration Subsystem (EIS/P&R) is designed to help resolve agency problems associated with maintaining and controlling sources of pollution (Figure 1.3-1). The subsystem provides the basic capabilities needed to record data describing sources of pollution in terms directly applicable to control and enforcement rules and regulations. The recorded data includes engineering data describing the ultimate source of pollution as well as the flow through the system to the atmosphere. The subsystem also accommodates descriptive data related to permits or registrations pertaining to each source.

Since the emission inventory identifies sources of pollution which are the ultimate objectives of control regulations and enforcement actions, the subsystem has been designed to support not only the basic functions but also a number of related functions. Prime among these are:

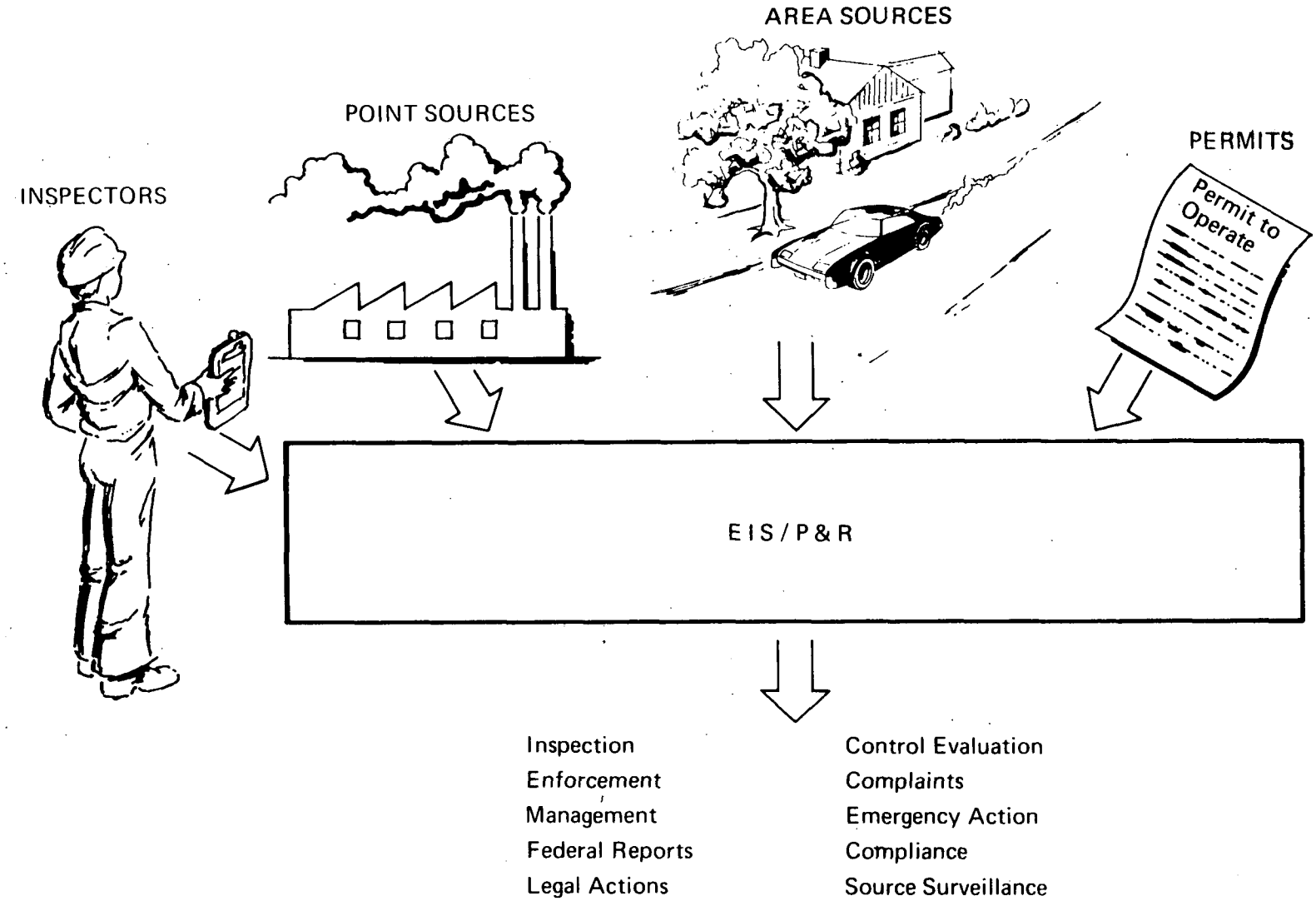

Figure 1.3-1. EIS/P&R Concept

a. Inspections
b. Complaints
c. Legal compliance
d. Source surveillance.

Inspections are supported by the subsystem and, in turn, support the subsystem. EIS/P&R provides a readily accessible medium for recording the results of inspections. The results may appear in the system as verified standard entries or as narrative accounts. The subsystem is supported in the sense that the content of the files are given additional validity and, thus, greater utility because of the verifications achieved during inspections. EIS/P&R can be easily used to verify or identify sources against which complaints have been made. For example, a geographical search of the file could be used to find the possible source or sources that may have given rise to a complaint. Conversely, the data on file may eliminate sources from consideration. Again the flexibility of the system can be used to record the complaint in narrative form once the appropriate source has been isolated.

Legal actions are also supported by EIS/P&R which provides a means for recording facts that could be introduced as evidence. Such facts could include as a standard part of the recorded data such items as measured and allowable emissions, pertinent permit information, and appropriate engineering data. Additional material can be included in EIS/P&R to cover such facts as registered complaints, source surveillance reports, and cross references to admissable legal documents. This information can be related to court calendars either by schedules incorporated in the special fields of the EIS/P&R record or by manual cross reference to the schedules maintained by the EMS. All of this capability provides greater management control as well as the means for presenting pending actions for management review.

EIS/P&R supplements EMS and the Compliance Data System by providing a direct association between sources, permits and compliance schedules. Compliance schedules can be inserted as a part of the emission inventory record by using

the specialized storage and retrieval capabilities of the subsystem. Thus, all aspects related to the degree of compliance of a source can be recorded within a single record of the source thereby permitting the retrieval and reporting of all pertinent data and compliance data. Such reports can be expected to be useful in establishing inspection schedules, monitoring compliance, highlighting problems, and guiding enforcement actions.

Source surveillance information consisting of emission measurements (e.g., stack tests) and special air quality measurements can also be recorded as part of a source record. This information can then be readily available for review when considering source enforcement actions of various types. Of course the interface with EMS for scheduling purposes helps organize the acquisition and use of valuable data.

The EIS/P&R consists of a number of programmed packages that perform the following functions:

a. Create an Emissions data base from NEDS formatted data

b. Create an Emissions/Permit data base from EIS/P&R formatted data

c. Add to, change, or delete data in the file

d. Selectively retrieve data from the file

e. Calculate estimated emission from stored emission factors

f. Generate detail emission reports

g. Generate detail permit reports

h. Generate summary emission reports

i. Generate Federal (NEDS) reports.

There are two types of input to the subsystem. They are cards punched according to the National Emissions Data System (NEDS) format, and cards punched according to the EIS/P&R format.

Both types are used for creating the original file, while the EIS/P&R formats are generally used for adding to, changing, or deleting data from the file.

Major subsystem functions are file maintenance, internal data edits, data retrieval and report generation. File maintenance includes those functions that permit an up-to-date file to be maintained. Three types of file maintenance transaction are allowed--add, change, and delete. The subsystems can:

a. Add new source records
b. Add data to existing source records
c. Change a data value in a source record
d. Delete a source record
e. Delete a portion of a source record.

In ordinary use, most data values in a record can be used in a query statement to develop simple or compound logical inquiries about the data contents of a source record. Such queries can be used to retrieve from one to all source records, depending only on the type of source information desired.

Information retrieved from the master file can be presented in three ways, namely:

a. Detailed reports
b. Summary reports
c. NEDS reports.

Detailed reports present all data about the source. Summary reports present a summary of emissions from several sources. NEDS reports are simply the source data presented in NEDS input formats.

1.4 ENFORCEMENT MANAGEMENT SUBSYSTEM (EMS)

The Enforcement Management Subsystem was developed in 1972 to provide state and local air pollution agencies with an effective means of controlling enforcement activities (Figure 1.4-1). The subsystem emphasizes management control of the enforcement function and establishes standardized methods of handling data. Presently, a number of changes and additions to improve operation of the system are underway, including a mechanized link with the Compliance Data System (CDS) which will permit state agencies to meet many EPA reporting requirements with a minimum of additional effort. Another improvement will permit better selection of output data.

Installation of the system at an air pollution agency should improve the efficiency of the enforcement function and increase management's ability to control these operations. A number of states have been operating the system for varying periods of time. In addition, four states are in the process of installing the EMS.

All enforcement staff members of an air pollution agency should benefit from the use of the EMS. Agency management will obtain improved operational control over the enforcement function. They will be able to prepare many periodic reports more efficiently than before.

Engineers, inspectors and other staff members should also benefit through easy access to specific source information and improved control and visibility over their routine activities.

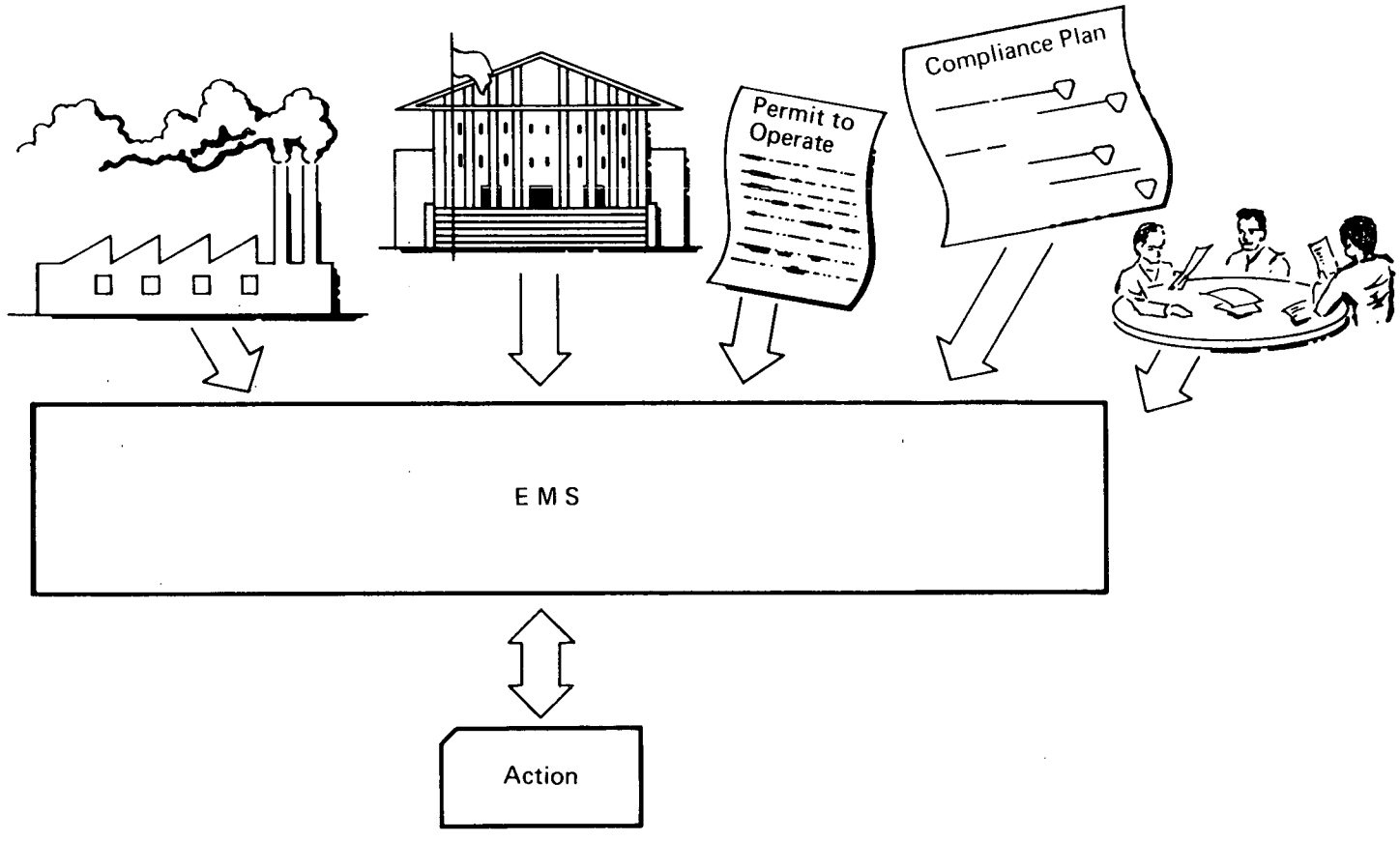

Figure 1.4-1. EMS Concept

EMS is concerned directly with enforcement activities such as inspections, variances, permits, and any action which the agency undertakes. It is concerned with both historical and future actions, and makes future workloads highly visible to management and section heads through the Future Schedule Summary Report. It improves the agency's ability to evaluate the upcoming workload in terms of available personnel, and permits scheduling changes in activities. For example, at any time, an agency can determine the status of its activities with respect to a given source. These activities can include a history of past contact, current activities, and other actions planned for years into the future, such as a compliance schedule or a routine every-two-years checkback inspection. This whole sequence of activity is available for inspection by staff members and management, showing what has been done, what is due, and what is planned.

The availability of this record of all present and planned contact between agency personnel and a specific source generally represents a significant improvement over most state files.

One member of a state agency described the EMS as a sophisticated tickler file. That description is largely correct. The ability to store future schedules, or to plan inspections or legally mandated increments of progress, and to present these activities sorted out by time or by staff group, has proven to be useful. Middle management should be able to look ahead and evaluate the upcoming actions that his group will be expected to perform within a given time period.

Looking at the realities of daily agency activity, it becomes apparent that technical personnel spend significant time digging out data to support external requests for informatives. Answers to citizens, to the legislature, to other state officials, to the EPA, and to others are needed. An up-to-date EMS data base permits the staff to leave much of the "digging out" of the information to the system. For example, one report permits a staff member to prepare a report listing all agency activity of a designated type, such as inspections, and allows the staff member to select the time period covered, even including future periods if desired.

EMS is a dynamic system, not one which simply maintains a static data base for reporting purposes. It features the Action Card. The Action Card is a "turn-around document," in that it is produced by the system and is also designated to be reentered with new information. For example, an increment of progress is scheduled for a certain date. The system prepares one of these cards identifying the upcoming increment of progress. This card goes to the person who performs the needed action. Next, the person who is responsible for the activity can write the results right on the card. If it is a routine action, the staff member need circle only one or two numbers. However, he has the capability of adding English language comments or even scheduling a reinspection or other activity by means of the card. When he has done this, these results are punched into the card itself and it is fed back into the system which updates its own records. If an additional action has been called for, the system will produce a second card at the appropriate time and the process is repeated.

A new feature of the EMS is now under development. This will provide a direct link to the national Compliance Data System. We believe it will substantially reduce the effort required by state agencies to meet federally mandated reporting requirements of increments of progress. The link works this way: at reporting time, the operator from the state agency who is running the system schedules a special computer program to be run, using his EMS data base. The special "converter" computer program extracts appropriate information from the state's EMS base, reformats it into CDS input format, and punches out a deck of cards representing that information. State personnel will have an opportunity to review the data being extracted through the printed report to ensure that all data which they need to pass on is included and that it is completed correctly.

The reports capable of being produced by the system are as follows:

a. Action card
b. Action Summary Report
c. Edit and Update Report

d. Future Schedule Summary Report
e. Geographic Locator Report
f. Source Action Summary Report
g. Source Registration.

Additional retrieval capabilities are now being implemented and these will be available within a short period of time. It will be possible, however, to install the existing system and directly go to the retrieval programming without substantial effort.

Of particular interest to heads of State agencies, is the basic capability of EMS to provide individual service to local agencies. For example, the centralized operation of EMS by the state provides a respository for the information pertaining to the local agencies operations. Periodic generation of reports for distribution provides a tailored schedule monitoring system for each local agency. In this way both state and local agencies can contribute data to the system and utilize its outputs.

1.5 THE AIR QUALITY DATA HANDLING SUBSYSTEM (AQDHS-II)

The AQDHS-II is, as a subsystem, similar to all other components of CDHS and is designed to assist agencies resolve problems dependent upon the distribution and trends of air quality. It has basic design functions which are supplemented by built-in flexibility to support related enforcement functions (Figure 1.5-1). Its basic functions include data storage, update and retrieval capabilities, as well as a broad range of statistical processing all of which support air quality monitoring. Related enforcement functions which are supported include:

a. Control strategy evaluation and development
b. Recording source surveillance measurements
c. Generation of SAROAD reports

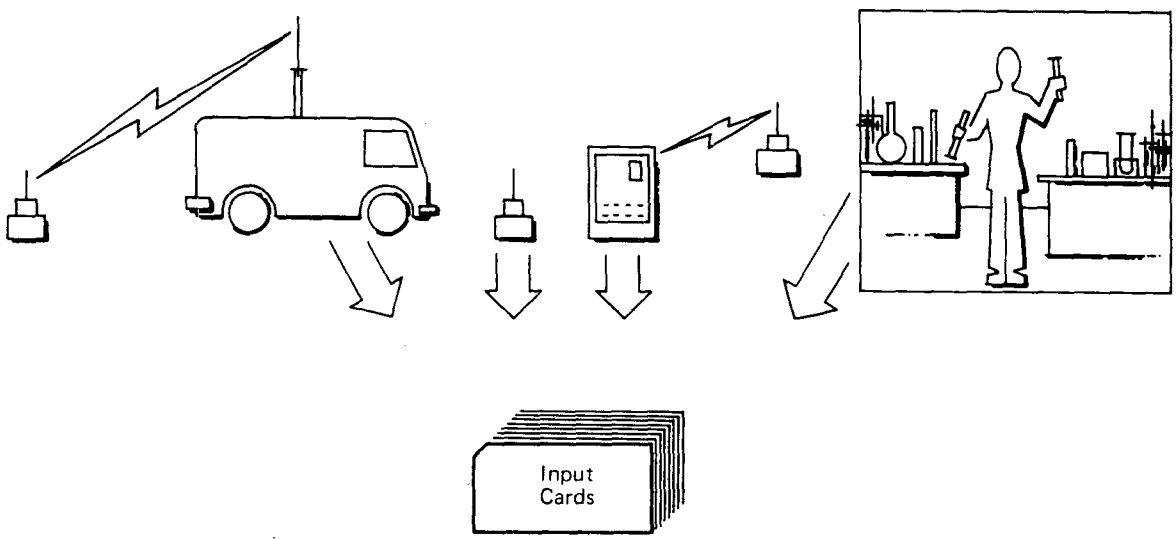

Figure 1.5-1. AQDHS-II Concept

d. Providing data for modeling
e. Providing data for legal actions.

AQDHS-II stores and maintains measurements of air quality which are input to the subsystem in either of two formats. The SAROAD format can be used and is recommended as a source of data for the initial generation of a file. The AQDHS-II format provides a full range of input capability including adding, changing or deleting information in the file.

The more significant processing functions of the subsystem are:

a. Data checking
b. Statistical calculations
c. Automatic generation of Federal reports
d. Report generation.

Input data is checked for validity in those instances where such action is possible. These checks are supplemented by normal data processing checks to ensure that data is entered in the proper form. For example, numeric fields are checked to see that only numbers have been key punched. Statistical calculations include both arithmetic and geometric means and standard deviations, maximum and minimum values, sliding averages for selectable intervals, and a broad range of percentile values. A special function provides for flags attached to air quality data to indicate whether the data has been reported to the Storage and Retrieval of Aerometric Data (SAROAD) system and whether the data has been changed since last updated to SAROAD. This is valuable in reducing resource needs when generating the quarterly SAROAD input. In addition to the SAROAD input, AQDHS-II generates three major reports for routine agency use. There is a detail report of raw data showing file contents on a site-by-site basis. There is a sliding average report showing a sliding average of a selected pollutant for the usual intervals taken over a twenty-four hour

period. Finally, there is a statistical analysis report which contains the results of the statistical processing including the three highest values in the sample set. All reports are titled with descriptive headings.

1.6 IMPLEMENTATION CONSIDERATION

All subsystems within CDHS have been designed using simplicity as the major principle. This has resulted in subsystems which can be operated with a minimum of modification on almost any general purpose computer. All programs are written in either ANS COBOL or ANS FORTRAN. All input transactions have standard punched card formats. All system files are organized for either magnetic tapes or disk storage.

CDHS implementation activities will take various amounts of time depending on local circumstances. However, a general schedule can be made which should be reasonably applicable. Such a schedule is shown (Figure 1.6-1) which allows somewhat less than five months to accomplish necessary actions to install one or more CDHS subsystems.

All components of the system were developed so that they can be installed and used by any agency having access to a minimally equipped computer. The basic equipment requirements are:

a. Three tape drives (or disk equivalent)
b. 100K Bytes (or equivalent) of CPU memory
c. Card Reader
d. Printer
e. Keypunch

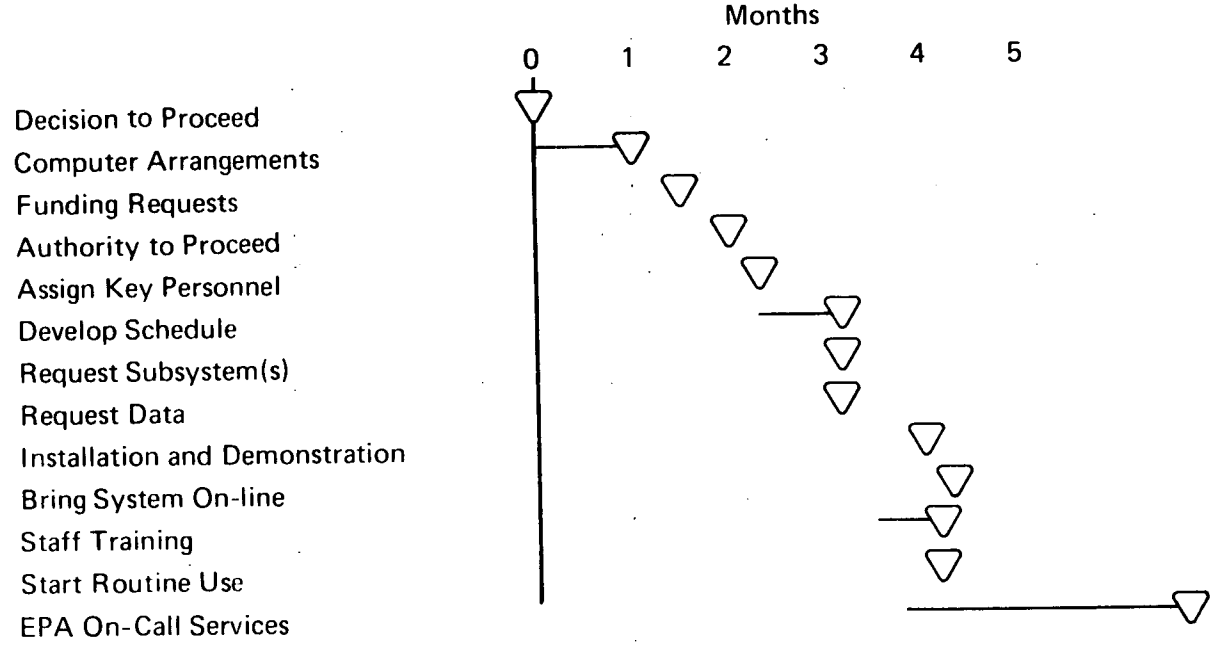

Figure 1.6-1. Sample Implementation Schedule

The CDHS components can, with little modification, be operated on any computer with the capability to:

a. Compile ANS COBOL and ANS FORTRAN

b. Sort files

c. Accept single sequential inputs and single sequential output files

d. Print reports and punch cards.

All operations are in batch mode; that is, a deck is submitted for a computer run and at some later time the results are returned to the user. Use of the system is scheduled and processed, and output is returned to the requestor according to local computer center procedures.

Each subsystem builds and maintains its own file from well defined input cards which carry a code directing the computer to add, change, or delete data from its files. Data in the files are automatically maintained in the most useful order by reference to identification data in each record.

All system master files are ordered sequentially and may be stored on magnetic tape or disk. Programs have only one master file input and a maximum of one updated master file output. Intermediate "answer" files may be retained so that several different reports may be prepared from the same source records.

A file entry can be made as soon as identifying information (plant information for EIS/P&R and EMS and site information for AQDHS-II) is available. Additional data may be added to or changed in the file by creating a change transaction with the proper identifying fields and providing the new or changed data. It is, however, more efficient to collect changes and additions to the file and perform file update activity periodically.

The arrangement of individual files of data have been structured to accommodate a large data base which is organized for ease of maintenance and use. This data base will logically include information provided from many diverse sources. Although this would normally create significant duplication of data within the individual files, organization in the hierarchical fashion serves to eliminate much of this duplication and consequently enhances the effective file storage capability.

Programs to build and maintain these hierarchical files are available as standard elements of CDHS. These programs offer the data base user a standard symbolic interface which permits him to manage and coordinate individual records in the system without developing an in-depth knowledge of the actual file organization.

The ability to group files and data from many diverse sources into one family of logical structures, coupled with the ability to provide conformity between these structures, makes this file organization scheme a highly attractive one for the air quality management system.

The CDHS master files are arranged using simple hierarchical structures. This means that in describing a particular phenomenon, the data which is characteristic of all aspects is presented first. Data which relate only to small facets are associated with the general descriptive information by being filed behind the header information. For example, the emissions data structure places process and engineering descriptive information subordinate to a point source which, in turn, is subordinate to a plant.

This type of filing is similar to manual filing systems in which a header ledger card is followed by trailer cards which describe particular events. In one access to the file all generic and detail descriptions can be retrieved.

In EIS/P&R, the hierarchical structure is plant, point, process, and permits. In AQDHS-II, the major control is pollutant and location while subset records describe measurements taken through time. The EMS describes a particular pollution source at its highest level of control. Detail data describes action taken to bring the emissions under control.

1.7 PERSONNEL SUPPORT

Two types of personnel support are required. The assistance of a computer specialist, a System Coordinator, is necessary to provide those services which deal directly with computer equipment, performing tasks such as establishing local machine procedures, allocating direct access space providing file backup support and other tasks requiring familiarity with computer procedures. This systems support is not normally a full time task and is not involved with the day-to-day use of the contents of files. Other uses for systems support are to provide assistance as required for coding special COBOL language retrievals, and restoring the system in the event of a failure. The requirement for programmer personnel has been minimized.

Data support personnel, headed by a Data Base Coordinator, are responsible for data oriented operations including securing input data, coding data input forms, coding retrieval statements, keeping track of the latest version of files, scheduling computer runs, and similar activities designed to ensure efficient use of the system. These people who must rely on the CDHS User's Guides supplemented by locally prepared instructions, may have little or no computer background training. The primary responsibilities of data support personnel are to maintain complete, current data in the files and to use CDHS facilities to prepare reports satisfying management requests.

The system requires no specialized training for CDHS operations personnel. However, local procedures will be required to organize the use of the subsystems. These deal with standard procedural decisions such as how many

generations of each file should be kept, how to minimize use of storage devices, when to update the files, and how to maintain data integrity.

A user should designate an individual whose responsibility includes the collection, storage and organization of data; that is, a Data Base Coordinator (DBC) is essential to the subsystem. The DBC sees that inspectors, engineers, and others responsible for collecting data present their data for inclusion in the EIS/P&R data base. He is also responsible for coordinating such activities as:

a. Assignment of identification numbers
b. Assignment of storage devices
c. Organizing and scheduling subsystem use
d. Generating periodic reports.

For small to medium size data files with moderate usage, one individual is probably more than sufficient. However, as the frequency of use increases and the files expand, the function may exceed the capabilities of one man. In any event he must be supported by keypunching which can be provided in many ways including:

a. A keypunch section
b. Keypunching by data collectors
c. Use of centralized keypunch services
d. Contracted services.

1.8 RESOURCES

The cost or resources required to operate CDHS vary in proportion to the size of the data base frequency of update and retrieval, the response time requirement, and the rates charged by the local computer center. For most agencies, CDHS

will use only a small proportion of a computer's daily schedule. Such usage expressed as CPU time may be on the order of one half-hour a week per subsystem for regular updates by a small agency with relatively infrequent need to extract data from the files. Larger agencies with many sources and sensor sites could be expected to use the system more frequently, and for planning purposes, should expect to need 1-2 hours per week per subsystem.

1.9 SYSTEM IMPLEMENTATION SCHEDULE

Successful installation of CDHS or any one of its subsystems is more assured if an agency carefully considers several scheduling activities and decisions. The first decision, of course, is that which affirms the need for one or more CDHS subsystems. Subsequently, the following actions should be taken:

a. Arrangements should be made for appropriate access to a suitable computer. Normally this is achieved by negotiation with the head of the state (or county, or city) central data processing center. These negotiations should consider both the time and frequency of computer use as well as the mechanisms and costs for obtaining access.

b. Based upon cost estimates developed during these negotiations, funding authorization should be established.

c. At this point, key personnel should be identified and given the responsibility for the implementation plan. Key personnel consist, at a minimum, of the Data Base Coordinator, and the System Coordinator. In all agencies a manager should be identified and given the appropriate managerial responsibility. The Data Base Coordinator should be assigned from the agency staff. The System Coordinator may be assigned from the agency staff, but may be, by special arrangement, a part of the data processing center staff.

d. The responsible key personnel should now schedule required actions consisting of at least the following elements:

1. Request for subsystem(s) from EPA through the Regional Office

2. Request for initial data base from EPA, NADB through the Regional Office

3. Subsystem(s) installation and demonstration

4. Bringing the subsystem(s) on-line by creating the initial data base

5. Provision of subsystem(s) training for appropriate agency staff in the selected subsystem(s)

6. Inauguration of routine use.

The accomplishment of this action will lead to an easy and well organized installation of CDHS.

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.	2.	3. RECIPIENT'S ACCESSION NO.
EPA-450/3-76-011		

4. TITLE AND SUBTITLE	5. REPORT DATE
CDHS Executive Summary Report	March 1976
	6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S)	8. PERFORMING ORGANIZATION REPORT NO.
The IBM Corporation - Federal Systems Division Gaithersburg, Maryland 20760	

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT NO.
The IBM Corporation - Federal Systems Division Gaithersburg, Maryland 20760	11. CONTRACT/GRANT NO. 68-02-2054

12. SPONSORING AGENCY NAME AND ADDRESS	13. TYPE OF REPORT AND PERIOD COVERED
U. S. Environmental Protection Agency Office of Air and Waste Management Office of Air Quality Planning and Standards Research Triangle Park, NC 27711	Final 14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

16. ABSTRACT

Management of air quality requires constant surveillance of air quality and careful recording of all significant pollution sources.

Measurements are required to trace progress of specific emission control attempts and of the general state of air quality. Detailed analysis is required to predict future air pollution conditions and to determine the contributory effect of any specific emission. These considerations must, in turn, be related to Air Pollution Control Implementation Plans to ensure that enforcement actions are achieving desired goals. These and the many unmentioned tasks in air quality control add up to a massive data handling requirement.

The CDHS was designed to provide using States (via the EPA Regional Offices) with a universal and compatible system for managing air data. It will also provide an efficient mechanism for timely, complete and accurate input of air quality data to the National Systems.

This report contains a brief description of the CDHS system and is intended to provide an executive overview of the system.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Computer Programs Computer Software Data Processing Air Pollution Data Handling Computer Systems Programs	CDHS AQDHS-II EIS/P&R EMS SAROAD NEDS CDS	

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES
Release unlimited	Unclassified	35
	20. SECURITY CLASS *(This page)* Unclassified	22. PRICE

EPA Form 2220-1 (9-73)

INSTRUCTIONS

1. **REPORT NUMBER**
Insert the EPA report number as it appears on the cover of the publication.

2. **LEAVE BLANK**

3. **RECIPIENTS ACCESSION NUMBER**
Reserved for use by each report recipient.

4. **TITLE AND SUBTITLE**
Title should indicate clearly and briefly the subject coverage of the report, and be displayed prominently. Set subtitle, if used, in smaller type or otherwise subordinate it to main title. When a report is prepared in more than one volume, repeat the primary title, add volume number and include subtitle for the specific title.

5. **REPORT DATE**
Each report shall carry a date indicating at least month and year. Indicate the basis on which it was selected *(e.g., date of issue, date of approval, date of preparation, etc.)*.

6. **PERFORMING ORGANIZATION CODE**
Leave blank.

7. **AUTHOR(S)**
Give name(s) in conventional order *(John R. Doe, J. Robert Doe, etc.)*. List author's affiliation if it differs from the performing organization.

8. **PERFORMING ORGANIZATION REPORT NUMBER**
Insert if performing organization wishes to assign this number.

9. **PERFORMING ORGANIZATION NAME AND ADDRESS**
Give name, street, city, state, and ZIP code. List no more than two levels of an organizational hirearchy.

10. **PROGRAM ELEMENT NUMBER**
Use the program element number under which the report was prepared. Subordinate numbers may be included in parentheses.

11. **CONTRACT/GRANT NUMBER**
Insert contract or grant number under which report was prepared.

12. **SPONSORING AGENCY NAME AND ADDRESS**
Include ZIP code.

13. **TYPE OF REPORT AND PERIOD COVERED**
Indicate interim final, etc., and if applicable, dates covered.

14. **SPONSORING AGENCY CODE**
Leave blank.

15. **SUPPLEMENTARY NOTES**
Enter information not included elsewhere but useful, such as: Prepared in cooperation with, Translation of, Presented at conference of, To be published in, Supersedes, Supplements, etc.

16. **ABSTRACT**
Include a brief *(200 words or less)* factual summary of the most significant information contained in the report. If the report contains a significant bibliography or literature survey, mention it here.

17. **KEY WORDS AND DOCUMENT ANALYSIS**
(a) DESCRIPTORS - Select from the Thesaurus of Engineering and Scientific Terms the proper authorized terms that identify the major concept of the research and are sufficiently specific and precise to be used as index entries for cataloging.

(b) IDENTIFIERS AND OPEN-ENDED TERMS - Use identifiers for project names, code names, equipment designators, etc. Use open-ended terms written in descriptor form for those subjects for which no descriptor exists.

(c) COSATI FIELD GROUP - Field and group assignments are to be taken from the 1965 COSATI Subject Category List. Since the majority of documents are multidisciplinary in nature, the Primary Field/Group assignment(s) will be specific discipline, area of human endeavor, or type of physical object. The application(s) will be cross-referenced with secondary Field/Group assignments that will follow the primary posting(s).

18. **DISTRIBUTION STATEMENT**
Denote releasability to the public or limitation for reasons other than security for example "Release Unlimited." Cite any availability to the public, with address and price.

19. & 20. **SECURITY CLASSIFICATION**
DO NOT submit classified reports to the National Technical Information service.

21. **NUMBER OF PAGES**
Insert the total number of pages, including this one and unnumbered pages, but exclude distribution list, if any.

22. **PRICE**
Insert the price set by the National Technical Information Service or the Government Printing Office, if known.