ISWS/CIR-174/90
Circular 174
STATE OF ILLINOIS
DEPARTMENT OF ENERGY AND NATURAL RESOURCES

Automated Data Systems for Ground-Water Information

by
SUSAN C. SCHOCK, KAY MUMM,
TRUDY K. DAHL, and SUSIE DODD

ILLINOIS STATE WATER SURVEY
CHAMPAIGN
1990

CIRCULAR 174

Automated Data Systems for Ground-Water Information

by
SUSAN C. SCHOCK, KAY MUMM,
TRUDY K. DAHL, and SUSIE DODD

with contributions from
KENNETH J. HLINKA, KRISTOPHER K. KLINDWORTH,
and DOROTHY M. WOLLER

Title: Automated Data Systems for Ground-Water Information.

Abstract: Systematic collection, storage, and organization of ground-water data are necessary as the volume and variety of records increase, so that the data will be accessible and useful. When an automated system is being planned, it is important to consider what will go into that system, the uses of the system, and the types of data outputs that will be provided. This publication discusses the types of ground-waterrecords maintained at the Illinois State Water Survey, and the system developed for filing the records. It then reports the results of a survey of state agencies in Michigan, Indiana, Kentucky, and Wisconsin regarding the types of ground-water data in their state database systems, financial support available for automated information systems, system features and characteristics, data sources, methods for making updates and corrections, and record identification systems. Recommendations are offered to database and computer groups regarding documentation, assignment of record identification numbers, archival of automated information systems. Geographic Information System (GIS) capabilities of information systems, and responsibilities of information managers.

Reference: Schock, Susan C, Kay Mumm, Trudy K. Dahl, and Susie Dodd. Automated Data Systems for Ground-Water Information. Illinois State Water Survey, Champaign, Circular 174, 1990.

Indexing Terms: Automation, data processing, data storage and retrieval, ground-water data, Illinois, Indiana, information retrieval, Kentucky, Michigan, water wells, well data, Wisconsin.

STATE OF ILLINOIS
HON. JAMES R. THOMPSON. Governor

DEPARTMENT OF ENERGY AND NATURAL RESOURCES
Karen A. Witter, M.S., Director

STATE WATER SURVEY DIVISION
RICHARD G. SEMONIN, Chief

2204 GRIFFITH DRIVE
CHAMPAIGN, ILLINOIS 61820-7495

1990

Printed by authority of the Slate of Illinois

(11-90-200)

Funds derived from grants and contracts administered by the University of Illinois were used to produce this report.

CONTENTS

ABBREVIATIONS

ENR - Illinois Department of Energy and Natural Resources
GIS - Geographic Information System
IDEM - Indiana Department of Environmental Management
IDMM - Illinois Department of Mines and Minerals
IDNR - Indiana Department of Natural Resources
IDPH - Illinois Department of Public Health
IEPA - Illinois Environmental Protection Agency
IGPA - Illinois Groundwater Protection Act
ISGS - Illinois State Geological Survey
ISWS - Illinois State Water Survey
KDEP - Kentucky Department of Environmental Protection
KDNR - Kentucky Department of Natural Resources
MDNR - Michigan Department of Natural Resources
ODNR - Ohio Department of Natural Resources
STORET - Storage Retrieval System (USEPA)
TRS - Township, range, section
USEPA - United States Environmental Protection Agency
USGS - United States Geological Survey
WDNR - Wisconsin Department of Natural Resources

AUTOMATED DATA SYSTEMS FOR GROUND-WATER INFORMATION

by
Susan C. Schock, Kay Mumm,
Trudy K. Dahl, and Susie Dodd

with contributions from
Kenneth J. Hlinka, Kristopher K. Klindworth,
and Dorothy M. Woller

INTRODUCTION

Data on subsurface geology, water quantity, and water quality are often collected in ground-water investigations related to wells and water supplies. Systematic collection, storage, and organization of these data are necessary as the volume and variety of records increase, so that the data will be accessible and useful. When an automated system is being planned, it is important to consider what will go into that system, the uses of the system, and the types of data outputs that will be provided. Software should be chosen to accommodate the data and the users' needs, and the hardware that runs the software should be chosen on the basis of these considerations.

Information managers must be aware of user needs and the constraints on the data. For example, when users want data for a specific location for mapping purposes and the information includes only addresses, the user may have to determine where on a 5-mile route an address is located. In anticipation of such a need, information managers may add a data element to convert addresses to points. To be successful, managers must look to the future. They must be sensitive to changes in user needs and to the constantly changing technology, such as laser disks for information storage and computer hardware advances for speed and efficiency of data retrieval.

Ground-Water Records at the Illinois State Water Survey

The number of ground-water records on file at the Illinois State Water Survey (ISWS) has grown to nearly 3.5 million, including approximately 300,000 paper copies of water-well logs. The accumulation of these records began almost a century ago when the ISWS was established in 1895. The variety of records and the number of individual pieces of paper require that staff persons be assigned to maintain the information in a usable form.

There has been a gradual evolution in the processes used to maintain these records. The most revolutionary change in record keeping has been the introduction of computers to information management. It is extremely important that automation of information be carried out with goals, plans, and flexibility. Computer technology changes quickly, and although some of the improvements are extremely useful to

information managers, the costs and effort in upgrading can make it difficult to stay current. The manager must consider whether a change is beneficial enough to warrant the work.

Among the early records collected at the ISWS was a set of well inventory forms, which were used from 1895 until the 1930s. In the early 1930s, a new type of well inventory form was developed and used for a door-to-door inventory of wells, which was conducted to locate as many water wells in Illinois as possible. Information was collected about the origins of those wells and perceptions of the well users about the water quality in them. The records, which categorized wells according to township, range, and section (TRS) location, were filed at the ISWS.

Beginning in the early 1900s, information pertaining to subsurface geology was collected in the form of water-well logs. These records are first submitted to various agencies in compliance with regulations, and the agencies then send them to the ISWS. For example, logs are submitted to the Illinois Department of Public Health (IDPH) and Illinois Department of Mines and Minerals (IDMM), and copies are sent to the ISWS.

Water quality records from chemical laboratory analyses of ground-water samples were included in the early records. These records were added to the automated information system beginning in the 1970s. They come from both the ISWS laboratories and Illinois Environmental Protection Agency (IEPA) laboratories.

From those early beginnings until today, the ISWS has maintained paper files of information relevant to municipal, private, industrial, irrigation, and other types of wells in the state. These records include:

Water-well logs

Affidavits of well abandonment

Water-quality chemical analysis reports

Water-well inventory forms

Water-related correspondence

Special projects reports

Well test reports

Maps

The information collected and organized at the ISWS has many uses both in conducting research and in providing services. As the public, private businesses, and other agencies have become aware of the records at the ISWS, they have begun to rely upon the ISWS staff to maintain those data and to make them available for their use.

Information Filing System at the Illinois State Water Survey

The system developed for filing records of information at the ISWS was based on location. The federal government had established the TRS system as the surveyed standard for location in the Midwest, so this system was adopted to locate sites in the paper filing system at the ISWS. It was also logical that county be the first major subdivision for location, because it is larger than TRS. Using county as the primary

subdivision and TRS as the second subdivision allowed records to be accessed in several different ways, depending on the needs of the user.

Certain information was used regularly, and certain types of information were usually used together. Special projects such as ground-water availability studies used water-level data and well logs to determine subsurface conditions and to make predictions about potential yields from aquifers. A series of publications on public water supplies was based on data from well logs and inventory forms, as well as water usage data.

Within the county/TRS system, records were separated into general categories of information. To accommodate the most common uses of information, public water supply information was separated from all other records in the files. Information related to public water supplies has always been readily available. This is because public water supplies are regulated by both the state and federal governments, and they must supply regular reports to demonstrate their compliance with those regulations. Therefore the volume of public water supply information is greater than that for other types of wells. A second category of information was ground-water availability and related documents. Yet another was domestic supplies information.

Industrial information cannot always be easily accessed when it is organized only according to the TRS scheme. In some cases, many industries are located in close proximity to each other, and in other cases, only one industry is located in an area. An industry can own wells that are located in several TRS locations. Therefore, industrial information is indexed both by name and by location. This is an example of how comprehensive paper files need to be cross-referenced so that records can be accessed in several ways.

One way to deal with cross-referencing is to make copies of the records that are pertinent to more than one segment of the paper files and place a copy in each pertinent place. Another option is to add sheets to the files indicating the location of the information. Each of these methods requires adding large volumes of paper to the system. Any changes made to the information that has been duplicated might not be made to all pieces, leading to confusion and invalid information. To overcome part of the cross-referencing problem, the ISWS developed a separate index of nondomestic site names. This index cross-references locations and industry names, and indicates where the records are to be found. All of the files and indexes are in paper form.

As technology for automating data became available, segments of the information collected at the ISWS were placed into computer files on mainframe computers available through the University of Illinois computer system. The first automation of records was carried out through that system, which was expensive, difficult to use, and not easily accessible by groups of users.

The earliest data to be automated at the ISWS were the water quality chemical analyses reports from both public and private water supplies. Several years later the water use (pumpage) data for public water supplies and self-supplied industries were automated as part of a joint U.S. Geological Survey (USGS) and ISWS program. Several publication series on public water supplies (Bulletin 21, Bulletin 40, and Bulletin 60) were the foundation of a comprehensive database on public water supply wells. This database is still a fundamental part of the ISWS data and information

system. With the introduction of personal computers to general office staff, the automation of records expanded rapidly.

In 1987 the Illinois Groundwater Protection Act (IGPA) was enacted. It called for the organization of all ground-water information by the Illinois Department of Energy and Natural Resources (ENR) as part of a larger ground-water protection program. Since the ISWS is the repository for ground-water records for the state, it is appropriate that this database be developed by the ISWS. The ground-water information group of the ISWS has been formulating plans for this database for several years. The group plans to integrate subsurface geology data with data on ground-water quality, water levels, aquifer properties, and water use. Consideration is being given to statistical and graphic analysis tools to be used with the data, as well as to the Geographic Information System (GIS) capabilities of the system.

Survey of Other States

To expedite the database plan and to make more efficient use of available knowledge, the ISWS information group decided to determine what other states have learned in automating similar data for their agencies. Arrangements were made to visit several states during fall and winter 1988 and spring 1989. The ISWS group contacted the Michigan Department of Natural Resources (MDNR) - Groundwater Quality Division; the Indiana Department of Natural Resources (IDNR) - Division of Water; the Kentucky Department of Environmental Protection (KDEP) - Division of Water, Groundwater Section; and the Wisconsin Department of Natural Resources (WDNR) - Groundwater Management Division. These agencies were selected because their states are nearby and are similar to Illinois in background problems and environment. The group found that these other midwestern states were aware of the need for automated data, and that each was at a different stage in the paper-to-computer transition. This report presents the group's findings regarding the database systems in these states.

TYPES OF DATA IN STATE DATABASE SYSTEMS

Illinois Database System

The ISWS has been developing databases for various aspects of ground-water research and for access by its service and information groups for many years. Until recently these databases were developed separately as parts of projects, or for specific programs of information collection such as the water pumpage data collection and the Bulletin 21, Bulletin 40, and Bulletin 60 publication series. The ISWS has several independent databases. These include a water levels database; an aquifer properties database; a water withdrawals (pumpage) database; the Public, Industrial, Commercial Survey Well Inventory; and a ground-water quality database based primarily, but not exclusively, on the chemical analyses from public water supply wells monitored by the Illinois Environmental Protection Agency (IEPA).

The broad spectrum of types of data in the Illinois automated system makes it unique among the systems known to the ISWS group. Many of the states contacted

have more data than water-well logs in their system. Several include chemical analyses and point source contaminant data. However, no other state has the comprehensive set of data and the broad coverage of information related to ground water that exist at the ISWS. The uniqueness of these data and the possibilities that exist for their use make it imperative for the ISWS to continue to vigorously pursue the development of its automated information system.

Michigan Database System

Michigan is estimated to have from 600,000 to 800,000 water-well logs on file. These records are scattered among numerous agencies and private industries, primarily in paper files. The MDNR has decided that these records should be collected, verified, and entered into its automated information system. As of mid-1990, data from 30 counties were being entered onto the computer. Data from four of those counties had been verified and finalized. Data files from other offices of MDNR also are available on the system. These include a few GIS-type coverages of highway systems and hazardous waste sites, for example. Provision has been made in the system for many kinds of data to be entered at a later date.

Indiana Database System

Indiana has approximately 250,000 water-well logs on file in paper form. The volume increases by approximately 4,000 to 6,000 per year. Well logs are filed separately from all other records. A limited number of chemical analysis reports of ambient water quality are included. These analyses are not automated, and access to them is limited. Water quality records for public water supplies are collected and kept by the Indiana Department of Environmental Management (IDEM). IDNR and IDEM share paper copies of their information upon request.

Kentucky Database System

Kentucky has an estimated 250,000 wells according to the 1980 census. Since well permitting was begun in 1985, 6,700 records have been received. A growth rate of approximately 2,000 per year is expected. As of mid-1990, no plans had been discussed to inventory wells installed before 1985.

Wisconsin Database System

WDNR has approximately 450,000 well logs dating from 1940 through 1987 stored on microfiche. Well logs from 1988 to the present are not stored on microfiche. There are plans to merge the log files into the larger, automated system. Chemical analytical results from water samples collected by WDNR are stored in Storage Retrieval System (STORET) format. STORET is a USEPA data file system for analytical results. Unlike the situation in Illinois, the Wisconsin State Geological Survey is the repository for water-well logs. Copies of these logs are sent to the WDNR from the State Geological Survey.

Paper Files versus Automated Records

All states have more paper records on file than they have automated records. One reason for this is that many of the records are reports or analyses of other data, which are difficult to automate in a usable form. Another reason is that the volume and variety of records prohibit the automation of all of them. Table 1 shows the status of the paper versus automated records in Illinois and the states visited.

Table 1. Paper Files versus Automated Records

	IL	*MI*	*IN*	*KY*	*WI*
Approximate number of well logs	300K	600-800K	250K	6-7K	450K
Approximate number automated	160K	30 counties	15K	6-7K*	250K

** Kentucky began data entry of these records in about 1987*

IDENTIFIERS: THE FUNDAMENTAL DATA ELEMENT

In all systems that contain large volumes of data, the data must be identified in a unique and accessible way. Each state consulted had developed some method for record identification, and therefore for cross-referencing. Effective identification number systems can be very difficult to develop. The following subsections present observations about this concern that were made by the ISWS group.

Location as the Basis of an Identification System

Each state consulted by the ISWS information group had encountered a similar situation to the one experienced by the Illinois group. Over the years, the variety and volume of records pertaining to single wells became complex and often confusing. Frequently, several wells are located in close proximity, making it increasingly difficult to differentiate wells by location. This in turn makes it less effective to use location as a basis upon which to uniquely identify wells and the variety of records related to those wells.

For example, if a well has been identified by a number based on location, and the correct location of that well is later determined to be different, the temptation is to change the ID to reflect the corrected location, since that is the basis of the system. However, there is a risk that every record associated with the well to which the first ID was attached may not be changed. This not only causes a loss of information for the well, but may cause that information to be associated with a different well in the original location. This is the primary risk run whenever location is used as an ID basis. In Illinois, this situation has occurred with water quality data.

Kentucky's wells are located in the records by latitude, longitude, and topographic quadrangle. This is similar to the method employed by the USGS. Prior to 1985, Kentucky had no regulations for permitting of wells, and therefore well records were not kept. Because the Kentucky records are relatively few in number and their uses are few at this time, the need for more specific locations and more definitive identifiers has not become clear. The USGS, on the other hand, deals with information on a national basis. It is less likely that the USGS will need detailed information at the level needed by a state or local agency. Therefore the use of latitude-longitude may be appropriate to the USGS needs. Wisconsin stores both TRS and latitude-longitude information with its records when it is available. Michigan uses TRS for location.

Using location as a well identifier can be a problem in information systems that include many different types of data. Well-specific identifiers are not adequate for information systems that include records other than well records. For example, well logs are well-specific, but ground-water availability reports are not. An automated information system must relate all the available information for a well to that well. This requires that more general or regional reports be related to several wells. Well-specific identifiers cannot easily be used for this purpose.

Local Well Identifiers

Local well owners or utilities may assign identification numbers to wells, especially (but not exclusively) in the case of public water supply wells. As an example of how these local well numbers can become confusing, the following scenario may be posed. The sale of wells to new owners or administrators is sometimes accompanied by the assignment of new identification numbers to fit into the new owners' scheme. This change of well number is not always communicated to the ISWS.

If the new well owner specifies that he/she has changed the well number, or if a water quality analysis report is accompanied by an indication of the old well name and number as well as by the new information, the ISWS updates its files to reflect the changes. But with any change of ID number there is the possibility that records related to a well may be reassigned an incorrect identification number, or may never have the old number replaced with the new one. Another possibility is that only some of the related records may have their numbers changed. The rest of the records could appear to be related to a different well if that well now has the same number that was originally assigned to the old well. This is another example of how identifiers that are directly related to a well can become a problem.

The ISWS observed several different approaches to well identification in the states it consulted. Wisconsin has used a system of well ID tags related to well installation since 1988. Identification numbers are attached near the water entry point into the house. The numbers are also printed on tapes that are attached to records from that well, including chemical analysis reports and logs. Any future records for that well can easily be added to the automated system with the associated ID. The ID is so simple to use that other agencies may begin to carry the same ED number, thus establishing one cohesive information system. In the future, IDs may be assigned and mailed to owners of wells installed before 1988.

Kentucky has implemented a water-well tagging system similar to the Wisconsin system. One tag is attached to the well, and a second tag is placed on the paper file after the well has been inspected. Each well is required to have analyses done for fecal coliform and nitrate levels, and the tag number is added to the chemical analyses forms.

The concept of well tagging has been discussed by various groups in Illinois, but the logistics of putting such a procedure into practice have prevented it from becoming a reality thus far.

Multiple Agency Identification Numbers

In each state consulted, certain agencies process and store a particular type of record for wells and well-related information, and others do the same for other types of records. Each agency has its own method for assigning identifiers to its records. Over the whole range of records pertaining to any well, several IDs exist, none of which is unique to that well alone. There is very little chance that the records from one agency can be merged with those from other agencies without tedious examination to assure that they are really pertinent to the same well. It is appropriate for each agency to identify its records. The ISWS, as repository for ground-water records, must keep track of all of these records with a comprehensive identification system. That system must relate the numbers from each of these agency systems to each other.

Unique Well Identifiers

The need for unique well identifiers was recognized by the ISWS at the earliest stages of record keeping. Each agency had its own numbering system, and some agencies had more than one number pertaining to the same wells. For many decades the ISWS attempted to keep track of all of these numbers on the paper records. The permit numbers assigned by the Illinois Department of Mines and Minerals (IDMM) or the Illinois Department of Public Health (IDPH) were recorded; thus there were at least three episodes of numbering involved in the assignment of permit numbers to applications for well drilling. IDMM assigned its own number sequence; IDPH assigned a set of sequential numbers related to the permit application; and when the application was changed and made more comprehensive, the numbering was started over.

The Illinois State Geological Survey (ISGS) assigned sequential numbers within each county for each permit reported to them by the other two agencies. During the 1980s, the ISGS stopped associating these numbers with all permits for wells. Instead they began to assign them to well logs that were received, since the volume of water-well logs is smaller than the volume of permits. In the late 1980s the ISGS briefly returned to their original practice. At that time the ISWS discontinued tracking the ISGS number in recognition that confusion could be created by the changes in the number assignment policy.

In the mid-1980s, the ISWS attempted to adopt the IDs from the ISGS sequential numbers as unique well IDs. This attempt yielded the information that

not only were the numbers not unique, but the information about the wells in the ISGS files was maintained in a different way from the system at the ISWS. A one-to-one comparison for confidently merging the well files from the ISWS and the ISGS was not possible. The hope that the ISGS identifiers could be used as unique identifiers for wells was abandoned by the ISWS.

Cross-Reference Table

The next attempt made by the ISWS to merge and cross-reference well information files was the creation of a cross-reference table that would contain a list of as many wells of all types as were known, along with the identifier numbers associated with each well. As expected, the same problems described above (the lack of confidence that the locations and other support information for some wells were correct) made it likely that the "total" list of wells would have actual duplicates that could not be identified as such. This concept of a cross-reference table still has some value, but the effort required to create it and clean it up has given it a low priority on the tasks-to-be-completed list for the ISWS ground-water information and data management groups.

Record Identifiers

The ISWS has developed a record-identifier concept that goes beyond the well-specific identifier. The unique well ID is still the main data element needed to correlate the information in the ISWS and other database systems. However, well logs are only one part, although a major part, of the data. There are many other types of records, reports, and analyses of information that are regularly associated with wells. Further, there are more general local and regional hydrologic and geologic studies that add large volumes of information to that known about any area of the state. Therefore the ISWS has further abstracted the idea of a unique ID that is a record ID number — possibly a bar code — rather than a well ID number. Through this process, the records pertaining to any well can be associated with that well, but any report or information pertaining to more than one well can also be associated with each of those wells through the record IDs.

SIMILARITIES AND DIFFERENCES IN DATABASE SYSTEMS

A questionnaire was used to guide the interviews in each state in order to achieve consistency in the information gathered. Through the interviews, it became apparent that the evolution of the ISWS ground-water database and information system was similar to developments in other states. Sources of information in the systems were compared, and philosophies of information management were discussed. After consulting with the other agencies, the ISWS ground-water information group made several observations about similarities and differences in system development, structure, and funding methods among the states. The major factors compared are discussed below.

Motivation for System Creation

The first questions asked in each state were about the creation of the automated system. In every case, the agencies had previously known that an automated system would be beneficial. Each agency was able to develop some level of automation based on widely varying amounts of funding.

It is difficult for state agencies to justify the use of time and resources needed to develop a computer system without some outside stimulus. In most states the legislatures later mandated that some form of ground-water protection program be initiated. In those states where programs were mandated, the agencies further developed automated systems based on the plans they had already inaugurated. Table 2 shows the variety of factors that brought about the automated systems in Illinois and the states visited.

Table 2. Motivation for Creation of Automated Data Systems

	IL	*MI*	*IN*	*KY*	*WI*
Single group initiative	X*	X	X	X	X
Intra-agency group initiative	X				
Inter-agency initiative	X†	X			
Response to legislated mandate	X‡	X	X	X	

* *Recognition within information group*

† *Recognition motivated by projects and legislation*

‡ *IGPA enacted September 1987*

Funding

One of the biggest problems for any group that needs a database or an information system is funding. Several kinds of financial support are necessary to create, maintain, and expand an automated information system. The initial costs of creating a system include the purchase of hardware and software. Planning for a system should follow specific procedures. The first step is to determine what the users of the system do with the information and how they wish to manipulate data. The next step is to determine the best type of software to fill those needs. Only after those decisions have been made should hardware be chosen. This is to assure that the hardware can run the needed software efficiently.

The initial costs of the system should involve the development of this needs assessment and the plan for the structure and functions of the system. The initial costs are usually the greatest one-time costs for a system, although the long-term costs of upkeep and expansion will amount to a larger total figure.

Long-term costs include maintenance and updating of hardware and software. They should also include salaries and support for the individuals involved with all aspects of the system. They should include archival of information and costs of data acquisition and editing. Producing summary reports from the system and sending

staff to conferences to keep current with the activities of other agencies in the information community should also be part of the costs.

Table 3 shows the funding that was reported for the systems in the states visited. In some cases, as in Kentucky, funding was available for the initiation of the system. In Kentucky and Indiana, small budgets were available to enable the agency to use an existing system from another agency. However, the funding to expand and maintain the system must come from within the agency. In other cases, as in Illinois, agencies initiated the system within a department, and state funds later became available to help support it, although those funds are not necessarily enough to carry out the entire task.

In Michigan a state grant was combined with multi-agency funds to initiate the system. Agency funds from general revenues are used for continuing support, and user fees contribute a small amount for system upkeep. Wisconsin has developed its system entirely from state funds. Agency funds from both general revenues and contracts were used to initiate the system. Both sources, combined with user fees and a specified amount from a fee which is charged by the state to users of chemicals, pay for the database system upkeep. Table 3 shows the support for initiation of the system with an X and for the continuation of the system with a Y.

Table 3. Financial Support for Automated Systems

	IL	*MI*	*IN*	*KY*	*WI*
Federal grant					
State grant		XY			
State tax or fee-based revenue	X				XY
Multi-agency funding					
Agency general revenue fund	XY	Y	Y		XY
Agency contract funds	XY		XY		XY
Single group funding from contracts	Y				XY
User fees from outside agency	Y	Y*	Y		Y

Note: X = initial funding; Y = continuing funding

** Other state agencies pay user fees to the system*

System Features and Characteristics

The systems developed by the states consulted by the ISWS have several features in common, although they vary widely in degree of sophistication. For example, almost all the systems use a combination of mainframe computers and PCs. Systems that rely completely on the mainframe for all information processing seem to have more limited accessibility than those that use some capabilities of the main-frame and some from the PC. The limits result from costs, which are continuous on a

mainframe, and method of access (modems or hardwires), which require phone lines or cables to connect the terminals and the mainframe. PCs allow users to work independently and to carry information from computer to computer by diskette without connection costs. Indiana was using a mainframe in 1988. By mid-1990, Indiana was using PCs and mainframes together and was considering expansion of its system capabilities.

Most states have established software for their own use. The software is based on a combination of commercially available packages and programs specifically produced for the data from that state. The needs of the agencies that use the system required that some programs be written for specific purposes not provided for by commercial packages. Commercially available software offers the capability to create programs and routines to process information. For example, in Michigan and Wisconsin, where a database management package was used to enter, edit, and format data, the systems were interfaced with a wide variety of options to analyze and present data. These states both opted to have additional software written, which worked smoothly for the specific structure of their data.

All the systems had at least rudimentary GIS functions or were planning to have this capability in their systems. The variety of GIS systems was one of the features most noticeable to the ISWS group. The Michigan system seemed to be the most advanced. That system allows the user to integrate graphics, maps, and statistics with any of the data, using a menu-driven system.

At the present time, the Illinois system cannot be accessed through a central menu-driven system. Several mainframe computers are available to Illinois users, and PCs are found throughout the ISWS. There are plotters and printers associated with all of these devices. To use the GIS in Illinois, a PRIME 9955II computer is accessed. To use graphics, any one of several PC packages or one of two mainframe computers is used. To use statistics, other PC packages or mainframe packages are used. Plotting of graphs is easier on PCs but is available on the mainframes with limitations. Movement of data from one system to another is time-consuming and sometimes complex and requires changing the format of the data several times. If a user wishes to produce tabular reports, maps, and plots, he/she may have to move the same data set to different computers for each type of output.

The Michigan system has more flexibility for the single user than the Illinois system. Far less computer knowledge is needed to use the system in Michigan. If the concepts behind the development of the Michigan system were applied to the Illinois data with the available system configuration, an extremely powerful data and information system could result.

Paper output is common to all systems. However, there is a great deal of difference between a screen-dump and a report that is tailored to the needs of the user. A system that allows users to generate computer files for their own use is more useful than one that allows the user to obtain only paper output. Sophistication of the output varies greatly. Table 4 shows some of the key characteristics of automated systems and the status of each state with respect to those characteristics.

Table 4. Automated System Characteristics and Features

	IL	*MI*	*IN*	*KY*	*WI*
PC or mainframe:					
PC		X		X	X
Mainframe	X	X	X	X	X
Relational database		X X			X X
Other software		X X			X
User guide available		X			X
Menu-driven system		X X			X
GIS compatibility		X X		X	X *
Other graphics		X X			X
System output					
Paper reports	X	X	X	X	X
Computer files	X X				X
Data entry					
Interpreted		X			X X
Codes used	X			X X	X
Verbatim		X			X X
Information verification process					
Field-checking		X	X	X	X
Cross-reference	X X				X
Phone and letter contacts	X				
Agency verifies system records		X †			X X
Maintenance and updating					
Agency maintains	X X				X X
True updating (keeping current)	X	X			
Correction procedure documented		X			X X
User charges	X X			X	
Archiving					X‡

* *GIS limited at the time of the consultation*
† *Some record verification performed by other agencies*
‡ *Well construction data microfilmed until 1978*

System Input

All the states acquire data for their systems from many sources. Well permits are a source of ground-water information in all the states consulted by the ISWS. Another common source of well and subsurface data is well logs, which come from the drillers. The relationship between agencies and the drillers is important. The submission of logs to the states is achieved only when the drillers cooperate. Table 5 displays some of the significant ways in which drillers and state agencies interact.

Another information source is water quality analyses reports. Water quality information for domestic/private wells is available when samples are required after new wells are installed, or when owners request analyses. General water quality analyses, such as those collected by the IEPA from public water supply wells in

Table 5. Information Sources Related to Well Installation

	IL	*MI*	*IN*	*KY*	*WI*
Well permitting system					
Statewide	X		X*		X
Only some counties	X†	X			
Fee for well permits	X	X	X		X
State education of drillers in the use of its forms		X		X	
Drillers licensed	X	X	X	X	X
Drillers regulated	X	X	X	X	
Well logs required to be submitted for each well drilled	X	X	X	X	
Penalties for non-submission of logs	X	X	X		
Required water quality tests for water from wells				X	
Penalties for non-submission of samples					
Well-tagging procedure				X	X

* *Statewide permitting administered by the counties*
† *Some Illinois counties began permitting in 1989*

Illinois, may be the only available broad-scope database of water quality. Regulatory agencies such as EPA and the Department of Public Health have programs which require the collection of water quality data. These data are an integral part of the total database system in Illinois. Most states collect monitoring data from their public water supply wells in response to federal regulations. But as shown in table 5, not all states require water quality tests for private/domestic wells.

Each state had features that seemed unique. These features were noted since they impacted the quality or quantity of data that eventually reached the information-collecting agency. For example, in Indiana, conservation officers are assigned to enforce compliance of the drillers with regulations. In other states, lack of compliance with regulations is found during the process of checking specific aspects of a regulation.

Kentucky's policies with respect to drillers involve a strong relationship between the state and the drillers. The state trains the drillers to fill out the required forms properly. As a result of this practice, Kentucky has a large volume of well-filled-out records. The Michigan Department of Health also educates its drillers to fill out its forms properly, with a high degree of success.

Kentucky has a unique form of data entry: prison inmates are trained to perform data entry for the KDNR. This practice serves the state in its need for data entry, in addition to providing rehabilitation and vocational training to the prisoners. It appears to be a very cost-efficient, mutually beneficial practice.

The Michigan approach to planning its system is very structured. They begin with a needs assessment and follow that with software choices and then hardware selection. Each step along the way is reviewed by several groups before the next step is taken. The varied interests of many groups are therefore considered during the planning stages. The end result has been a system with a wide variety of capabilities that can meet many varied needs.

The Wisconsin well tagging process, though not entirely unique, is one of the better developed systems. The use of stickers that accompany the tag allows the home owner, installer, or sample collector to attach the identification number directly onto the papers entering the information flow. This assures that information is associated with the correct well.

The Michigan data retrieval system is especially interesting. The Welldig program for data retrieval and the Cmap program for mapping allow retrieval and visual representation of data from several sources at one time. The capabilities of interfacing selected data and zooming in on a portion of those data are very effective tools. Although this capability exists in the GIS on mainframes in other systems, it is accessible on PCs in the Michigan system, making the composite maps quickly available to the user. With the wide variety of data in the Illinois system, this easy access feature could be extremely powerful for research and service.

Updates and Corrections

In several states, data are entered in the automated system only after field-checking. For example, Michigan has made a decision to enter only field-checked data into its system. With larger systems, such as that in Illinois, which include many types of data from all parts of a state, staff and funding limitations prohibit field-verification of every piece of information that comes into the information system. This does not mean that the unverified information should not be kept in the system. Field checking creates a high-quality data set, but does not yield a comprehensive set. Most automated data systems are used, or will be used in the future, for purposes that require broad coverage of information.

The ISWS recommends that all available information be entered into a system. Those records that have been field-checked can contain a field to show this fact. If users need data that are absolutely correct, they can select only the field-checked records. However, if, for example, they need to have an estimate of the number of wells that may be impacted by a new industrial installation, they can select all records in a specific area, knowing that all records may not be absolutely correct.

The ISWS is particularly sensitive to the difference between updating and correction activities in automated systems. Updating is defined by the ISWS group as adding to the information system. This includes adding new records as well as adding changes of status when the old status was once correct but has changed. This is different from correction of records. Correction, to the ISWS group, means that information formerly thought to be true has been found to be false and has been changed to reflect that situation.

Wisconsin continually updates data, but only in the computer database. Michigan verifies data entry and asks for review information from agencies that have

supplied data, but it does not change records that are found to be incorrect after the data for those counties have been finalized in the system. In our experience, we have found that if records are not maintained on a continuing basis, the data set created from those records soon becomes inadequate for research. This is especially true for municipal and industrial well records, where changes are an everyday occurrence. By having a mechanism in place to cover the regularly occurring changes in information, the system is kept current. Automation and the use of the record ID and a cross-reference table allow multiple records to be changed simultaneously. To maintain records in such a way is extremely tedious and time-consuming. However, this procedure is necessary to maintain the integrity of the database system.

RECOMMENDATIONS TO DATABASE AND COMPUTER GROUPS

General Recommendations

It is recommended that three areas of effort be part of the information collection and database creation process:

1) A list should be compiled of the minimum information needed in any database to allow its use.

2) Standard formats should be established for those minimum data elements.

3) Perhaps most importantly, documentation must be produced to accompany any data so that definitions and limitations, as well as strengths and specifications, are clear.

All systems and programs should be documented so that anyone can use them. Automated information systems are complex and usually have a variety of capabilities and functions. Therefore they should be carefully documented for several reasons. First, users need to have a ready list of the functions available, as well as the necessary instructions for performing their tasks. Second, systems operators need documentation so they may maintain or enhance a system without destroying any of its capabilities. Third, potential funding groups can use documentation to determine whether a system performs the functions that fulfill their needs, or whether it can be expanded to meet their needs through the use of additional funds. For these and other reasons, documentation is a fundamental part of any automated information system.

The requirement of these minimum elements does not preclude anyone from collecting other information, or from saving it in any format desirable. It merely establishes the basis upon which any interested party can access the data with confidence, and upon which incorporation into the overall database system can be achieved.

It is further recommended that the assignment of record identification numbers be started, and that a cross-reference table be created, maintained, and made

available. Once a workable system of record numbers becomes available, information of all types will be usable for specific wells, for regions, or for categories of wells. For example, a researcher can get all information related to a specific well or may obtain all information related to sand and gravel wells of a certain depth range throughout the state.

A strong recommendation is made that a regular archival system be established. The volume of records is only the first reason for archival. The ISWS and other groups interviewed have records on file that exist nowhere else, and some of these historic records are deteriorating in present storage. If a disaster were to strike, the loss of those records would cause irreparable damage.

One general observation was made by the information group after the trips to other states. It was clear that a system with a wide variety of data available for many uses was both feasible and extremely beneficial. Not only GIS capabilities, but other types of graphic display, statistical analyses, and report generation make the data in the system a more powerful tool than any paper filing system ever could be. There is a clear need to develop graphic capabilities to display many kinds of data without requiring complex movement and re-entry of data. Among the graphic capabilities should be a simple system to create hydrographs that can be incorporated into documents in-house. Such simple displays as histograms, bar charts, pie charts, and contour plots should be directly available to users of the large databases. Such capabilities would stimulate the use of existing data to produce proposals and to enhance research.

The ISWS has a clearly defined ground-water data management group, of which the information group is a part. This group is already involved with the databases. Plans for expansion of the information in the system, and especially for an increase in the available uses of the system, are the key to a strong program that can benefit others, as well as the ISWS. In the interviews with other states, it was clear that where a specific group has been designated to develop an automated system, the plans become better defined and user needs are considered in the planning. Where there is no designated planner, systems usually develop in a more piecemeal way, and sometimes software and data structures must be reworked several times. Michigan developed a planning document for its system, which took user needs into consideration.

Archival of the Information System

Because many of the documents in the ISWS information system and systems in other states are unique, there is an advantage to archiving them in case of loss of the files through disaster. In much the same way, the automated information systems should also be archived, or backed up on a regular basis, and a copy of that backup should be kept off-site for safety's sake. For several decades the ISWS tried to maintain a set of microfilm copies of its paper records. However, because of the cost and time involved in filming the entire volume and variety of records, a complete set of films was never made. Other states consulted by the ISWS group had similar situations and either had no archives or very skeletal copies of records. Deterioration of older films has rendered many of those records unreadable.

As recently as three or four years ago, the ISWS group considered having a complete set of microfilms or microfiche archives made of its paper files. The effort would have cost at least $30,000 to $50,000 for the initial double set of films, and then a continuing cost for updates. Retrieval of those records from film would be difficult and costly, but this would serve as a way to preserve some very old and delicate records, as well as the large volume of logs and reports. Even this proposed work would not have involved making copies of all the records.

In its continuing quest for an answer to this problem, the ISWS contacted the Ohio Department of Natural Resources (ODNR), Division of Water, Ground Water Resources Section. The ISWS had learned that the ODNR section had a system composed of computers and laser disks for the storage and retrieval of its water-well logs and maps. The system allows for scanning of water-well logs and maps into the system and for their storage on video disks. The scanned records are then linked to a relational database management system into which key fields can be entered for searching and retrieving records. The ODNR system managers were extremely interested in further linking the disk-based system to a GIS for map production.

The ODNR system is expensive, but it offers a very appealing solution to several information management problems. The first problem solved with a system such as the ODNR system is the obvious archival problem. Scanning a set of records and storing them on laser disks provides a set of easily stored, archived documents. The ODNR management feels that such a form of record storage will be given the status of legal document storage in the near future. This is a factor not previously considered by the ISWS group. Computerized records and photocopies of microfilmed records are not legal documents.

A second problem solved by the ODNR type of system is that a variety of sizes and shapes of documents can be scanned for storage on disk. Maps, recorder charts, and decaying documents have presented a serious problem when other forms of archival have been considered.

The ability to retrieve records rapidly from a vast amount of data is a further advantage of the ODNR system. Since the laser disks are linked to a relational database management system, they can be searched for a variety of key fields and subsetted before they are brought to the computer screen for viewing or taken to a file for copying or other data reduction. This makes the daily, constant use of the paper files unnecessary. Space and time would be saved by using such a system. This capability also allows the archival system to be much more than that; it would now be a daily use system as well. Therefore, the cost of the system is more justifiable.

The concept that an automated information system can fill several functions is not new, but it has not been a primary part of the planning for the ISWS system. Archival of documents used to mean paper duplication of paper records only. That duplication was expensive and served little useful purpose other than to preserve a set of the records. With automation, copies of many of the records, or at least the information from many of the records, can be preserved without duplication of the paper files.

The integration of relational database management, GIS capabilities, statistics and data reduction, computer modeling, and video scanning of complex records

can offer a system responsive to changes in the status of information. Information can now be a readily available, useful tool in responsive service, research that draws on multiple sources of data, and planning for both the public and private sectors.

Structured Computer Files

The consultations with other states have reinforced the idea that information systems must have several capabilities in order to be effective tools for use by all of the staff and requesters of information. A system must have accessible data, updating capacity, correctability, GIS compatibility, and an easy connection with software in order to model and analyze the data from all parts of the system.

The GIS capability seems to be the one that has drawn the greatest amount of attention recently. This is not surprising considering the fact that maps can be used to focus attention on the features of concern, with little need for extensive explanation. In all the states interviewed, such a capability has been considered, and in some cases it has been acquired, as shown in table 4.

The ISWS has been aware that the improper use of data with a GIS can create serious problems. The power of the GIS is obvious; the dangers are a little harder to see at first. The quality of the information in the system, and the care with which those data are combined, will determine the true quality of the products from that system. Kenneth Hlinka and Susan Schock have written a paper on this subject, "GIS Applications in Ground-Water Data Management: The Problem of Scale," which is available from the authors of this report on request. The ISWS group recommends that the system developed for its data be driven by the entire scope of needs of the users with the added sophistication of an available GIS, rather than focusing the system on the production of maps with little concern for the quality and proper use of those data.

Responsibilities of Information Managers

Information managers have high hopes for the future. The technology is changing rapidly, and automation has brought information to the attention of planners and researchers with a new emphasis. Future planning for information systems will be influenced by several factors, not the least of which is funding. As legislatures determine that research into water-related areas requires structured data, mandates for information systems are being written. However, these federal and state regulations are frequently written without adequate funding to accompany them. In those cases, the regulations are followed by a brief period of interest at the state or local level, and a temporary influx of funds, usually at a level far below what is needed.

There are periods of waxing and waning interest in the issues related to various aspects of water research, but consistent funding at the needed levels to carry out and maintain a viable automated system with all the necessary parts is not common. As in Illinois, Wisconsin, and Michigan, agencies attempt to maintain a system with internal funding, but cannot achieve the scope of capabilities they would like for their data use. The ISWS group believes that one reason for this is that there is not a true recognition of the broad-spectrum value of the data in automated

systems, nor is there an understanding of how beneficial a well-developed and well-maintained system can be in meeting the needs of the users. Information managers everywhere must try to make agency heads and legislators aware of how much benefit can be derived from putting such a system in place. They must also give as clear a picture as possible of what is needed to put such a system in place and to properly maintain it. If managers can do that, all users of information will be given tools with which to collect, correct, and analyze their data to the benefit of all the people in their state.