INTERACTIVE DATABASE OF PRELIMINARY ASSESSMENT SCORESHEETS FOR CROSS-SITE COMPARISONS

by

JOEL D. GUZMÁN
B.S. Civil Engineering
University of California, Davis, 1996

Submitted to the Department of Civil and Environmental Engineering
In Partial Fulfillment of the Requirements for the Degree of

MASTER OF ENGINEERING
IN CIVIL AND ENVIRONMENTAL ENGINEERING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 1997

Signature of the Author _____
Department of Civil and Environmental Engineering
May 9, 1997

Certified by _____
Patricia Culligan-Hensley
Assistant Professor of Civil and Environmental Engineering
Thesis Supervisor

Certified by _____
Feniosky Peña-Mora
Assistant Professor of Civil and Environmental Engineering
Thesis Supervisor

Accepted by _____
Professor Joseph Sussman
Chairman, Department Committee on Graduate Studies

Interactive Database of Preliminary Assessment Scoresheets for Cross-Site Comparison

by

Joel D. Guzmán

Submitted to the Department of Civil and Environmental Engineering on May 9, 1997 in Partial Fulfillment of the Requirements for the Degree of Master of Engineering in Civil and Environmental Engineering

Abstract

Computerized information systems assist people and improve processes by increasing access to information, automating tasks, and aiding with decision making. This work addresses an information system designed to assist with the identification of federal environmental remediation projects. The system specifically improves the preliminary site assessment phase of the Superfund process. The Superfund program was created as a result of the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) passed by Congress in 1980. CERCLA established broad authority for the government to respond to problems posed by the release, or threat of release, of hazardous substances, pollutants, or contaminants. The computerized information system will aid in the gathering of background information about particular sites under remedial investigation. This information can then be used to determine the next step in the Superfund process.

This work focuses heavily on the development of an architecture for a site comparison relational database. Completed preliminary assessment scoresheets of facilities can be a valuable source of information for potential hazardous waste sites currently under investigation. The Structured Query Language is used to perform the site comparisons. An example of its implementation, including the design of a graphical user interface, is also examined using Microsoft Access 95.

The information system was developed in conjunction with the Information Technology, Master of Engineering students in Civil and Environmental Engineering. The completed information system will provide both "executive information" and "decision support". "Executives" or other decision makers and engineers will be able to use the system to impose accountability for answers to individual scoresheet sections and review the reliability of information that is used to complete the preliminary site assessment scoresheets. To support individual decisions made for completing scoresheets for new potential hazardous wastes sites, the information system provides links to data sources on the Internet and previously recorded preliminary site assessments.

Thesis Supervisors: Patricia Culligan-Hensley
Feniosky Peña-Mora

Acknowledgments

I would like to thank the many individuals who contributed to the completion of this thesis.

For the guidance and support of my thesis advisors, I thank Feniosky Peña-Mora and Patricia Culligan-Hensley.

Thank you to David Marks and Shawn Morrissey for acting as mediators during the Information Technology Group Project definition process.

For providing consultation on the important issues involved in environemental remediation procedural processes, the contributions of Kenneth Till are greatly appreciated.

To the other Information Technology group members including Kai Kuo, Anna Lukasiak, Carrie Morton, and Rony Mukhopadhyay, I am grateful to have been a part of one of the pioneering efforts to apply information technology to an environmental engineering application.

Lastly, I thank my family members including Dave, Anne-Marie, and Daniel and especially my parents for the support and love they have exhibited towards me throughout my life.

Table of Contents

1. 0 INTRODUCTION 8

1.1 OVERVIEW OF SUPERFUND (CERCLA) PROCESS 9

1.1.1 History 9

1.1.2 Current Superfund Process 11

1.2 PRELIMINARY ASSESSMENT UNDER CERCLA 12

1.2.1 Objective 12

1.2.2 Scope 12

1.2.3 The Preliminary Assessment Scoresheets 13

2. 0 INFORMATION SYSTEM 15

2.1 DEFINITION OF AN INFORMATION SYSTEM 15

2.2 USE OF INFORMATION SYSTEMS IN DECISION MAKING PROCESS 16

2.3 APPLICATIONS OF INFORMATION SYSTEMS IN THE ENVIRONMENTAL FIELD 17

2.4 PROPOSED INFORMATION SYSTEM FOR PRELIMINARY SITE ASSESSMENT UNDER CERCLA 18

2.4.1 Architecture Under current Preliminary Site Assessment Processes 19

2.4.2 Future Architecture of Preliminary Site Assessment Processes 21

3. 0 RESULTS FOR ARCHITECTURE UNDER CURRENT PA PROCESSES 23

3.1 GRAPHICAL USER INTERFACE FOR EXISTING PRELIMINARY SITE ASSESSMENT SCORESHEET 23

3.1.1 Overview of the Graphical User Interface of the Electronic Scoresheet 23

3.1.2 Structure of the Electronic Scoresheet 24

3.1.3 Implementation of the Electronic Scoresheet 26

3.1.4 Alternative Implementation Scheme 28

3.1.5 Future Recommendation 28

3.2 DEVELOPMENT OF INTERNET SEARCH ENGINE FOR GROUNDWATER PATHWAY CRITERIA LIST 30

3.2.1 Availability of Information on the Internet for Completion of PA Scoresheets 30

3.2.2 Accessing Internet Search Tools Through the Use of Common Gateway Interface 31

3.2.3 Multi-Keyword Searching 32

3.2.4 Displaying Results 33

3.3 DEVELOPMENT OF DATA STORE SEARCH ENGINE FOR SURFACE WATER PATHWAY CRITERIA LIST 35

3.3.1 The Data Store Search Engine's Role in the System 35

3.3.2 Description of USGS Data Source 37

3.3.3 Matching Existing Data Source Information with Scoresheet Questions 38

3.3.4 Conclusion 39

4. 0 SYSTEM ARCHITECTURE FOR THE INTERACTIVE PRELIMINARY ASSESSMENT SCORESHEET DATABASE 40

4.1 OVERVIEW 40
4.1.1 Motivation 40
4.1.2 Design and Population of Data Storage 40
4.1.3 Database Queries and Comparison Criteria 41
4.1.4 Using the System 42
4.1.5 Making Use of the Results 42
4.1.6 Interactive Database Organization 43
4.2 INFORMATION SYSTEM DESIGN - RELATIONAL DATABASES 43
4.2.1 A Non-relational Database Example 43
4.2.2 Avoiding Data Redundancy 45
4.2.3 Establishing Table Relationships 47
4.2.4 Maintaining Flexibility 49
4.2.5 Keeping the Data Fields Simple 51
4.2.6 Other Relational Database Design Principles 51
4.3 DATABASE DESIGN FOR PRELIMINARY ASSESSMENT SCORESHEETS 54
4.3.1 General Information Tables 56
4.3.2 Source Evaluation Tables 59
4.3.3 Ground Water Pathway Tables 59
4.3.4 Surface Water Pathway Tables 61
4.4 AN INTRODUCTION TO THE STRUCTURED QUERY LANGUAGE 62
4.4.1 Single Table Query 63
4.4.2 Multiple Table Query 65
4.4.3 Variables, New Tables, and Aggregate Functions 67
4.4.4 The Top Keyword 68
4.4.5 The Keywords DISTINCT and DISTINCTROW 69
4.5 PERFORMING QUERIES IN THE PA DATABASE 71
4.5.1 Location Comparison SQL 72
4.5.2 Type of Facility Comparison SQL 74
4.5.3 Hazardous Waste Comparison SQL 75
4.5.4 Waste Source Type Comparison SQL 77
4.5.5 Methodology For Multiple Parameter Comparisons 79
4.5.6 Methodology For Pathway Comparisons 79

5. 0 IMPLEMENTATION OF THE PRELIMINARY ASSESSMENT SCORESHEET DATABASE FOR CROSS-SITE COMPARISONS 80

5.1 USING MICROSOFT ACCESS 95 80

5.1.1 Graphical User Interface and Forms 80

5.1.2 Controls 81

5.1.3 Macros and Events 82

5.1.4 Subforms 84

5.2 IMPLEMENTING THE INTERACTIVE PA SCORESHEET DATABASE 85

5.2.1 Switchboards 86

5.2.2 Location Comparison 90

5.2.3 Type of Facility Comparison 92

5.2.4 Hazardous Wastes Comparison 94

5.2.5 Waste Source Type Comparison 95

6. 0 CONCLUSIONS AND RECOMMENDATIONS 97

6.1 ELECTRONIC INFORMATION 97

6.1.1 Availability 97

6.1.2 Standardization 97

6.2 TECHNOLOGY 98

6.2.1 Data Source Collection 98

6.2.2 Linking the EIS with the DSS 98

6.3 IMPROVEMENTS TO THE INTERACTIVE DATABASE OF PA SCORESHEETS FOR CROSS-SITE COMPARISON 99

6.4 EXPANDABILITY - BEYOND THE PRELIMINARY ASSESSMENT PHASE 99

6.5 IMPACT OF THE SYSTEM ON THE PRELIMINARY ASSESSMENT PHASE 100

7. 0 REFERENCES 101

APPENDICES 103

APPENDIX A. PRELIMINARY ASSESSMENT SCORESHEETS 103

APPENDIX B. USGS DATA FILES 151

APPENDIX C. COMPLETE LIST OF WATER-USE ELEMENTS 152

APPENDIX D. MICROSOFT ACCESS PRELIMINARY ASSESSMENT SCORESHET FORMS 160

List of Figures

FIGURE 1.1.2A SUPERFUND PROCESS 10
FIGURE 2.4A INFORMATION SYSTEM FOR PRELIMINARY SITE ASSESSMENT 19
FIGURE 2.4.1A FLOW OF INFORMATION 20
FIGURE 2.4.1B USER'S PROSPECTIVE 21
FIGURE 3.2.3A EXAMPLE OF INTERNET SEARCH ENGINE 33
FIGURE 3.2.4A GROUND WATER PATHWAY SECTION OF ELECTRONIC PA SCORESHEET 34
FIGURE 3.2.4B ELECTRONIC PA SCORESHEET WITH SEARCH RESULTS WINDOW 35
FIGURE 3.3.1A DATA STORE SEARCH ENGINE SYSTEM DIAGRAM 36
FIGURE 4.2.1A NON-RELATIONAL DATABASE MODEL FOR PRELIMINARY ASSESSMENT 44
FIGURE 4.2.2A AVOIDING REDUNDANCY -BREAK TABLE INTO TWO SEPARATE TABLES 46
FIGURE 4.2.2B INTRODUCTION OF PRIMARY KEY 47
FIGURE 4.2.3A MANY-TO-ONE RELATIONSHIP 48
FIGURE 4.2.3B ONE-TO-ONE RELATIONSHIP 49
FIGURE 4.2.4A MAKING TABLES MORE FLEXIBLE 50
FIGURE 4.2.5A KEEPING DATA FIELDS SIMPLE 51
FIGURE 4.2.6A ENFORCING REFERENTIAL INTEGRITY 52
FIGURE 4.2.6B ELIMINATING MANY-TO-MANY RELATIONSHIPS 53
FIGURE 4.3A SECTIONS OF THE PRELIMINARY ASSESSMENT SCORESHEET 54
FIGURE 4.3.1A SAMPLE OF TYPEOFFACILITY/OPERATION TABLE RECORDS 57
FIGURE 4.3.1B ALL RECORDS CONTAINED IN THE SOURCE TYPE TABLE 58
FIGURE 4.3.1C RESTRICTED VALUES FOR DATA FIELD STATEID 59
FIGURE 4.3.3A DISTANCE INTERVALS IN 4DISTANCEINTERVALS TABLE 60
FIGURE 4.3.4A WATERBODYTYPE TABLE RECORDS 62
FIGURE 4.4A TABLE USED FOR ILLUSTRATION OF SQL 63
FIGURE 4.4.1A SIMPLE SQL RESULT 64
FIGURE 4.4.1B SORTING IN SQL 64
FIGURE 4.4.1C SELECTING ALL DATA FIELDS 65
FIGURE 4.4.1D SELECTING MULTIPLE DATA FIELDS 65
FIGURE 4.4.2A TABLES FOR MULTIPLE TABLE SQL ILLUSTRATION 66
FIGURE 4.4.2B MULTIPLE TABLE SQL RESULT 67
FIGURE 4.4.3A GROUNDWATER-SCORESHEET SAMPLE RECORDS 67
FIGURE 4.4.3B RESULT OF BASIC MATHEMATICAL OPERATIONS 68
FIGURE 4.4.3A RESULT OF AGGREGATE FUNCTIONS 68
FIGURE 4.4.4A GROUNDWATER-SCORESHEET SAMPLE RECORDS 69

FIGURE 4.4.4B USE OF TOP KEYWORD 69
FIGURE 4.4.5B USING DISTINCT IN SQL 71
FIGURE 4.5.1A "LOCATION COMPARISON" SQL STATEMENT: 73
FIGURE 4.5.1B "LOCATION COMPARISON" QUERY SUMMARY 74
FIGURE 4.5.2A. SAMPLE OF TYPEOFFACILITY/OPERATION TABLE RECORDS 75
FIGURE 4.5.2B "TYPE QUERY" SQL STATEMENT: 75
FIGURE 4.5.3A SAMPLE OF HAZARDOUSSUBSTANCES TABLE RECORDS 76
FIGURE 4.5.3B "HAZARDOUS RANKING QUERY" SQL STATEMENT 77
FIGURE 4.5.4A ALL RECORDS CONTAINED IN THE SOURCETYPES TABLE 78
FIGURE 4.5.4B "WASTE SOURCE COMPARISON" QUERY SUMMARY 78
FIGURE 5.1.2A CONTROLS IN A SAMPLE FORM 81
FIGURE 5.1.3A A SAMPLE MACRO 82
FIGURE 5.1.3B LINKING AN EVENT WITH A MACRO 84
FIGURE 5.1.4A EXAMPLE OF SUBFORM 85
FIGURE 5.2.1A MAIN MENU OF THE PA DATABASE 87
FIGURE 5.2.1B MAIN MENU FOR REVIEWING PA SECTIONS. 88
FIGURE 5.2.1C MAIN MENU FOR PERFORMING FACILITY COMPARISONS 90
FIGURE 5.2.3A TYPE OF FACILITY COMPARISON FORM 93
FIGURE 5.2.4A HAZARDOUS WASTE COMPARISON FORM 95
FIGURE 5.2.5A WASTE SOURCE COMPARISON FORM 96
FIGURE D1. GENERAL FACILITY INFORMATION FORM 161
FIGURE D2. SOURCE EVALUATION FORM 162
FIGURE D3. MAIN SURFACE WATER PATHWAY FORM 163
FIGURE D4. SURFACE WATER PATHWAY CRITERIA LIST FORM 164
FIGURE D5. SURFACE WATER PATHWAY SCORESHEET FORM 165
FIGURE D6. MAIN GROUND WATER PATHWAY FORM 166
FIGURE D7. GROUND WATER PATHWAY CRITERIA LIST FORM 167
FIGURE D8. GROUND WATER PATHWAY SCORESHEET FORM 168
FIGURE D9. GENERAL SITE INFORMATION FORM 169

1.0 Introduction

1.1 Overview of Superfund (CERCLA) Process

1.1.1 History

In the past, there was little understanding of what effect certain wastes have on human health and the environment. Consequently, numerous abandoned hazardous waste sites contributed to the pollution of the earth's soil, water and air. Some common hazardous waste sites include abandoned warehouses, manufacturing facilities, processing plants and landfills. In 1980, Congress established the Superfund Program to clean up these sites in response to a growing concern over the health and environmental risks posed by hazardous wastes. The Superfund program was created as part of the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA), which established broad authority for the government to respond to problems posed by the release, or threat of release, of hazardous substances, pollutants, or contaminants. In 1986, CERCLA was amended by the Superfund Amendments and Reauthorization Act and by the National Contingency Plan (NCP). At present the U.S. Environmental Protection Agency (EPA), in cooperation with individual states and tribal governments, administers the Superfund Program.

The Superfund Trust Fund was established to support the cost of cleanup of hazardous waste sites under the Superfund program. The Trust Fund is supported from taxes on the chemical and petroleum industries and is used primarily when those companies or people responsible for contamination at Superfund sites cannot be found, or cannot perform or pay for the cleanup work.

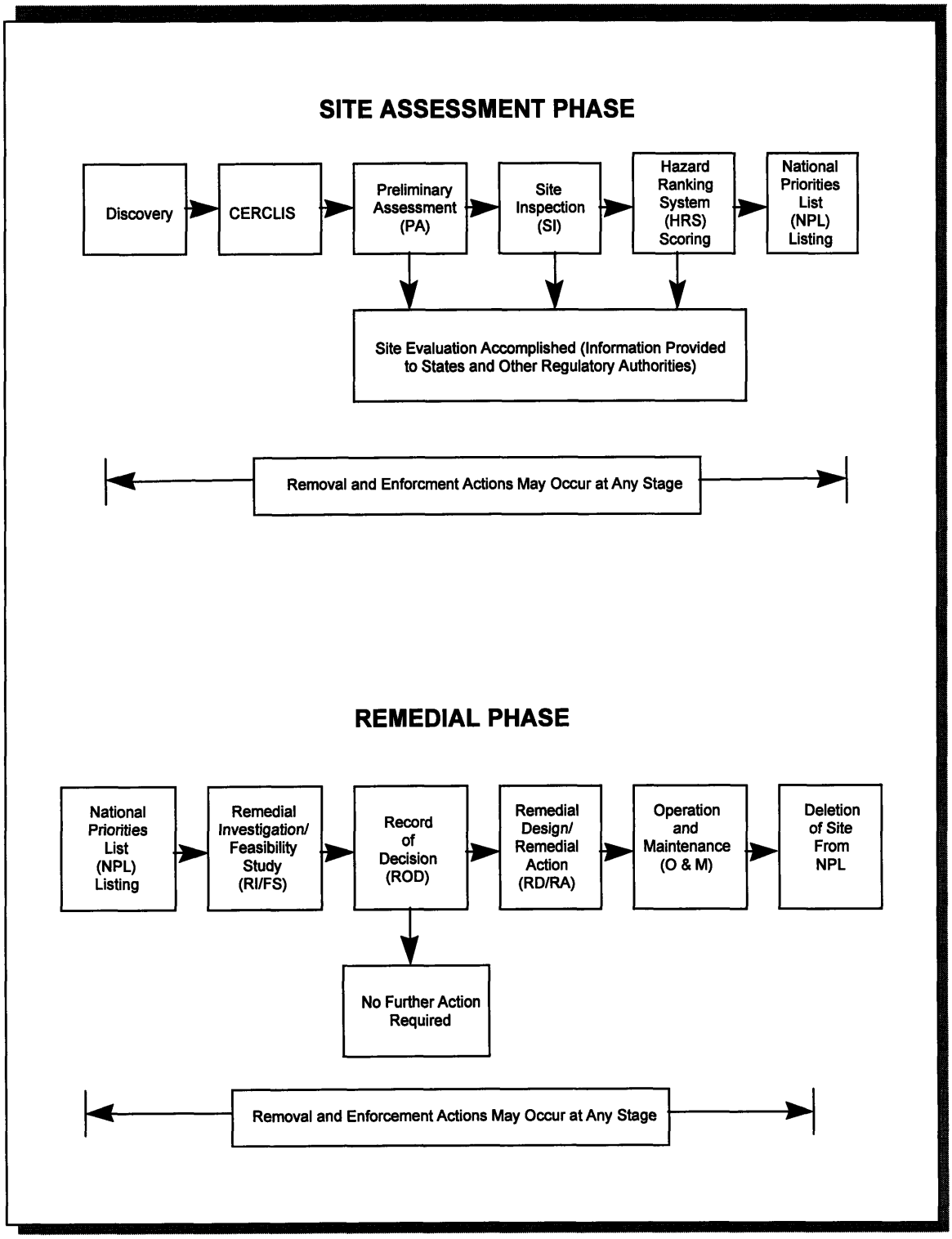


Figure 1.1.2a Superfund Process

1.1.2 CURRENT SUPERFUND PROCESS

The superfund process consists of two main phases: site assessment and remedial response action (see Figure 1.1.2a). Site assessment is the evaluation of all sites to determine those sites for which some response action may be required. If appropriate, the result of the site assessment process is the listing of a hazardous waste site on the National Priorities List (NPL). For sites that are placed on the NPL, the second phase of the superfund process, the remedial response action, is performed. During this phase, the nature and extent of contamination is determined, followed by the selection and implementation of any necessary cleanups at the site. If threats to human health are imminent, immediate or short-term responses may be performed during either of these two main phases.

The site assessment phase begins with notification to the EPA of possible releases of hazardous substances. Sites are then entered into the Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS), which is the EPA's computerized inventory of potential hazardous substance release sites. The site assessment phase continues with the Preliminary Assessment and the Site Inspection stages. The Preliminary Assessment stageuses relatively limited data that is readily available to identify sites that may pose a threat to human health and the environment, and therefore require further investigation. If the Preliminary Assessment phase recommends further investigation, only then is the Site Inspection performed. The purpose of the Site Inspection is to determine which sites have a high probability of qualifying for the NPL. Once a site has been placed on the NPL, the site will undergo the remedial response action, as explained previously.

Recently, the EPA developed the Superfund Accelerated Cleanup Model (SACM) to allow for immediate action combined with continuing study as necessary. The SACM improves upon the traditional Superfund process, which requires a prolonged initial phase of study and assessment. Under SACM, the EPA can institute actions to address threats to the health and safety of the surrounding population and environment as soon as those threats are identified. Listing sites on the NPL continues to be a prerequisite to using certain remedial action authorities to clean up contaminated sites.

1.2 Preliminary Assessment under CERCLA

1.2.1 OBJECTIVE

Based on limited data, the Preliminary Assessment (PA) phase is designed to recommend whether or not a site should undergo further investigation.

1.2.2 SCOPE

As noted in section 1.1.2, the Environmental Protection Agency (EPA) maintains a computerized inventory (CERCLIS) of potentially hazardous sites that have been "discovered" by the EPA regional offices, state agencies or private citizens. Every site in CERCLIS must undergo the PA. The PA is performed using readily available information about a site and its surrounding area. The report generated from the PA summarizes the information gathered, and based on this, concludes that either (i) the site poses no threat to human health or the environment; (ii) there is a potential threat and the site needs further investigation; or (iii) emergency actions are necessary. If the site is determined to be potentially hazardous, the PA report will often be referred to throughout successive stages of the superfund process.

The PA report consists of three parts: the data and site characteristic form, the narrative report and the PA Scoresheets (Appendix A). The data and site characteristics form, entitled "Potential Hazardous Waste Site Preliminary Assessment Form," is a four page summary of the PA scoresheets and the narrative report. The narrative report summarizes all the information researched and presents it in a predetermined structure. The last section of the narrative report should summarize the most important characteristics of the site and explain the major points of concern. The final section, the PA scoresheet, is described in the following section.

The Guidance for Performing Preliminary Assessments under CERCLA (U.S. EPA 540/G-91/013, Sept. 1991) defines the scope of the Preliminary Assessment as sufficient to complete the following tasks:

- Review existing information about the site.
- Conduct a site and environs reconnaissance.
- Collect additional information about the site with an emphasis on target information.
- Evaluate all information and develop a site score.
- Prepare a brief site summary report and site characteristics form.

Filling out the three sections takes an average of 120 hours for each site, and the information can be presented informally (i.e. legible handwriting as opposed to type written).

1.2.3 THE PRELIMINARY ASSESSMENT SCORESHEETS

The PA scoresheets are distributed as a workbook made up of checklists, worksheets, factor value tables, and scoring forms, each with brief instructions and guidelines for scoring (Appendix A). Some regions may require additional scoresheets, but there is a set of standard scoresheets that must be filled out for all regions.

The scoresheets are divided into six sections; General Site Information, Source and Waste Evaluation, and four more sections corresponding to the four hazardous substance exposure routes called pathways; Ground Water Pathway, Surface Water Pathway, Soil Exposure Pathway and Air Pathway. Each pathway section is loosely divided into three sections based on factor categories; likelihood of release (relative likelihood of a hazardous substance migrating from the site through the specific pathway), target (presence of people, physical resources or environmental resources that may be threatened by release of a hazardous material from the site), and Waste Characteristics (an estimation of the type and quantity of the wastes at the site). The particular importance of each factor can vary with the pathway, but, for example, primary targets are weighed heavily in the score regardless of pathway.

The scoresheets are set up so that the left hand pages of the workbook provide instructions for filling out the right hand pages, and often explain the questions asked in greater detail, or help the environmental engineer transfer data obtained into a numerical score for a particular section by providing tables and formulas. There is also a review for internal consistency included in the workbook, which is designed to eliminate inconsistencies in the report, which may undermine its overall validity. The EPA stresses, however, that the reviews

and guidelines are merely to assist the environmental engineers in the scoring process, and much of the time the engineer will be expected to use his or her professional judgment in the actual scoring.

In this manner, many sections or pages are assigned a total score, which is combined at the end to determine the overall score of the site. Many of the pages, however, simply ask for an explanation of certain aspects of the site in paragraph form, rather than a numerical score. The total time to research the information and score a site averages about 100 hours, and writing the reports averages about 20 hours. Sites determined to be ineligible for CERCLA response (i.e. sites where there is no danger of hazardous waste leakage, not simply a lack of targets) may submit abbreviated PA reports. The scoresheets need not be submitted for CERCLIS analysis. However, the first two pages of the Potential Hazardous Waste Site Preliminary Assessment Form and the narrative report remain a requirement.

Finally, the decision (i.e. further action or no action) made concerning the PA is usually based on the overall site score. In general, a score of 28.50 or higher receives a recommendation for further investigation, while a score of less than 28.50 receives a "No Further Remedial Action Planned" (NFRAP) recommendation.

2.0 Information System

This section covers the definition of an information system, how information systems are used in decision making processes, and gives examples of applications in the environmental field. Then, an architecure is presented of the development of an information sytem under the current preliminary assessment process as well as in the future.

2.1 Definition of an Information System

An information system can simply be defined as a system for retrieving appropriate and relevant data from a source and transferring it to a designated target with different format. A more advanced information system would also consist of a unit for processing the data and adding meaning to it.

Although extremely complex, the human brain is the smallest information system in terms of scale. Our brains contain a large amount of information that we call knowledge. When we encounter a problem, part of our brain requires information from another. After the knowledge is processed, it is expressed through speech, writing or body movement. A think tank—generally a group of people with similar skill-sets that come together to solve a problem—is an even more complex setup of an information system. Although they have a greater number and resources, think tanks introduce complex problems of conflict resolution and sharing.

A more diversified information system is demonstrated in a company. People with different interests and skill-sets are put together to help the company perform better as a whole. Information is stored not only in employees' brains, but also on paper and on other formats. Processors of the information system can include engineers and scientists, office administrators, accountants and managers, depending on what type of organization the company is.

Other large-scale information systems are universities and governments. These information systems are extremely complicated. More importantly, however, is the fact that knowledge transfer in these complex systems comes from different sources. For example, when a student has a question, the information and knowledge sources can be professors, teaching assistants or roommates. The reliability of this information can vary according to the source. Because of the difficulty in ensuring reliability of information, computerized information

systems have been implemented all over the world. Initially, many of these systems were developed to alleviate the human work force from some rather routine processes such as automatic payroll systems and inventory tracking systems. These systems, like specialists, performed tasks within their knowledge boundaries.

As computer processing power and storage capabilities continuously grow, computerized information systems, referred to as information system from now on, are becoming more sophisticated. People are beginning to tackle issues surrounding reliability by improving communication; electronic-mailing and newsgroup systems are examples of such communication improvement that have been recently introduced.

2.2 Use of Information Systems in Decision Making Process

In addition to performing routine tasks, Information System can also support the decision making process through two modules: the Decision Support System and the Executive Information System. The Decision Support System and Executive Information System are being applied not only to help users make better decisions, but also to reduce the processing time.

A Decision Support System (DSS) is an information system designed to provide employees access to information crucial to their decision-making processes. The scope of a DSS is rather broad; any system providing its user knowledge can be categorized as a DSS. For example, Tiger Creek, a paper manufacturer, introduced in 1983 an Expense Tracking System (ETS) to allow operators at the mill to make better technical adjustments by studying cost impact information provided by the ETS (Bronsema, 1984).

Frito-Lay Inc., a food distribution giant, developed a DSS through the use of Hand-Held Computers (HHC). This DSS provides delivery and shelving employee better information on how to re-distribute and re-shelf a store. Thus, the employees can not only base their decisions on past experiences, but also obtain accurate and up-to-date account information for the particular store at which they work. Using this system, new employees learn quickly and soon become a productive part of the team (Applegate, 1989).

In addition to their DSS, Frito-Lay also implemented an Executive Information System (EIS). In general, an Executive Information System is an information system designed to

provide senior managers access to information relevant to their management activities. This includes information concerning the company finance and accounting, the employee work-schedule, and the marketing, annual and quarterly reports. Because managers use the information gathered by an EIS to make fast and accurate decisions, companies using a well-designed EIS will gain competitive advantages over their competitors (Leidner, 1993).

Decision Support Systems and Executive Information Systems provide a new way to do business. They have simplified information searches and the presentation of information. As an information intensive and massive decision-based industry, the Environmental Engineering Industry can benefit greatly from both DSS and EIS systems.

2.3 Applications of Information Systems in the Environmental Field

Environmental project management can be a very difficult task because so many factors must be taken into account. Environmental decision making involves understanding not only the immediate impact of human activity on the environment, but also issues like human health, economic costs, current and pending regulation and fairness. In principle, all of these interrelated factors have a bearing on any decision made relating to the environment.

To deal with these complex problems, the environmental engineering industry could greatly benefit by utilizing information technology. In general, there are three domains in which information technology can make a real difference. The first domain is in the modeling of complex environmental processes. Air and water quality modeling are good examples. The second domain is in information management. Integrating information from diverse sources is necessary in order to make sound decisions. Important sources of information range from field-monitored data, to simulation results, to documents on regulatory policy. Finally, the last domain involves modeling the decision process itself and providing the structure and support to enable policy makers to make timely, balanced decisions that are consistent with what we know about the environment.

Satisfying the first criterion of environmental project management, analysis programs available in the market range from air quality modeling tools to groundwater migration modeling tools. In terms of information management, many United States government agencies are

actively developing standardized information systems for storing geographic data, so called Geographic Information Systems (GIS). Using Global Positioning Systems (GPS), GIS databases store information about specific locations using their northings, eastings and elevations.

At present, most environmental engineering Decision Support Systems, which tend to be hybrid systems of modeling and information management, are in the development stage. For example, the International Institute for Applied Systems Analysis has developed a working beta of a DSS named the Decision Support System for Evaluation of River Basin Strategies (DESERT). In a user-friendly environment based upon a Microsoft Windows interface, DESERT provides integration of important stages of decision support including data management, model calibration, simulation and optimization, and presentation of results (Somlyódy, 1996).

Although most DSSs are under development and are therefore not commercially available, most of the Executive Information Systems, which combine all three functions of environmental project management, are only in the conceptual design phase. The Environmental Programs Group at MCNC's North Carolina Supercomputing Center is developing the Environmental Decision Support System (EDSS) that includes all three aspects of environmental project management, making it more like an executive information system. Working closely with the Environmental Protection Agency (EPA), EDSS focuses on a "next-generation" air quality modeling system (Bilicki, 1996).

2.4 Proposed Information System for Preliminary Site Assessment under CERCLA

The global objective of this project was to develop information systems that can support decision making during the complex process of hazardous waste site remediation. An Information System is proposed for Preliminary Site Assessment under CERCLA, which was used as a spedific focus for the project objective. This system has two components as shown in Figure 2.4a.

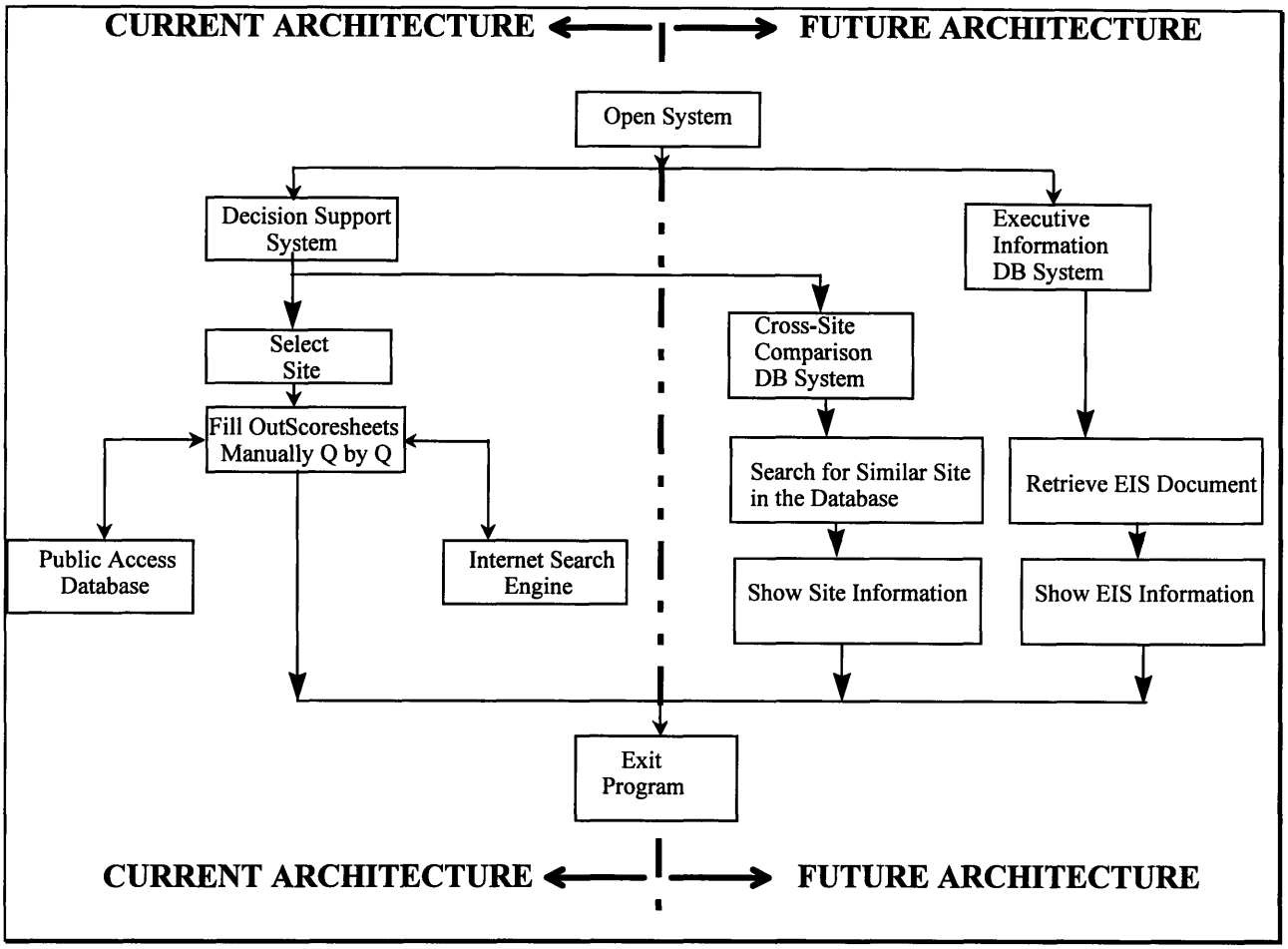


Figure 2.4a Information System for Preliminary Site Assessment

The first proposes a system to assist with the current PA process. The next proposes a system to futher enhance the support the current process. Both processes are discussed below.

2.4.1 Architecture Under current Preliminary Site Assessment Processes

Use of the system requires following a series of steps to complete the preliminary assessment scoresheet. First the user enters general information about a site for a new study or selects an site that is in the process of undergoing preliminary assessment. Then, the user selects which question to answer, and the system accesses the data store that is relevant to that question and returns the answer. Often, the answer will be accompanied by a recommendation for evaluation and a confidence level assigned to the data source thus accessed. The user will then set the confidence level to the answer. In the future, when the DSS and the EIS information systems for the PA are connected, the results of the query will be recorded in a database within

the system. If no answer is available in standard electronic form, the system will access its own intelligent search engine to point out possible storage locations for the information on the Internet.

For example, one of the (multiple choice) questions on the scoresheet is, "Is precipitation in the area heavy?" with possible answers, "Yes," "No," and "Unknown." Asking for the answer to this question will send the system to a database that has precipitation information for the region. In this case, the system will return a numeric value for the average amount of yearly rainfall in the area of that site. This value will be accompanied by a confidence rating for the source and a recommendation (in this case a statistic saying what level of precipitation is considered heavy). The user may then take this recommendation (or not), thereby answering the question, and then assign a confidence level to the answer. The user may also choose not to answer the question at this particular time by indicating that no decision is made, and go on to the next question or exit the system.

When the user reaches the end of the scoresheet, he or she will be able to ask the system tocalculate a score for the site based on the questions answered. The user will then be able to obtain a more permanent copy of the work by printing out the scoresheets.

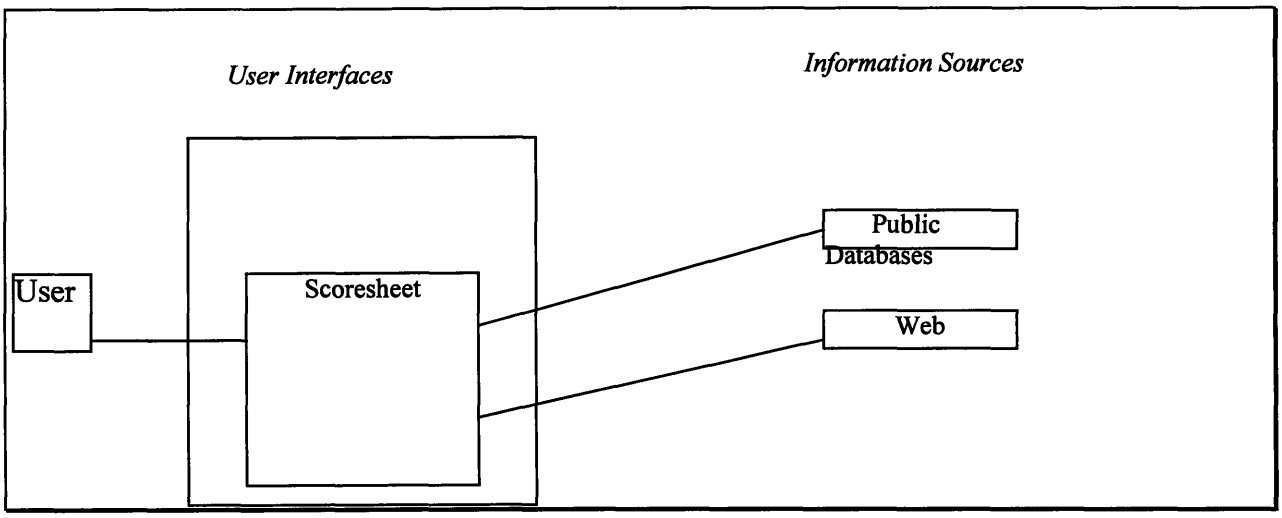


Figure 2.4.1a Flow of Information

The following give the steps to filling out the scoresheets.

1. The user will fill out the scoresheet question by question. The program will provide the relevant supporting information, with an assigned confidence level to each piece of

information.

2. After filling out the scoresheet, the user can SCORE the site.

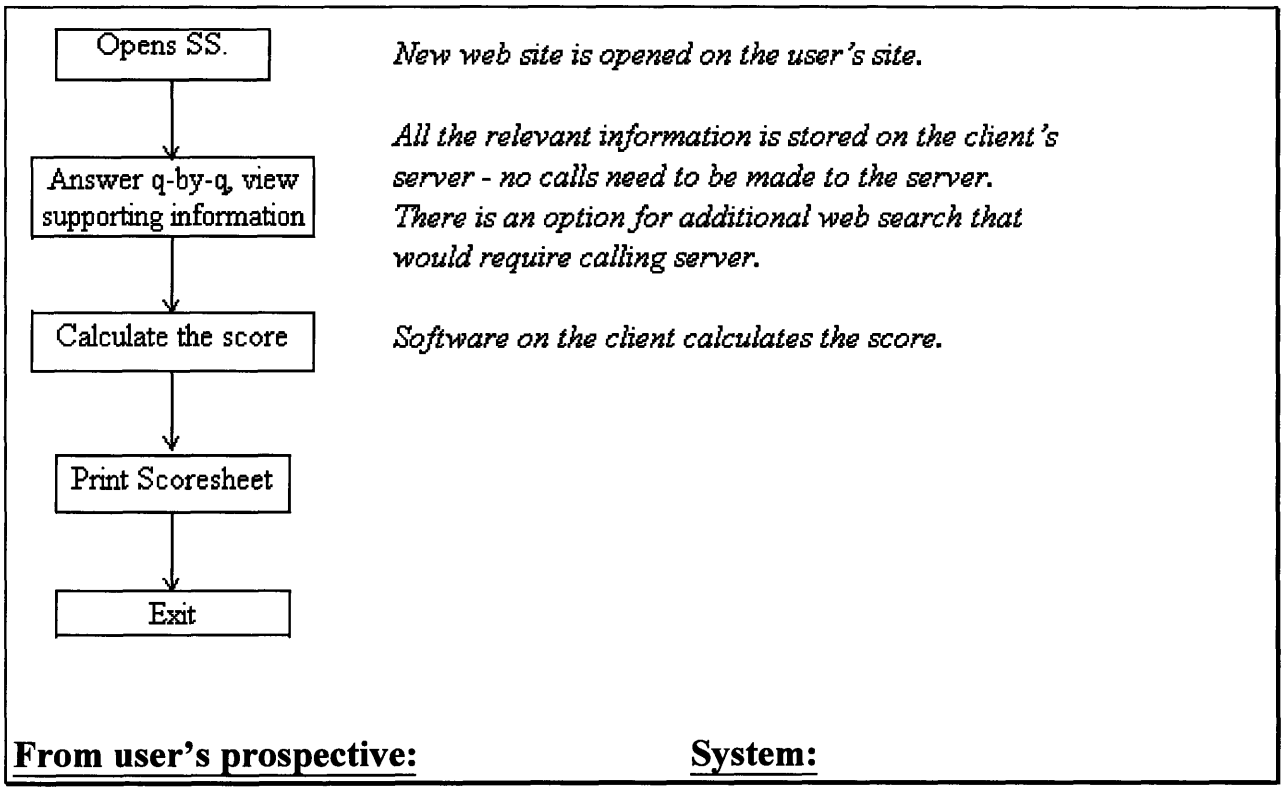


Figure 2.4.1b User's Prospective

2.4.2 Future Architecture of Preliminary Site Assessment Processes

While the proposed information system will greatly improve the current PA process by making information more reliable and readily available, the potential exists for future enhancements to both the PA process and the information system supporting it. Under the current system, information from a variety of electronic sources facilitates the question-by-question completion of the PA scoresheets. In addition to calculating a final numerical score that determines the next stage of the CERCLA process, the information system could be enhanced by somehow reusing the completed site/facility assessments. Once the proposed information system has been used to complete one or more PA's, a database of previous studies will exist. This additional data source could provide valuable comparisons for performing new studies. If a new site has characteristics that are in some way similar to sites that have been studied previously,

completing the PA process may be facilitated by referring to these previous studies. Some characteristics that could be common between sites include geographic location, suspected release of hazardous wastes, and type of facility or operations engaged on the site. Comparing new sites with completed studies would provide an additional check for reliability and support for the current site under assessment.

Expanding beyond the scope of aiding engineers and scientists, the Decision Support System can also incorporate the concept of an Executive Information System to provide better information for senior management. For example, a senior manager could use the system to manage the individual progress of preliminary site assessment of various sites the person oversees. An EIS also opens new opportunities in document tracking and decision management. These new functions allow senior managers, especially the Chief Executive Officer and the Chief Financial Officer, to better monitor and control the company performance and growth rate.

3.0 Results For Architecture Under Current PA Processes

Specifically, the tasks performed during this componenet of the project included the development of a graphical user interface for the existing PA scoresheets, the development of a data store search engine for the Groundwater Criteria Pathway list and the development of a data store search for the Surface Water Criteria Pathway List. The results from each of these tasks are described under separate sections below.

3.1 Graphical User Interface for Existing Preliminary Site Assessment Scoresheet

3.1.1 Overview of the Graphical User Interface of the Electronic Scoresheet

The Graphical User Interface for the electronic scoresheet presents information filtered out from the system that is determined to be useful to the user. The interface for the electronic scoresheet provides (1) better definition of fields to make it easier for the user to fill out clearly stated questions, (2) integration between the parts of the system that provide supporting information, and (3) automation capabilities for calculating the score of the site. Each of these points is described below:

First, the electronic scoresheet has better defined fields than the original paper PA Scoresheet. Many of the long and vaguely posed questions are rewritten in shorter form. New fields are created to simplify answering of the questions. In many instances, a list of possible answers is provided from which the user has an option of selecting one of the choices from the list or entering a new value that is not on the list. Although much reformatting was done to simplify and clarify the electronic scoresheet, at no instance was information was omitted. The electronic scoresheet therefore reduces the complexity of filling out the document.

Second, the Graphical User Interface provides integration of the many parts of the system. It displays results of a query for supporting information from both the external databases and other web based sources. When answering each question, the user has an option of looking up additional information related to that particular question. In the future, the electronic scoresheet will be almost entirely filled out automatically using information filtered out from other parts of the system.

Finally, the electronic scoresheet has capabilities of automatically calculating the score and selecting values from tables based on the user input. This eliminates tedious calculations by hand and an understanding of the layout of the tables.

3.1.2 Structure of the Electronic Scoresheet

Since the current procedure for preliminary site assessment undergoes constant change, the format of the scoresheet is also subject to changes. When developing the current format of the scoresheet document, flexibility was a main factor in determining the structure and technology used to implement the system.

Currently, the scoresheet is broken into parts corresponding to physical aspects of the environment: ground water, surface water, soil, air, in addition to two sections about the site and the source of contamination, and a final conclusion section. Each one of the eight parts of the electronic scoresheet corresponds to one section of the existing scoresheet document. The first part asks for the name and address of the site and the investigator. Based on this information, both external and internal databases are searched for relevant information that could be used to fill out subsequent parts of the scoresheet. This part can be extended to allow logging into a system.

The next two pages contain more detailed general information about the site and the possible source of the contamination. Those pages will be filled out partially by the system, based on the name of the site, and partially by the investigator. Based on this information, an internal database is created that will help with filling out the following parts of the scoresheet. Figure 3.1.2a shows an example of the electronic general information form.

The next four parts will provide the score for individual aspects of the environment, including ground water, surface water, soil pathway, and air pathway. The user will be asked to fill out all of the information that is not filled out automatically by the system. For some of the questions, the user can do an additional search for information. For each of the pages, the calculation of the score is automated, based on the answered questions. Figure 3.1.2b shows an example of one secion of the electronic ground water pathway scoresheet.

The last part integrates all the information from the individual parts and the final score is calculated automatically by the application. Figure 3.1.2c shows the electronic form of the final score section.

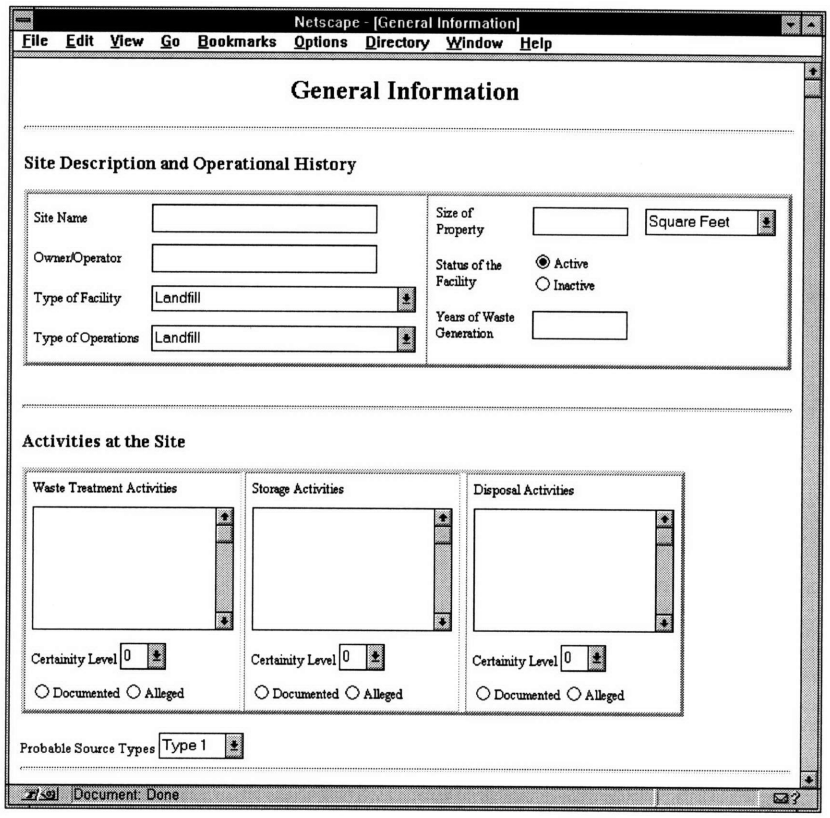


Figure 3.1.2a Interface for General Information Scoresheet

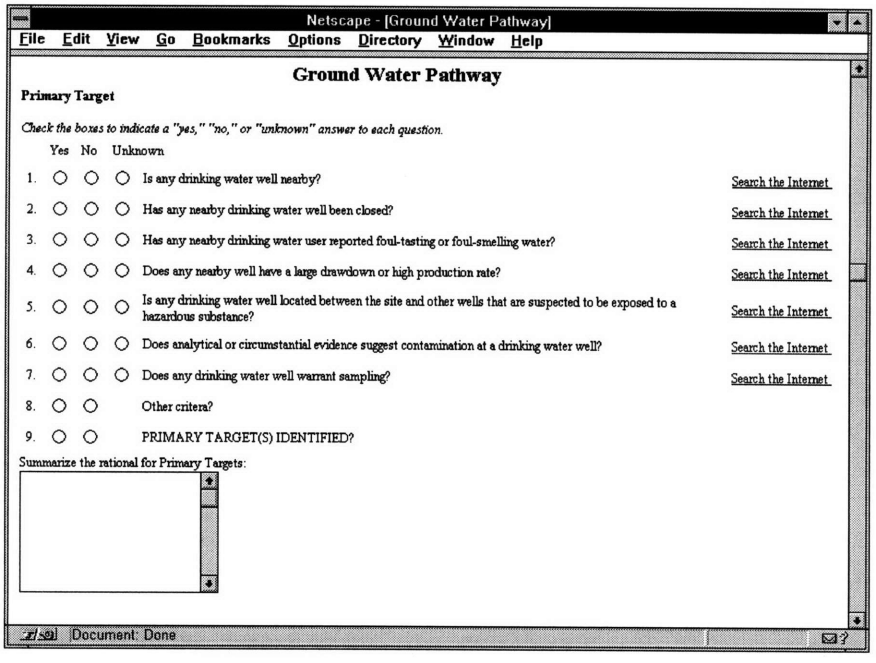


Figure 3.1.2b Interface for Ground Water Pathway

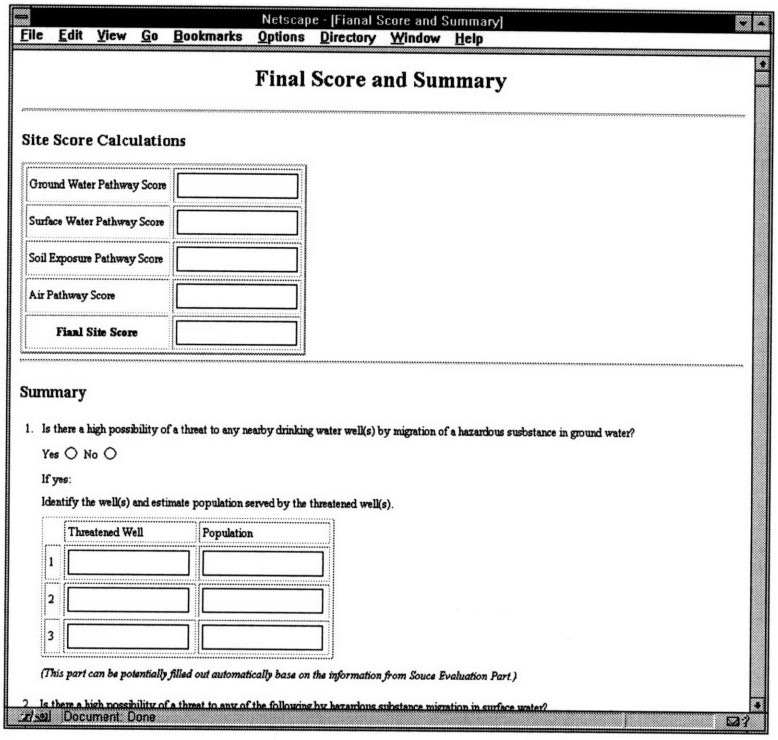


Figure 3.1.2c Interface for Final Scoresheet

3.1.3 IMPLEMENTATION OF THE ELECTRONIC SCORESHEET

Implementation of Graphical User Interface

To provide flexibility to the user and the developer, the electronic scoresheet is a web-based application. This grants platform independence and allows data to be stored on distributed sites. The pages are created dynamically using CGI Perl script to allow (1) filling out default values relevant to the site that are a result of the initial and sequential searches, and (2) provide connectivity between the pages.

In its present state, the user is allowed to fill out the scoresheet only sequentially page by page. The values from the previous sections are carried over to the following sections. The system can be updated to allow the user to browse back and forth between the pages in any order. HTML frames are used to implement this feature. One frame that contains an index of all the pages allows the user to select any page at any instance, while the other contains the actual page.

The fields in each form are a part of the HTML <FORM>. Text fields, check boxes, radio buttons, text areas, and lists are used to present information. Most of the formatting of the

pages is done through usage of the tags <TABLE> and <LIST>.

Upon pressing the "Next Page" button, the values of the fields are sent to CGI bin, where the data is written into temporary files, and a new page is created. Writing information into a temporary file is necessary because the HTML page itself has no capacityto store any information. Thus upon exiting a page, all the information is lost and can not be recovered. Pearl script was used for storing and reading back information to and from the temporary files. This also allows the creation of dynamic pages with information from previous pages using default values provided by the system.

Calculations and Automation of Tables

Java Applets are used to perform calculations on data from fields that are located on the same web page. For example, scores for individual parts are calculated using Java Applets. Parameters are sent from the HTML page to the Java Applet, which performs the operation and displays the results in the Java Applet fields called "Score". Based on the user input, appropriate values from tables are automatically selected and displayed in the result fields.

Java Script would be an alternative technology to Java Applet, but currently it is only supported by Netscape, and thus this application would be limited only to this particular browser.

Temporary Storage of Information

For temporary storage of the information entered on each HTML page, a temporary file is created via CGI Perl script. The web browser has no way of storing information, so each time a web page is exited, all the information that was on it will be lost. To recreate the page again with all the values in the fields, the CGI Perl script reinstates this page using information from the temporary files. Although not currently implemented this same method will also allow the user to go back and forth between the pages in the future when it is anticipated that this feature will be available.

Session Registration and Garbage Collection upon Exit

Using CGI scripts to store temporarily information and to create web pages dynamically to provide connectivity between the pages has a serious drawback, as this does not register sessions of particular users. Thus, if at any time the user decides to exit the application other than when the scoresheet is completed, there is no way to destroy the temporary files automatically. This problem can be eliminated by sending all the information to a central database, which requires building a bridge between the database and the CGI bin.

3.1.4 ALTERNATIVE IMPLEMENTATION SCHEME

Another way to implement the graphical user interface of the system is to use one Java Applet for the entire scoresheet. Passing variables between the various parts of the scoresheet is very easy, as all the variables are stored within one application. Since no temporary files are necessary, the problem of garbage collection upon an unexpected exit could be eliminated, as traces of Java Applets are destroyed when the application suddenly exits.

A serious drawback of using a single Java Applet for the entire scoresheet is its lack of flexibility. Changing parts of the application is complicated, involving adding functions to handle events, and recompiling the source code. Development of such a large size Java Applet is also difficult to debug. Finally, creating the user interface is complicated, as the layout of the web page is harder to control in Java Applet.

3.1.5 FUTURE RECOMMENDATION

In its final form, provided the information is available, the majority of the scoresheet can be filled out automatically, and then it can be viewed and verified by the investigator. Upon submission of the first page, a search for data will return necessary information that would assist with filling out the entire scoresheet. The user will have a chance to request additional information and check the confidence level of the source. This last function is already partially implemented in the current version of the application for a limited number of questions. The next two sections will cover the information search in more depth.

In the future, if the site is to be used as a commercial site, a better way of storing information than in temporary files will have to be created. Also, the problem of garbage file destruction, that will result from an unexpected termination of the application, needs to be addressed.

Creating a web-based application has many advantages. The user is no longer restricted to a particular platform and does not have to worry about installation of the software, provided that he/she already uses a web browser.

For more information on the graphical user interface of this system, refer to Lukasiak, 1990.

3.2 Development of Internet Search Engine for Groundwater Pathway Criteria List

Five years ago, the Internet was a word that many of us weren't even familiar with. Today, the Internet has become an integral part of our business, academic, and social everyday environments. Because it has the potential to be a powerful tool in gathering, sorting, and retrieving data, This explains the importance of the Internet in decision support systems.

Through the use of modern Internet search engine utilities, one can now sort through millions of documents held in a large number of locations around the globe in one single mouse click. There is an increasing number of commercial web sites now available to the public for the purpose of finding documents on the Internet that contain key words or phrases that qualify the information the user desires. Using these search engines in a decision support system gives the user the opportunity to view documentation relating to the current decision at hand, which they may possibly not have had access to in the past.

3.2.1 Availability of Information on the Internet for Completion of PA Scoresheets

For the purpose of completing a PA Scoresheet for an environmental clean up, there are many Internet web sites available to aid in the decision making process. These both general websites[1], containing information which can be applied to virtually any clean-up site, and in some cases specific websites, where information pertaining to only one particular clean-up location can be found. It is important to note that in both cases, documents contained in the website are maintained by the party owning that particular domain and the reliability of information found is often indeterminate.

One example of a general website where non-site-specific information pertaining to environment clean up can be found is maintained by the US Environmental Protection Agency (EPA), "http://www.epa.gov/". This site contains information varying from state and local

[1] A "website" is generally considered a domain location (i.e. www.epa.gov, www.mmr.org, etc...) where any series of "web pages" (actual documents such as index.html, etc...) are located.

environmental protection laws, to educational resources, to links to specific clean-up sites. Because the EPA maintains this website, the information found here can be assumed to be accurate and reliable. However, the final decision with respect to data reliability must be made by the engineer completing the PA Scoresheet. Thus, it may be useful to contact the webserver administrator to verify the status of the information found on a particular website.

3.2.2 Accessing Internet Search Tools Through the Use of Common Gateway Interface

As mentioned earlier, there are many Internet search utilities available free to the public. Some examples of these include Excite (http://www.excite.com), Yahoo! (http://www.yahoo.com), and Alta Vista (http://www.altavista.com). Each of these companies has developed programs that search their extensive databases of URL's (Universal Resource Locators) to return a series of web pages that contain the search string queries entered by users. In each case, the pages returned may vary due to differences in the databases maintained and the search programs created by the different companies. For this reason, it may be desirable to use multiple search engines in order to increase the chances of finding exactly the information required. This is known as "metasearching". An example of this technique can be found at "http://metasearch.com/".

Search engines, such as those listed above, use the Common Gateway Interface (CGI)[2]. CGI protocol allows anonymous users to access and run programs located on their web server and send the information back to the user's web browser. Furthermore, the use of CGI allows variables to be passed to these programs, as in the case of search strings or user names. The usual method for accessing these search programs is through HTML forms where values for each of the variables may be entered and the program may be run with a mouse or key click. Alternatively, one may run the program *directly* by entering the variable names with their values following the URL of the CGI program at the "Go to:" line of your web browser, or through the Open URL dialog box. An example of this is:

"http://search.yahoo.com/bin/search?p=common+gateway+interface".

Knowing how to access these search engines directly, HTML documents can be

---

[2] For more information on Common Gateway Interface, see "http://hoohoo.ncsa.uiuc.edu/cgi/".

generated dynamically that contain links to specific search results pages, not just search engine home pages. Using this one CGI program, the user can pass just one search string and have direct access to results from a variety of commercial search engines. This places a wide variety of Internet documentation relating to their search in one convenient location.

3.2.3 MULTI-KEYWORD SEARCHING

As anyone who has used an Internet search engine most likely knows, the search results returned are sometimes not exactly the results being sought. Often, the user will have to wade through a variety of unrelated web pages to find exactly what it is they are interested in finding. This usually happens because one (or all) of the search words used may also be found in documents pertaining to a completely different subject matter (an example of this is the word "environment", which could pertain to a wide variety of topics). In order to limit the pages returned to only those pertaining to the exact topic being searched, it is often useful to "parameterize" the search.

Mulit-keyword searching involves adding a series of search words to a search string variable that will help to better describe the information sought. It is helpful if the words used are likely not to be found on any site pertaining to a different subject matter. An example of this would be to add the word "groundwater" to a search for the word "environment". Sites pertaining to topics such as "political environment" or "social environment" will most likely not contain the word "groundwater". Therefore, these unrelated sites will not appear at the top of your search results window and the user need not bother wading through countless sites about President Clinton or the newest craze.

Following are two examples from the Ground Water Pathway sheet of the PA Scoresheet showing the questions asked, the call made to run the search program, and the list of keywords used.

Question: Are sources poorly contained?

HTML call: href="../../../scripts/gwp_test.pl?searchstring=MMR%2bgroundwater%2bGround%2bWater%2b Groundwater%2bPlume%2bplume%2bcontamination%2bsource%2bcontained%2bMassachusetts%2bMilitary%2b Reservation%2bwww.mmr.org" target=search

Keywords: MMR groundwater Ground Water Groundwater Plume plume contamination source contained Massachusetts Military Reservation www.mmr.org

Question: Is waste quantity particularly large?

HTML call: href="../../../scripts/gwp_test.pl?searchstring=MMR%2bgroundwater%2bGround%2bWater%2b Groundwater%2bPlume%2bplume%2bwaste%2bquantity%2bMassachusetts%2bMilitary%2bReservation%2bwww .mmr.org" target=search

Keywords: MMR groundwater Ground Water Groundwater Plume plume waste quantity Massachusetts Military Reservation www.mmr.org

Figure 3.2.3a Example of Internet Search Engine

3.2.4 DISPLAYING RESULTS

As mentioned above, links to search results will be displayed in a web browser window in HTML format. Because users will need to run the program many times (perhaps once for each question answered), and will then need to return to the PA Scoresheet document window to record their decisions, it is inconvenient to use the same browser window for both the scoresheet and the search results. For this reason, when the search program is run, a new browser window will be opened. This allows the engineer to follow long search paths without the hassle of going back to the original PA Scoresheet document. Figures 3.2.4a and 3.2.4b show screen captures of a theoretical user session, one with just the scoresheet browser window open, and one with both scoresheet and search windows open.

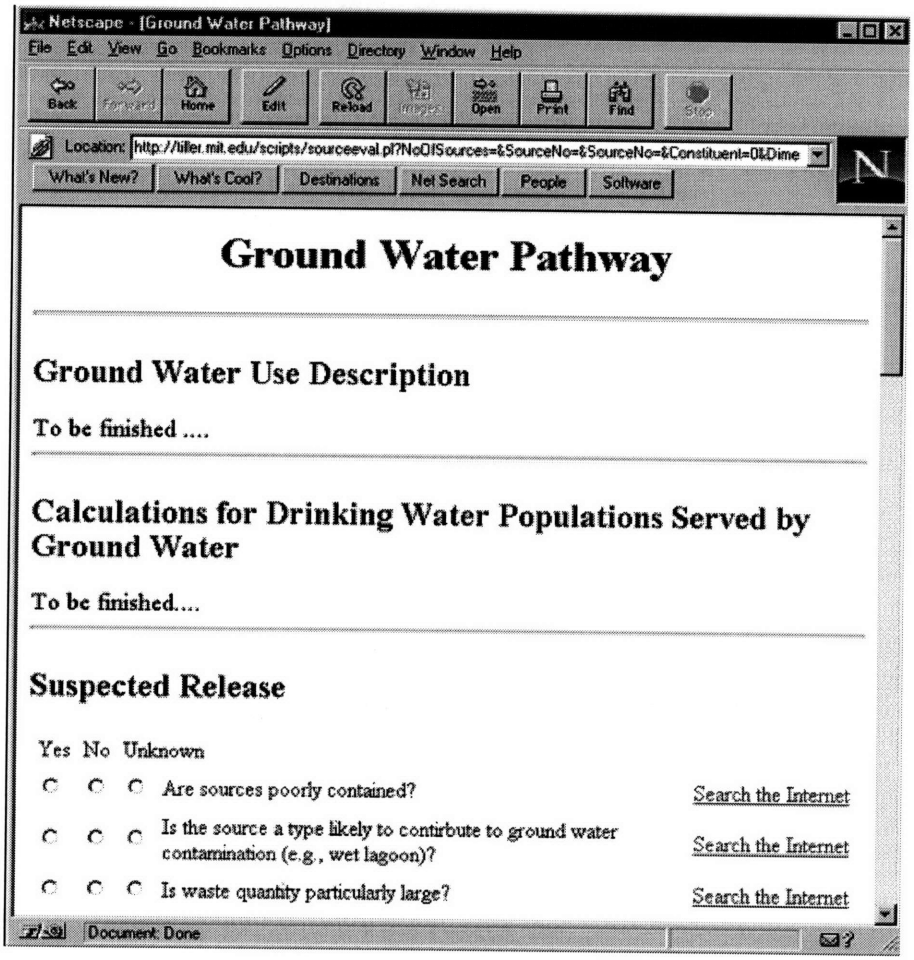


Figure 3.2.4a Ground Water Pathway Section of Electronic PA Scoresheet

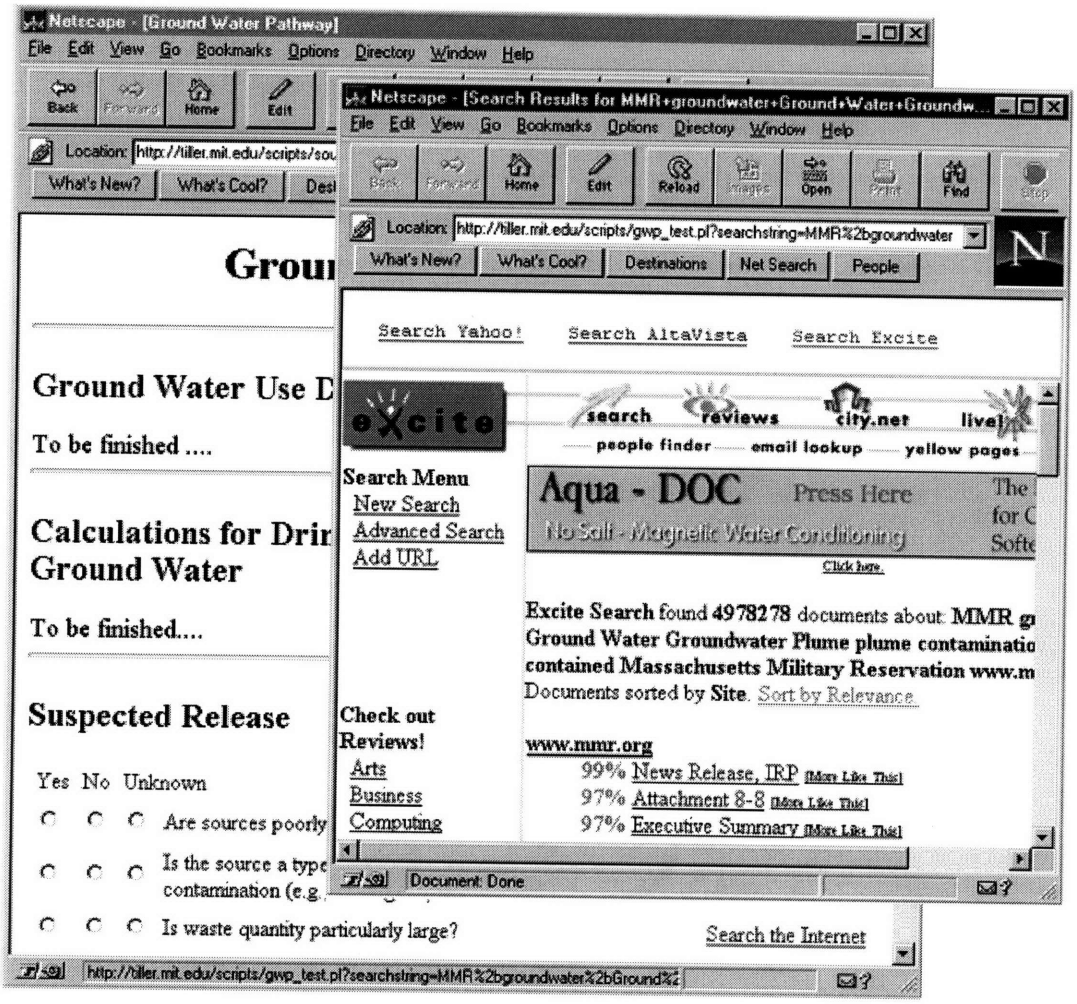


Figure 3.2.4b Electronic PA Scoresheet with Search Results Window

3.3 Development of Data Store Search Engine for Surface Water Pathway Criteria List

3.3.1 The Data Store Search Engine's Role in the System

The data store search engine contains static links to data sources that are in a known, standard format. This section of the system is diagrammed in Figure 3.3.1a, which is a cropped section of Figure 2.4.1a. PA Scoresheet questions that can be answered by these data sources are marked with a "Query" button placed next to them on the electronic scoresheets. This button initiates the search of the appropriate data sources. These include data sources that are in a parable standard format on the World Wide Web, on a connectable CD-ROM, or any other source where the information is in a format that allows the computer to extract specific

information from the database without the user's help. Generally, this means that the data is held in a spreadsheet or database format, as opposed to a written document, or a less formatted information list.

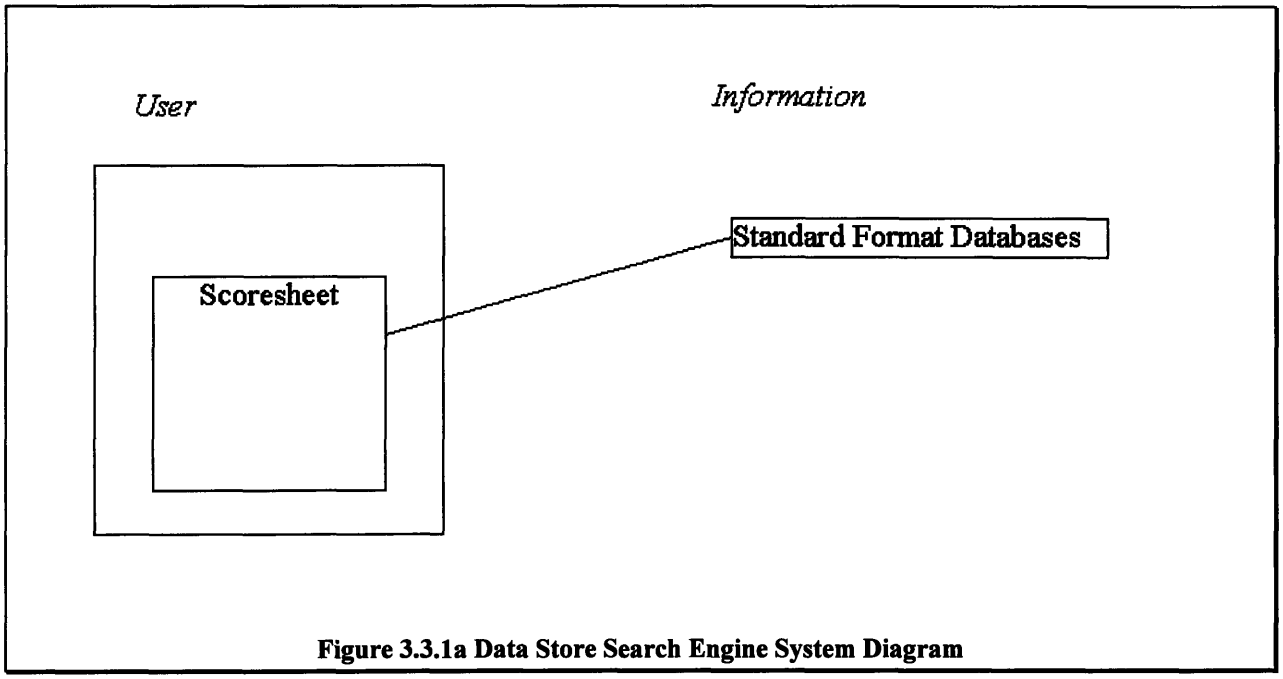


Figure 3.3.1a Data Store Search Engine System Diagram

The extent of the implementation of the data store search engine described here was limited by the current availability of documents containing PA Scoresheet information in a standard format. The environmental consultant on this project, Kenneth Till, was able to locate one such data source; a United States Geological Survey (USGS) web site that provides water-use information for fifty states in the US. For this project, the USGS data source file for the state of Massachusetts was connected to the PA Scoresheets for the MMR.

3.3.2 DESCRIPTION OF USGS DATA SOURCE

Each of the available USGS data source files contains information for a particular state in the United States. The data were gathered in 1990, and placed in standard text files written in spreadsheet format, with each row corresponding to a county in that state. Although the file currently connected to the PA Scoresheet is for the state of Massachusetts, combining the different state files into one countrywide file for conducting more generic queries would be an easy task, if this system were to be used for different Superfund sites in the future.

The column headings of the data file are codes for the water-use data elements present. A brief description of the file format is in Appendix B and a complete list of the water-use data elements along with their codes and descriptions is given in Appendix C. The elements used in the current system, together with the PA Scoresheet questions contained in Appendix A that they have been used to answer are as follows:

- *ps-popgw* (total population served by the ground water in the area), used to answer questions 3 and 4 on page 8 of the PA scoresheet
- *to-totsw* (total surface water used in millions of gallons per day), used to answer the first question in the "Suspected Release" column of page 11 of the scoresheets
- *do-sstot* (total domestic water withdrawals), used to answer the first question in the "Primary Targets" column of page 11 of the scoresheets
- *ps-popsw* (total population served by the surface water in the area), used to answer questions 4 and 5 on page 12 of the scoresheets

3.3.3 Matching Existing Data Source Information with Scoresheet Questions

Many of the questions on the scoresheet are currently unanswerable from existing data sources. However, much of the required information could one day be compiled into a spreadsheet or a database format.

For the system developed here, the scoresheet questions that have been linked to the existing data stores will be distinguishable by a "Query" button placed next to them. Pushing the button will open a new browser window that will display the results of the query. The results of any queries run from that point forwards will also appear in the same browser window.

When the answer is received from the data source, some of the PA Scoresheet questions require interpolation on the part of the user in order to translate the answer into relevant scoresheet information. For example, with the question "Is surface water nearby," (question 1, column 1, page 11), the system currently returns the amount of surface water used in the area. If this number is greater than zero, than the user will answer, "Yes," and if not, "No." Obvious answers such as this are not automatically filled in by the system, however, inorder to ensure that the user takes an active role in answering the question.

When the user pushes this "Query" button, they are actually executing a program that is on the same server as the web page they are viewing. As mentioned in Section 3.3.2, the system currently accesses a USGS file containing Massachusetts's water-use information. It is envisioned that when the system is complete, the user will log on to a particular account and select the Superfund site that they wish to assess. This selection will automatically tell the system which data files to access. The particular button pushed will tell the system which question is being asked, and the program will then determine how to run the query for that particular piece of information. The query will be run on the data source (also located at the site). The information will be parsed into HTML so the web browser can read it, and post it for the user.

It is important to note that for a particular scoresheet question, the program to execute the query and the data source components of the system must be on the same server. However, this server does not need to be the same server as that used for the main system user interface. This means that if there is an organization that maintains a standard format data source, the program to access their data will need to be on their server, otherwise their data files will need to be downloaded to another server. If the source is public, there should not be a problem bringing

data into a server controlled by EIDSS administrators. If it is a private data source, permission must be obtained to access it. Once the permission is obtained, the relative locations of the data source and the program can be easily placed as required.

3.3.4 Conclusion

In general, the nature of the information requested on the scoresheets is not conducive to being placed in a database. It is far more likely that if data sources are created in the future, whether they are text files on the web, or CD-ROMs, they will be in spreadsheet format, as the one currently used to demonstrate the development of a data source search engine here. The demonstration system that has been implemented in this project could be expanded to access other text data sources very easily, with the addition of approximately ten lines of code. This expansion will be simple, because the entire framework required to make this type of connection has been completed. If data sources of other types are found, it is estimated that the amount of code required to incorporate them will be similar. However, there will be some additional configuring that the system administrator would need to perform. The difficulty of this process will vary with the type of data source to be included in the search engine.

For more information on the the development of the data store search engine for the surface water criteria list, refer to Mukhopadhyay, 1997.

4.0 System Architecture For the Interactive Preliminary Assessment Scoresheet Database

4.1 Overview

4.1.1 Motivation

The feasibility of using databases located on the Internet and from local CD-ROM to complete preliminary site assessment scoresheets is limited by the actual availability of information from these data sources. One of the most useful sources of data may be previously completed PA scoresheets. Once the proposed information system discussed in this project has been used to complete one or more PA's, a database of previous scoresheets will be available for site comparisons. If a new site under investigation has characteristics that are in some way similar to sites that have been studied previously, completion of the new PA may be facilitated by referring to the database of these previous studies. Comparing new sites with completed studies will provide an additional check for reliability and support for the current site under assessment. This section summarizes the proposed architecture for an interactive database of PA scoresheets for cross-site comparison. This architecture will enhance the utility of the currently proposed decision support system. The database of scoresheets will additionally be used for implementation of the executive information system as well, which may be referenced in Kuo, 1997.

4.1.2 Design and Population of Data Storage

In order to take full advantage of the completed PA scoresheets, some data storage must be developed to keep track of the completed studies. A relational database, which is a set of data tables linked by common data fields, is well suited for this purpose because it can both adapt to change and is efficient when good design principles are followed (Hawryszkiewycz, 1990).

Any number of commercially available relational databases can be used to develop the tables that store the completed PA scoresheets. As more PA scoresheets are completed and entered into the database, more information will be available for site comparisons and the support of completion of other PA scoresheets. The best way to populate the database is to provide a direct connection between the proposed information system that will allow storage of user input provided over the Internet. The ease of accomplishing this task will depend on

development of current and new technologies. For example, both Microsoft Access '97 and Oracle provide interfaces to allow connection to their databases through the Internet, but they are limited by the flexibility they provide with respect to the manipulation of the databases and are difficult to set up.

4.1.3 Database Queries and Comparison Criteria

The database should be designed with particular queries or retrieval of information from the interactive database in mind. Knowing what kinds of questions will be "asked" of the database will affect the design of the tables and data fields. At a minimum, the database should allow easy viewing of the various sections of the PA sections. These include General Site Information, General Facility Information, Source Evaluation, Ground Water Pathway, Surface Water Pathway, Air Pathway, Soil Pathway, and Scoring. The user should be able to select a facility and site and then be able to view any of the mentioned sections.

The real benefit of the database will be derived from its ability to perform comparisons between sites based on specific criteria such as geographic location, types of hazardous wastes, type of facility or operations, and types of waste sources.

The most flexible option will be the selection of any type and number of parameters for performing site comparisons. For example, sites could be ranked by both geographic location and the type or number of similar hazardous wastes. An additional option of comparing sites by individual sections (ground water, surface water, air, and soil pathways) will allow ranking of sites based on answers within these sections of the PA.

4.1.4 USING THE SYSTEM

To illustrate how the interactive database could be utilized, consider the case where the geographic location is used as a parameter for comparison of sites. A geographic location specified by latitude and longitude coordinates would first be entered. A query would then be run outputting a listing of sites ranked by proximity to the specified geographic location. Alternatively, the user could select to first choose the geographical location of a particular site. The resulting query would return a specific number of sites ranked by proximity to the chosen site. The use of the type of facility or operations, types and sources of wastes could be used in similar fashion. The user explicitly inputs a value for which the sites are ranked, or the user selects a site on which to run a comparison. In the case where comparisons are based on individual sections of the PA scoresheet, such as the ground water pathway or surface water pathway, a completed or partially completed PA will have to be selected first as the basis for the comparison. The result of all of these queries will be a listing of the most similar PA scoresheets. The user will then be able to select any of the sites returned by the query and then view the contents of the completed PA for the selected site.

4.1.5 MAKING USE OF THE RESULTS

If a facility is located very near some other site or has been contaminated by the same hazardous substances, or is similar in any other way to another sites, then the user may find it useful to peruse the information that was used to fill out the completed scoresheets of previous sites. The most useful information may be found in the "memo" sections contained in each of the major sections of the PA scoresheet. These sections, which give further explanation as to the rationale behind answering some of the questions, may cause whomever is responsible for completing the current PA scoresheet to think of some factors that are not explicitly covered by the PA scoresheet questions themselves. Additionally, where sites share particular characteristics, the scoresheets can be compared to see if the data are reasonable. For example, if two sites were located very near each other, it would be expected that the population distribution recorded in both scoresheets should be similar as well.

4.1.6 Interactive Database Organization

The rest of this chapter describes the system architecture for building an interactive database of preliminary assessment scoresheets for cross-site comparison. The information given here is intended to be independent of the system implementation, while Chapter 5 provides an example implementation of the system using the relational database, Microsoft Access 95.

In what follows, an explanation of the basic fundamentals of relational database design followed by a description of the preliminary assessment scoresheet database. An introduction to the Structured Query Language is then presented to provide an understanding of how to use the database in order to make site comparisons. Finally, this description of the system architecture is concluded by explaining the Structured Query Language statements that would be used to perform the site comparisons in the database.

4.2 Information System Design - Relational Databases

Although the preliminary assessment database is based on a standard government form, it is important to allow the flexibility to modify the design both during the development of the database and in the future. It is equally important to have an efficient design for storage and retrieval of data. The relational database model can both adapt to change and is efficient when good design principles are followed (Hawryszkiewycz, 1990). This section will discuss how to design a good relational database. Section 4.3 will describe the design of the preliminary assessment scoresheet database.

4.2.1 A Non-relational Database Example

If someone were asked to design a way to store information about a site that was undergoing preliminary assessment, part of the database might look like Figure 4.2.1a.

PreliminaryAssesment Table

Site Name	Site Address	Facility Name	Facility Address	Hazardous Waste #1	Hazardous Waste #2
MMR	3450 Herbert Rd. Bourne, MA 02542	LF-1	6787 Worcester St. Barnstable, MA 02541	TCE	1,2 DCE
MMR	3450 Herbert Rd. Bourne, MA 02542	SD-5	8984 Newbury St. Barnstable, MA 02541	PCE	EDB
MMR	3450 Herbert Rd. Bourne, MA 02542	FS-12	4444 Hayway Rd Falmouth, MA 02541	TCE	PCE
Oroville Army Airfield	4500 Santiago St. Oroville, CA 95966	LF-1	6775 Butte Wy. Oroville, CA 95966	TCE	EDB
Oroville Army Airfield	4500 Santiago St. Oroville, CA 95966	CS-10	7456 Chico St. Marysville, CA 95968	VCL	TCL

Figure 4.2.1a Non-relational Database model for Preliminary Assessment

In this example, the site name and location is followed by the name and location of a facility located on the site and a list of hazardous wastes suspected to have been released in this facility. The assumed relationships are that sites can contain different facilities, while facilities contain a number of suspected hazardous wastes. The next section will explain why this is not an efficient way to store the information about the preliminary assessment in this form, but this example introduces a couple of basic concepts about storing information in a database. Figure 4.2.1a is an example of a "table" consisting of a set of columns and rows. A column represents a data element present in the table, while a row represents an instance of a record, or entry, in a table. Site Name, Site Address, and Facility Name can be thus be referred to as data elements or fields. Each data field holds the same data type. Both Hazardous Waste elements, for example, may only contain abbreviations for some kinds of hazardous waste. Even without employing a relational database model, and without a clear understanding of good design principles, however, this database is deficient in a number of ways.

The rest of this section presents the basic database design techniques to improve the current database design. Figure 4.2.1 is an example of a very rigid database design, where data is stored redundantly and data fields are more complex than they should be. When designing a relational database, any redundant storage of data is to be avoided. By using multiple tables connected by certain relationships, the database is allowed to maintain its flexibility. Flexibility in queries is achieved by keeping the data fields simple. Lastly, certain techniques are used to

ensure the "data integrity" of the database.

4.2.2 Avoiding Data Redundancy

One of the goals of good database design is to minimize the of storage of redundant data. There are two reasons why this is desirable. The overall size of the database increases as the number of records in the database increases. This makes the management of information more difficult. Thus, performing searches and retrieving specific pieces of information takes more time and consumes greater computer memory and processing resources. Secondly, if data is stored in more than one location, any modification of the data will be necessary for every instance of the data. For this reason, it is highly desirable to store information only once.

In Figure 4.2.1a, the name and address of the site is recorded in each row or record in the table. Recording this information is not only inefficient in terms of trying to minimize the size of the database, but if the site addresses had been entered incorrectly or if we wanted to refer to the site names by an alternate name, every instance of the site name and address would have to be changed. This could be quite an unwieldy task if there were many records stored in this database.

The solution to this problem is to break the table into two tables that are connected by a *key*. Figure 4.2.2a shows the improved design which includes one more data field called SiteID.

SiteInformation

SiteID	Site Name	Site Address
1	MMR	3450 Herbert Rd. Bourne, MA 02542
2	Oroville Army Airfield	4500 Santiago St. Oroville, CA 95966

Link Between tables

FacilityInformation

SiteID	Facility Name	Facility Address	Hazardous Waste #1	Hazardous Waste #2
1	LF-1	6787 Worcester St. Barnstable, MA 02541	TCE	1,2 DCE
1	SD-5	8984 Newbury St. Barnstable, MA 02541	PCE	EDB
1	FS-12	4444 Hayway Rd Falmouth, MA 02541	TCE	PCE
2	LF-1	6775 Butte Wy. Oroville, CA 95966	TCE	EDB
2	CS-10	7456 Chico St. Marysville, CA 95968	VCL	TCL

Figure 4.2.2a Avoiding Redundancy -Break Table into Two Separate Tables

It is through the SiteID that the information between the two tables is linked. Now when changes need to be made to the information in the site table, they only need to be performed once since the information resides in a single location. The concept of the key is central to relational database design and provides the solution to the problem of redundancy. In the SiteInformation table, the SiteID is called the *primary* key because its value uniquely identifies every record in the SiteInformation table. That is, there is only one occurrence of each value of the SiteID. The site address could also be used as a primary key. The choice of which data field is identified as the primary key is largely arbitrary, but it is generally more convenient to use a simpler data field. Links between tables are made between primary keys in one table and *foreign* keys in another. In the above example, the SiteID in the FacilityInformation is the foreign key.

In the current design, the FacilityInformation does not contain a convenient primary key. SiteID is not a primary key because there are multiple instances of the same value (e.g. there are three 1's and two 2's). Nor can the Facility Name be used as a primary key since there are two instances of LF-1. The address could be used as the primary key, but, as was mentioned earlier, simple data fields are more convenient to provide links between tables. The combination of two columns from the FacilityInformation table could uniquely determine its records, such as SiteID

and Facility Name, but it would be simpler if one column could be identified as the primary key so that other tables could be linked to the FacilityInformation table. One common way to introduce a primary key into a table is to simply add another data field containing unique values. The data field Fac/SiteID has been added to the FacilityInformation table in Figure 4.2.2b.

FacilityInformation

Fac/SiteID	SiteID	Facility Name	Facility Address	Hazardous Waste #1	Hazardous Waste #2
1	1	LF-1	6787 Worcester St. Barnstable, MA 02541	TCE	1,2 DCE
2	1	SD-5	8984 Newbury St. Barnstable, MA 02541	PCE	EDB
3	1	FS-12	4444 Hayway Rd Falmouth, MA 02541	TCE	PCE
4	2	LF-1	6775 Butte Wy. Oroville, CA 95966	TCE	EDB
5	2	CS-10	7456 Chico St. Marysville, CA 95968	VCL	TCL

Figure 4.2.2b Introduction of Primary Key

The tradeoff, however, is that this requires more data storage at the expense of the ease of linking tables.

4.2.3 Establishing Table Relationships

When multiple tables are used and linked through keys, it then becomes necessary to establish the type of relationships between the tables. Figure 4.2.3a shows the relationship between the SiteInformation and FacilityInformation table.

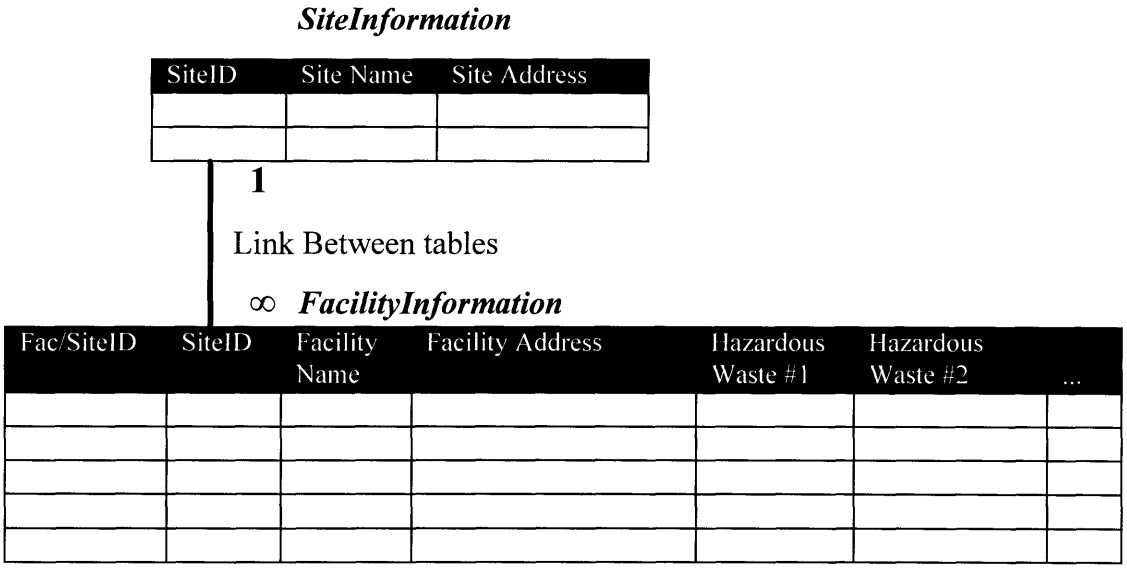


Figure 4.2.3a Many-to-One Relationship

The "one" and "infinity" markers indicate that for every site there can exist multiple facilities. The other type of relationships that are possible are *many-to-many* and *one-to-one*. An example of a one-to-one relationship will exist between the FacilityInformation table and the GroundWater-Scoresheet table. Assuming that each facility only undergoes one preliminary assessment, there will only be one table that holds the data related to completing the Ground Water Scoresheet. Figure 4.2.3b shows the one-to-one relationship where each table contains the same values for the Fac/SiteID.

FacilityInformation

Fac/SiteID	SiteID	Facility Name	Facility Address	Hazardous Waste #1	Hazardous Waste #2
1	1	LF-1	6787 Worcester St. Barnstable, MA 02541	TCE	1,2 DCE
2	1	SD-5	8984 Newbury St. Barnstable, MA 02541	PCE	EDB
3	1	FS-12	4444 Hayway Rd Falmouth, MA 02541	TCE	PCE
4	2	LF-1	6775 Butte Wy. Oroville, CA 95966	TCE	EDB
5	2	CS-10	7456 Chico St. Marysville, CA 95968	VCL	TCL

1

1 ***GroundWater-Scoresheet***

Fac/SiteID	Primary Score	Secondary Score	WHPA Score	T Score
1	50	0	15	60
2	0	0	15	20
3	50	25	35	80
4	100	50	20	0

Figure 4.2.3b One-to-One Relationship

The one-to-one relationship between these tables means that they could actually be joined into one larger table, but it is considered better design conceptually to keep related data items together in separate tables.

4.2.4 MAINTAINING FLEXIBILITY

Using the updated tables shown in Figure 4.2.2a, the FacilityInformation table could be improved even further. The FacilityInformation table currently contains two data fields for recording the names of hazardous wastes suspected to have been released at the facility. How would the data be stored if the facility contained only one hazardous waste, or more than two hazardous wastes? In the first case, the second column could just be left blank resulting in wasted space. In the case where there are more than two hazardous wastes, an additional record with the same Fac/SiteID, Facility Name, Address would be entered. As discussed earlier, we want to avoid storing redundant data wherever possible, so a better method should be devised to handle

both of these situations.

The way to do this is to break the FacilityInformation table into two separate tables.

FacilityInformation

Fac/SiteID	SiteID	Facility Name	Facility Address
1	1	LF-1	6787 Worcester St. Barnstable, MA 02541
2	1	SD-5	8984 Newbury St. Barnstable, MA 02541
3	1	FS-12	4444 Hayway Rd Falmouth, MA 02541
4	2	LF-1	6775 Butte Wy. Oroville, CA 95966
5	2	CS-10	7456 Chico St. Marysville, CA 95968

1

∞ ***FacilityHazardousWaste***

Fac/SiteID	Hazardous Waste
1	TCE
1	1.2 DCE
1	**PCE**
2	**EDB**
3	TCE
3	PCE
4	TCE
4	EDB
5	VCL
5	TCL

Figure 4.2.4a Making Tables More Flexible

Notice that in the FacilityHazardousWaste table that LF-1 now holds three hazardous wastes, while SD-5 only has one recorded hazardous wastes. There are no blanks spaces and no redundancy of data. Only by splitting the table was this flexibility allowed. Again, the tradeoff is that the database now contains an extra table where the Fac/SiteID needs to be recorded for every record.

4.2.5 KEEPING THE DATA FIELDS SIMPLE

The facility address can be broken down into street address, city, state, and zip code. This will prove to be useful when performing queries on the data. A complete description of queries is discussed in the next section, but the definition of a query as simply retrieving the values of specified fields satisfying specified conditions will be sufficient for now. If the user wanted to review the list of facilities located in a particular city, a data field containing the city name would need to exist. Figure 4.2.5a shows this breakdown of the FacilityInformation field Facility Address.

FacilityInformation

Fac/SiteID	SiteID	Facility Name	Street Address	City	State	Zip
1	1	LF-1	6787 Worcester St.	Barnstable	MA	02541
2	1	SD-5	8984 Newbury St.	Barnstable	MA	02541
3	1	FS-12	4444 Hayway Rd	Falmouth	MA	02541
4	2	LF-1	6775 Butte Wy.	Oroville	CA	95966
5	2	CS-10	7456 Chico St.	Marysville	CA	95968

Figure 4.2.5a Keeping Data Fields Simple

The principle of reducing data fields into their simplest elements should be applied to data fields that contain the processing of any other fields as well. There is no need, for example, to keep any kind of "total" data fields if the values are the result of performing some mathematical operation on other data fields. When all data fields contain only one piece of data the table is referred to as having achieved *first normal form* (Hawryszkiewycz, 1990).

4.2.6 OTHER RELATIONAL DATABASE DESIGN PRINCIPLES

There are two additional issues involved in designing a good relational database. Both deal with relationships between tables. The enforcement of *referential integrity* refers to ensuring that the value entered into the foreign key of one table actually exists as one of the values of the primary key in another table. For example, the enforcement of referential integrity would not allow the record with Fac/SiteID in Figure 4.2.6a to be entered into the FacilityInformation table based on the records contained in the Site Information table.

SiteInformation

SiteID	Site Name	Site Address
1	MMR	3450 Herbert Rd. Bourne, MA 02542
2	Oroville Army Airfield	4500 Santiago St. Oroville, CA 95966

1

∞ ***FacilityInformation***

Fac/SiteID	SiteID	Facility Name	Street Address	City	State	Zip
1	1	LF-1	6787 Worcester St.	Barnstable	MA	02541
2	1	SD-5	8984 Newbury St.	Barnstable	MA	02541
3	1	FS-12	4444 Hayway Rd	Falmouth	MA	02541
4	2	LF-1	6775 Butte Wy.	Oroville	CA	95966
5	2	CS-10	7456 Chico St.	Marysville	CA	95968
6	**3**	CS-4	6666 Invalid Wy.	Mistake	MA	02541

Figure 4.2.6a Enforcing Referential Integrity

Because there is no record in the SiteInformation table with a SiteID of 3, this is an invalid entry. Enforcing referential integrity is a specific way to ensure the data integrity of the database.

The second relationship issue involves eliminating many-to-many relationships. When many-to-many relationships exist between tables, it becomes impossible to avoid redundant storage of data. This will become clearer after review of the structure of a set of tables where the many-to-many relationships have been replaced by one-to-many relationships. Figure 4.2.6b shows such an example.

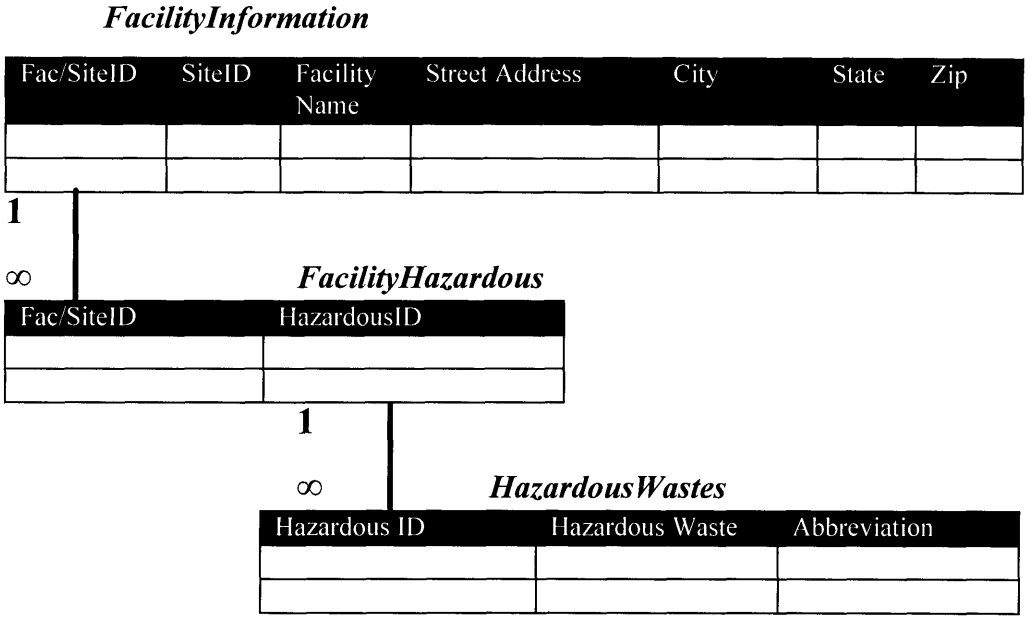


Figure 4.2.6b Eliminating Many-to-many Relationships

In Figure 4.2.6b, the table FacilityHazardous has been created to record the hazardous wastes contained at each facility. This table eliminates the many-to-many relationship that previously existed betweeen the FacilityInformation table and the HazardousWastes table. To see why the FacilityHazardous table was created, imagine, that the table HazardousWastes was directly connected to the table FacilityInformation. Some data field must be created to act as the key that would join the two tables. In the first solution, a HazardousID could be added to the FacilityInformation table. The problem however is that this does not allow the recording of multiple hazardous wastes contained on one facility without rerecording all of the same facility information for each hazardous waste. The alternative would be to connect the two tables through a Fac/SiteID field added to the Hazardous Wastes table. However, the same problem of storing redundant data would occur whenever the same hazardous waste was stored at different sites. A record would be needed for each facility even though the hazardous waste information (stored in the data fields HazardousID, Hazardous Waste, and Abbreviation) would stay the same.

With an understanding of the principles of relational database design, the specific design of the preliminary assessment scoresheet database can now be presented.

4.3 Database Design for Preliminary Assessment Scoresheets

The Preliminary Assessment Scoresheet is broken into seven sections, as seen in Figure 4.3a.

Preliminary Assessment Scoresheets	
A. General Information	E. Soil Exposure Pathway
B. Source Evaluation	F. Air Pathway
C. Ground Water Pathway	G. Site Score Calculation
D. Surface Water Pathway	

Figure 4.3a Sections of the Preliminary Assessment Scoresheet

The architecture presented here does not specifically include the last three sections, Soil Exposure Pathway, Air Pathway, or Site Score Calculation. The Site Score Calculation includes information that is entirely based on data recorded in other tables, so does not require the inclusion of any tables devoted to Site Score Calculations. The Soil Exposure Pathway and Air Pathway sections are very similar to the Ground Water Pathway and Surface Water Pathway sections. Providing descriptions of all four tables would mean a great deal of redundancy in anexplanation of the design. Therefore, only the Ground Water Pathway and Surface Water Pathway are discussed in this report. The flexible nature of relational databases, however, easily allows the addition of further tables.

This section discusses the design of the tables, data fields, and relationships between tables for the preliminary assessment scoresheets. Refer to Appendix A to see the actual Preliminary Assessment Scoresheets. Figure 4.3b shows the complete set of tables, data fields, and relationships between tables. This section explains the rationale for the architecture presented in Figure 4.3b.

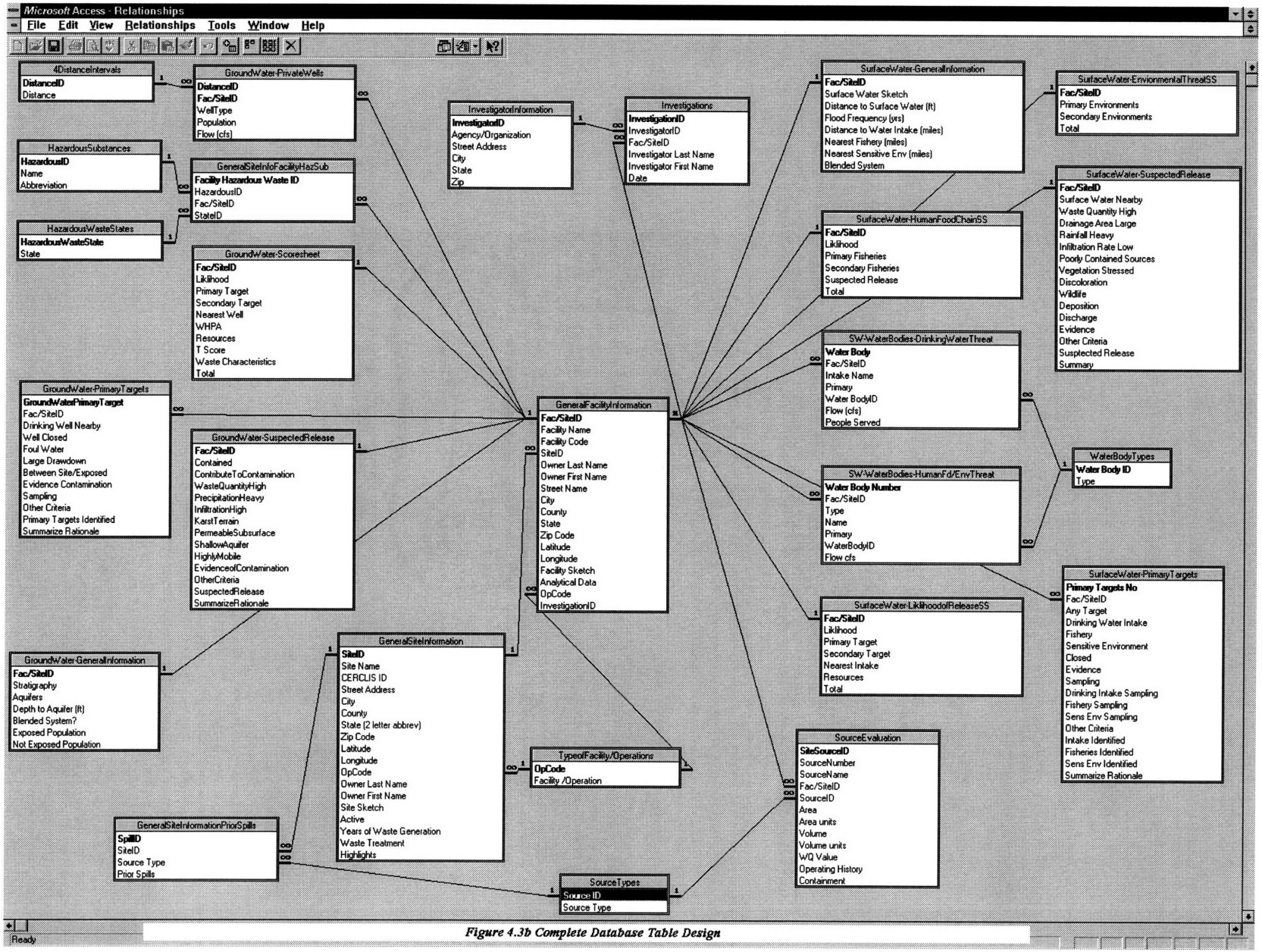


Figure 4.3b Complete Database Table Design

The database consists of sets of general information tables, tables specific to the various pathways, and supporting tables that restrict values to be entered in certain data fields. All of these tables will be explained in the remainder of this section.

A number of links shown in Figure 4.3a go through other tables. These "links" are not meant to infer any type of relationship between these tables. The links pass through other tables simply because of space constraints in the figure.

4.3.1 GENERAL INFORMATION TABLES

The General Information section of the PA scoresheet is used to provide a site description, operational history, and discuss probable waste substances of concern. Although the PA Scoresheet contains a General Information section that is not particular to a "site" or "facility", the database has been designed to accommodate the existence of sites that contain numerous facilities. The Massachusetts Military Reservation contains the facilities, Ashumet Valley, LF-1, and SD-5, for example. The one-to-many relationship that exists between the GeneralSiteInformation and GeneralFacilityInformation tables is shown in Figure 4.2.b through the common data field SiteID. Referential integrity should be enforced between these tables. That is, no SiteID should be entered in the GeneralFacilityInformation table that does not correspond to a site record in the GeneralSiteInformation table. The GeneralFacilityInformation and GeneralSiteInformation tables do not contain exactly the same data fields since some data fields are unique to each site, thus do not need to be stored along with each facility record.

The GeneralFacilityInformation table contains the field Fac/SiteID that is used to connect almost every other table in the database to a particular site and facility. Because so many of the other tables in the database contain a Fac/SiteID, the GeneralFacilityInformation table is a central one in the database.

Type of Operations

Most of the names in these tables clearly define what information they hold, but a few require further explanation. The OpCode data field in both tables is used to connect to another table called TypeofFacility/Operations located just below the GeneralFacilityInformation table in Figure 4.3b, whose records consist of an id (code and id are used interchangeably) and the type of operation or activities that characterize the site or facility. Figure 4.3.1a shows a sample of

records that could be contained in the TypeofFacility/Operation table.

OpCode	Facility/Operation
1	Aircraft operations
2	Aircraft maintenance
3	Aircraft support
4	Aircraft operations, maintenance, and support
5	Mechanized Army Training
6	Army Maneuvers
7	Mechanized Army Training, maneuvers, and support

Figure 4.3.1a Sample of TypeofFacility/Operation Table Records

Any descriptions for the type of operations could be added to this table. Referential integrity is enforced between the "GeneralInformation" tables and the TypeofFacility/Operation table to ensure that only valid types of operations are entered in the GeneralInformation tables. Because any one description of a type of operation can apply to a variety of sites and facilities, a one-to-many relationship exists between the TypeofFacility/Operations table to the GeneralInformation tables.

Prior Spills Table

The General Information section of the PA scoresheets asks for the history of prior spills at the site being investigated. A GeneralInformationPriorSpills table located in the lower left hand corner of Figure 4.3b is created to hold this information. This table contains the data field SpillID to act as a primary key, a SiteID to associate it with a particular site, a Source Type whose description is contained in the SourceTypes table, and finally a Prior Spills data field which is a verbal description of the details of the spill. The Source Type table contains the primary key Spill ID, and a second field which is a text description of the type of the source. Figure 4.3.1b shows the records that make up the Source Type table.

Source ID	Source Type
1	Landfill
2	Surface Impoundment
3	Drums
4	Tanks and Non-Drum Containers
5	Contaminated Soil
6	Pile
7	Land Treatment
8	Other

Figure 4.3.1b All Records Contained in the Source Type Table

The definitions of these source types are explained in the PA scoresheets in Appendix A.

Investigation Tables

It is also necessary to record information about those responsible for performing the preliminary assessment investigations. The InvestigatorInformation and Investigations tables both located directly above the GeneralFacilityInformation table in Figure 4.3b are used for this purpose. The InvestigatorInformation table keeps track of the Agency/Organization, Investigator Address, etc. while the Investigations table keeps track of which agency was responsible for the investigation, the name of the main investigator and the date of the investigation. In the GeneralFacilityInformation table, an InvestigationID is included to associate the site and facility with a particular investigation.

Facility Hazardous Wastes

The last information that is recorded in the General Information section of the PA scoresheet is a description of the hazardous wastes that may have been stored, handled, or disposed of at the facility. The FacilityHazSub table (see top left hand corner of Figure 4.3b) consists of a Facility Hazardous Waste ID used as a primary key, Fac/SiteID, HazardousID which is a link to a list of Hazardous wastes listed in the records of the HazardousSubstances table, and a StateID. The StateID refers to the HazardousWasteStates table listing the possible combinations of storing, handling and disposing hazardous wastes. Similar to the TypeofFacility/Operations and SourceType tables, the HazardousWasteStates tables simply restricts the values that are entered in other tables to the set of records contained within these tables. These tables are not actually required in the database, but it is one method in which to

restrict entries in data fields to a particular set of values. For example, the HazardouseStateID field of in the FacilityHazSub table may only contain the values in the HazardousWasteState table as in Figure 4.3.1c.

StateID	State
1	Stored
2	Handled
3	Disposed
4	Stored and Handled
5	Stored and Disposed
6	Handled and Disposed
7	Stored, Handled, and Disposed

Figure 4.3.1c Restricted Values for data field StateID

4.3.2 Source Evaluation Tables

The Source Evaluation section of the PA scoresheets is used to describe the various sources of waste generation. This description includes the type of source, physical characterization and relevant quantities. Because any facility may have multiple sources of waste, there is a one-to-many relationship that exists between the GeneralFacilityInformation and SourceEvaluation tables. The SourceEvaluation table consists of a SiteSourceID used as the primary key, SourceNumber which starts at "1" for the first waste source being described for a facility, Fac/SiteID, and other information that is required in the Source Evaluation section of the PA scoresheet.

The SourceID in the SourceEvaluation table is restricted to the values contained in the SourceTypes table that contains the records shown in Figure 4.3.1b.

4.3.3 Ground Water Pathway Tables

The Ground Water Pathway is used to record information about the possible threat to the environment or human health by way of the ground water. In this section, information about nearby ground water usage, stratigraphy, aquifers, rainfall, drinking water populations, etc. is gathered.

The first table used to record information about the ground water pathway is the Ground-

Water GeneralInformation table. In this table, information about stratigraphy (data field, Stratigraphy), aquifers(data field, Aquifers), and several other parameters relevant to questions that are actually asked in the scoresheet section are stored. The data fields Depth to Aquifer, Blended System?, Exposed Population, and Not Exposed Population are recorded in the GroundWaterGeneralInformation table rather than the GroundWater-Scoresheet table so that only numeric values are recorded in the GroundWater-Scoresheet table. This is done simply to keep the contents of the "score" sheet table consistent with its name. The distribution of municipal and private wells within a 4-mile radius is maintained by the GroundWater-Wells table.

This first data field of this table is a DistanceID, which specifies a distance interval away from the center of the facility where the well is located. Figure 4.3.3a shows the records in the table 4DistanceIntervals that define the possible distance intervals.

DistanceID	Distance
1	0 -1 mile
2	1- 2 mile
3	2-3 mile
4	3-4 mile

Figure 4.3.3a Distance Intervals in 4DistanceIntervals Table

The next data field in the table GroundWater-Wells is the Fac/SiteID which links the record to a particular Facility and Site. This is followed by the data field Population, the population served by the well within the specified distance, then the data field Flow(cfs), the amount of water serving the population, and finally the data field WellType, which is restricted to take on either the value "private" or "municipal".

The next two tables in the ground water section are the GroundWater-PrimaryTargets and GroundWater-SuspectedRelease. Both of these tables are used to complete the Ground Water Pathway Criteria List on the preliminary assessment scoresheet. This section is used as a guide in developing hypotheses concerning the occurrence of a suspected release, and the exposure of, specific targets to a hazardous substance. Each of these two tables simply contains data fields that correspond to the questions requiring "yes", "no", or "unknown" answers. Each table also includes a data field called SummarizeRationale where the investigator may record at length

his/her reasoning for suspecting a hazardous waste release in the GroundWater-SuspectedRelease table together with the reasoning for specifying a certain well as a primary target of concern in the GroundWater-PrimaryTarget table. These tables also include a Fac/SiteID.

Finally, the table GroundWaterScoresheet consists of data fields that are used to calculate a final Ground Water Pathway Score. Referring to the Ground Water Pathway Scoresheet in the Appendix A, the data fields in this table correspond directly to all of the subscores contained in this section.

4.3.4 SURFACE WATER PATHWAY TABLES

The design of the tables for the Surface Water Pathway section of the PA scoresheets is very similar to the Ground Water Pathway section. For this section there exist corresponding tables for General Information, Primary Targets, and Suspected Releases. The main difference is the break down of the Surface Water Scoresheet into two separate tables. One holds information relating to the Human Food Chain Threat and Environmental Threat, while the other is the Likelihood of Release (and Drinking Water Threat) table.

The design principle of storing only data values used to calculate a total surface water score was used in this case so that the SurfaceWater-GeneralInformation table contains data fields that store answers to some of the questions in the Surface Water Pathway Scoresheet. In addition, the SurfaceWater-GeneralInformation table contains a Surface Water Sketch, which is an image showing the runoff route, intakes, fisheries, sensitive environments, etc.

The SurfaceWater-PrimaryTarget and SurfaceWater-SuspectedRelease tables consist of a series of data fields to store the "yes", "no", and "unknown" values to the questions on the Surface Water Pathway Criteria List. Similar to the Ground Water tables, these tables also contain a Summarize Rationale data field for an explanation of suspecting a hazardous waste release through the surface water and an explanation for specifying a certain location or population as a primary target that may have been exposed to the hazardous waste.

The table SurfaceWater-LikelihoodofReleaseSS consists of five data fields corresponding exactly to the five scores recorded on this section of the scoresheet. These data fields are called Likelihood, Primary Target, Secondary Target, Nearest Intake, Resources, and Total. The table SW-Waterbodies-DrinkingWaterThreat is used to hold information about the drinking water

intakes which is a part of this section of the scoresheet. This table consists of the data fields Intake Name, WaterBodyID, Flow(cfs), People Served, and Primary, which takes on the values of either "yes" or "no" indicating whether or not the intake is a primary target. The WaterBodyID is a link to the table WaterbodyTypes, which contains the records shown in Figure 4.3.4a

WaterBodyID	Type
1	River
2	Estuary
3	Lake
4	Stream
5	Canal
6	Pond
7	Bay

Figure 4.3.4a WaterBodyType Table Records

The tables SurfaceWater-HumanFoodChainSS and SurfaceWater-EnvironmentThreatSS are very similar to the SurfaceWater-LikelihoodofReleaseSS that was just described. They contain data fields matching exactly with the scores required in these sections of the scoresheet. While the "Drinking Water Threat" section includes a list of Water Intakes that was recorded in the SW-Waterbodies-DrinkingWaterThreat, both the Human Food Chain Threat and Environment Threat scoresheets have subsections listing information about fisheries and sensitive environments, including their names, the amount of flow, and the water body type. All of this information is stored in the table SW-WaterBodies-HumanFd/EnvThreat that includes a data field called Type to specify whether it is referring to a "fishery" or "sensitive environment", which are the only allowed values for data field called Type.

4.4 An Introduction to the Structured Query Language

Once the data fields, tables, and relationships between tables have been created, it is then possible to perform certain operations or queries on the database. For example, facilities that contain a particular hazardous waste may be of particular interest. How does one retrieve just the desired information from all of the data stored in the database tables? The structured query language (SQL) is used to accomplish such a task. SQL is based on the relational model of database management proposed in 1970 by Dr. E.F. Codd (Patel and Moss, 1996). Although

SQL involves the creation and modification of tables, this section will outline the basics of performing queries in SQL. The specific SQL statements that are a part of the preliminary assessment database will be presented in the next section.

Figure 4.4a, which is simply a copy of Figure 4.2.5a, is shown again here to be used as an example for making some sample SQL statements.

Fac/SiteID	SiteID	Facility Name	Street Address	City	State	Zip
1	1	LF-1	6787 Worcester St.	Barnstable	MA	02541
2	1	SD-5	8984 Newbury St.	Barnstable	MA	02541
3	1	FS-12	4444 Hayway Rd	Falmouth	MA	02541
4	2	LF-1	6775 Butte Wy.	Oroville	CA	95966
5	2	CS-10	7456 Chico St.	Marysville	CA	95968
6	3	CS-4	6666 Invalid Wy.	Mistake	MA	02541

Figure 4.4a Table Used for Illustration of SQL

4.4.1 SINGLE TABLE QUERY

Suppose one wishes to see a list of facilities located in the state of Massachusetts. To do this we would make use of SQL's primary command SELECT. To illustrate the use of the SELECT statement, consider the following syntax of the SELECT statement:

```
SELECT column_names
FROM table_names
WHERE predicates
```

The result of the following SQL statement:

```
SELECT Facility Name
FROM FacilityInformation
WHERE State = "MA"
```

would return the following result:

Facility Name
LF -1
SD-5
FS-12
CS-4

Figure 4.4.1a Simple SQL Result

Suppose we wanted to see the list of facilities in alphabetical order. The ORDER BY directive is used to accomplish this. The new SQL statement becomes:

```
SELECT Facility Name
FROM FacilityInformation
WHERE State = "MA"
ORDER BY Facility Name
```

As expected, the result of this statement is:

Facility Name
CS-4
FS-12
LF-1
SD-5

Figure 4.4.1b Sorting in SQL

The default sorting order is ascending. To sort the records in descending order append the keyword DESC at the end of the ORDER BY statement.

To return all data fields within in a table, simply use a "*" as in:

```
SELECT *
FROM FacilityInformation
WHERE State = "MA"
```

The result of this statement would be:

Fac/SiteID	SiteID	Facility Name	Street Address	City	State	Zip
1	1	LF-1	6787 Worcester St.	Barnstable	MA	02541
2	1	SD-5	8984 Newbury St.	Barnstable	MA	02541
3	1	FS-12	4444 Hayway Rd	Falmouth	MA	02541
6	3	CS-4	6666 Invalid Wy.	Mistake	MA	02541

Figure 4.4.1c Selecting All Data Fields

It is also possible to select multiple data fields separated by commas in the SELECT statement:

```
SELECT Facility Name, City
FROM FacilityInformation
WHERE State = "MA"
```

returning the following result:

Facility Name	City
LF-1	Barnstable
SD-5	Barnstable
FS-12	Falmouth
CS-4	Mistake

Figure 4.4.1d Selecting Multiple Data Fields

4.4.2 Multiple Table Query

A slightly more complicated query, which would involve two related tables would be to list the names of facilities belonging to the site MMR. To construct this query it is necessary to know how the tables are linked. Figure 4.4.2a shows the two tables that contain the necessary information to perform the query.

SiteInformation

SiteID	Site Name	Site Address
1	MMR	3450 Herbert Rd. Bourne, MA 02542
2	Oroville Army Airfield	4500 Santiago St. Oroville, CA 95966

1

∞

FacilityInformation

Fac/SiteID	SiteID	Facility Name	Street Address	City	State	Zip
1	1	LF-1	6787 Worcester St.	Barnstable	MA	02541
2	1	SD-5	8984 Newbury St.	Barnstable	MA	02541
3	1	FS-12	4444 Hayway Rd	Falmouth	MA	02541
4	2	LF-1	6775 Butte Wy.	Oroville	CA	95966
5	2	CS-10	7456 Chico St.	Marysville	CA	95968
6	3	CS-4	6666 Invalid Wy.	Mistake	MA	02541

Figure 4.4.2a Tables for Multiple Table SQL Illustration

Data fields from two different tables are required to perform the query. *Dot notation* is used to reference a data field within a table as in:

"table_name"."data_field_name"

The connection between the two tables is accomplished by using the directives INNER JOIN and ON. The following SQL statement would return the desired result:

```
SELECT FacilityInformation.[Facility Name]
FROM SiteInformation INNER JOIN FacilityInformation ON
        SiteInformation.SiteID = FacilityInformation.SiteID
WHERE SiteInformation.[Site Name] = "MMR"
```

As can be inferred from the example, INNER JOIN is used after the FROM keyword and uses the following syntax:

FROM *table1_name* **INNER JOIN** *table2_name* **ON** *table1_name.datafield1* = *table2_name.datafield2*

In order for this to be a valid SQL statement, *datafield1* and *datafield2* do not have to have exactly the same name (i.e. SiteID), but they do have hold values of the same data type. Common examples of data types are numbers, dates, or text. The SQL statement above would return the result:

Facility Name
LF -1
SD-5
FS-12

Figure 4.4.2b Multiple Table SQL Result

4.4.3 VARIABLES, NEW TABLES, AND AGGREGATE FUNCTIONS

It is often very useful to perform basic mathematical operations on different data fields or columns. These operations are possible in SQL. The result of these operations may be stored in variables using the keyword AS. Figure 4.4.3a will be used as an example to show some basic calculations.

GroundWater-Scoresheet

Fac/SiteID	LiklihoodScore	PrimaryScore	SecondaryScore	TScore
1	300	250	500	60
2	500	250	400	150
3	500	250	600	200
4	400	250	200	150

Figure 4.4.3a GroundWater-Scoresheet Sample Records

The SQL statement:

```
SELECT Fac/SiteID, (LikelihoodScore + PrimaryScore + SecondaryScore +
Tscore) AS Total Score
FROM GroundWaterScoresheet
```

returns the result:

TotalTable

Fac/SiteID	Total Score
1	1050
2	1300
3	1550
4	1000

Figure 4.4.3b Result of Basic Mathematical Operations

The use of the keyword INTO allows the saving of the query into a new table. If the statement, **INTO** TotalTable, were added after the SELECT statement, the query result would then get saved into a new table called TotalTable.

In addition to being able to perform operations on data fields in one record, it is also possible to perform "aggregate functions" on the same data field in all records. They are called aggregate functions because they summarize the results of a query, rather than listing all of the rows. These operations include AVG, MAX, MIN, SUM, and COUNT. Suppose the previous query had in fact been saved INTO the table TotalTable. The SQL statement:

```
SELECT AVG(TotalScore) AS AverageScore, MAX(TotalScore) AS
       MaximumScore, SUM(TotalScore) AS SumOfScores,
       COUNT(TotalScore) AS NumberOfRecords
FROM TotalTable
```

would return the following:

AverageScore	MaximumScore	SumOfScores	NumberOfRecords
1225	1550	4900	4

Figure 4.4.3a Result of Aggregate Functions

4.4.4 THE TOP KEYWORD

Say that next we wanted to list the facilities with the top 2 TScores from the GroundWater-Scoresheet table listed here:

GroundWater-Scoresheet

Fac/SiteID	LiklihoodScore	PrimaryScore	SecondaryScore	TScore
1	300	250	500	60
2	500	250	400	140
3	500	250	600	200
4	400	250	200	150

Figure 4.4.4a GroundWater-Scoresheet Sample Records

The use of the keyword TOP *x* placed directly after the keyword SELECT returns the first two records of a query result. The following SQL statement would return the highest two TScores from the GroundWater-Scoresheet table as seen in Figure 4.4.4b:

```
SELECT TOP 2 Fac/SiteID, TScore
FROM GroundWater-Scoresheet
ORDER BY Tscore DESC
```

Notice that the ORDER BY line is included to first sort the records from highest to lowest values of TScore.

Fac/SiteID	TScore
3	200
4	150

Figure 4.4.4b Use of TOP Keyword

4.4.5 THE KEYWORDS DISTINCT AND DISTINCTROW

The keywords DISTINCT and DISTINCTROW are used to avoid returning records that contain duplicate data fields or rows. For example, suppose we are interested in sites that contain a facility with a TScore greater than 100. First notice in Figure 4.4.5a that the facilities with Fac/SiteID's of 2, 3, 4 all have TScores above 100.

SiteInformation

SiteID	Site Name	Site Address
1	MMR	3450 Herbert Rd. Bourne, MA 02542
2	Oroville Army Airfield	4500 Santiago St. Oroville, CA 95966

1

∞ ***FacilityInformation***

Fac/SiteID	SiteID	Facility Name	Street Address	City	State	Zip
1	1	LF-1	6787 Worcester St.	Barnstable	MA	02541
2	1	SD-5	8984 Newbury St.	Barnstable	MA	02541
3	1	FS-12	4444 Hayway Rd	Falmouth	MA	02541
4	2	LF-1	6775 Butte Wy.	Oroville	CA	95966
5	2	CS-10	7456 Chico St.	Marysville	CA	95968
6	3	CS-4	6666 Invalid Wy.	Mistake	MA	02541

GroundWater-Scoresheet

Fac/SiteID	LiklihoodScore	PrimaryScore	SecondaryScore	TScore
1	300	250	500	60
2	500	250	400	150
3	500	250	600	200
4	400	250	200	150

Figure 4.4.5a Tables Used for DISTINCT Illustration

Two of these facilities belong to the site MMR. Since we are only interested in seeing which sites contain facilities with TScores above 100, we do not need to list the MMR twice. What we need to do is explicitly ask for records where the site is not duplicated. The following SQL statement will achieve the desired result:

```
SELECT DISTINCT SiteInformaiton.[Site Name]
FROM (SiteInformation INNER JOIN FacilityInformation ON
        SiteInformation.SiteID = FacilityInformation.SiteID) INNER JOIN
        [GroundWater-Scoresheet] ON FacilityInformation.[Fac/SiteID] =
        [GroundWater-Scoresheet].[Fac/SiteID]
WHERE GroundWater-Scoresheet.TScore > 100
```

The query result is:

Site Name
MMR
Oroville Army Airfield

Figure 4.4.5b Using DISTINCT in SQL

The keyword DISTINCTROW is used when all data fields in a record are required to be unique.

Finally, look at the use of two consecutive INNER JOINS in the SQL statement just described. The first inner join essentially creates a temporary table that combines the tables SiteInformation and FacilityInformation through the common data field SiteID. This temporary table is then combined to the GroundWater-Scoresheet table through the common field Fac/SiteID.

With an understanding of the basic syntax for making queries using SQL, the specific SQL statements of the preliminary assessment scoresheet database can now be presented.

4.5 Performing Queries in the PA Database

The ultimate goal of the preliminary assessment database is to be able to use existing scoresheets stored in the database to aid in the completion of new studies. Answers to specific questions, scores, and explanations of these responses on completed scoresheets may provide useful information and checks for reliability when answering questions related to the current facility under review. Under the architecture presented here, there are six criteria that will be used as basis for comparison between studies. Figure 4.5a shows the six comparison factors that the investigator should be able to use to compare facilities:

Comparison Criteria
Geographic Location
Type of Facility
Hazardous Wastes
Waste Source Types
Multiple Parameters
By Pathway

Figure 4.5a Comparison Criteria

When performing these comparisons, the user selects the values that will be used for comparisons or uses the values from a selected facility. The result of the query is a ranked listing of facilities by similarity to the values chosen for comparison. The user would then select the facility in which he or she wishes to review a certain section of the completed scoresheet.

An understanding of relational database design and the use of SQL now makes it possible to explain the SQL statements that will allow these comparisons. The SQL statements described in this section will contain certain labels of the form, *Forms![Form Name]![Control]*. The exact meaning of this syntax will be explained in the next section, but for now, these labels will just indicate variables that need to be replaced by numeric or textual values in order for the SQL statements to be valid. For example, in the generic SQL statement:

SELECT *table.[data field]*
FROM *[table]*
WHERE *table.[data field]* = *Forms![SampleForm]![SampleControl]*

the label *Forms![Form Name]![Control]* would have to be replaced by an appropriate numeric or textual value before the SQL statement would actually return valid results.

4.5.1 Location Comparison SQL

The "Location Comparison" query shown in Figure 4.5.1a returns the top five facilities closest to the coordinates specified by the variables *[Forms]![Location]![Input Longitude]* and *[Forms]![Location]![Input Latitude]*. The data fields Fac/SiteID, Facility Name, Longitude, and Latitude are selected from the table GeneralFacilityInformation. The data field Site Name is taken from the GeneralSiteInformation table.

The bulk of the query is devoted to converting the geographical coordinates from degrees, minutes, and seconds into just degrees. The coordinates are expressed in the form DDMMSS.S

where DD equals degrees, MM is minutes, and SS.S is seconds to the tenth of a second. The coordinates are converted to degrees to perform the comparisons. Most of the intermediate variables used in this query are used to determine the degree equivalence of the coordinates input by the user and contained within the facility records. The distance between the coordinates in *[Forms]![Location]![Input Longitude]* and *[Forms]![Location]![Input Latitude]* where the latitude and longitude coordinates of each facility is approximated by using the Pythagorean Theorem. Figure 4.5.1b summarizes the "Location Comparison" query including a short description of what the query does, the tables and data tables, the data fields linking the tables, variable naming conventions, and explanation of the mathematical functions used in the query.

```
SELECT DISTINCTROW TOP 5 GeneralFacilityInformation.[Fac/SiteID], GeneralFacilityInformation.[Facility
Name], GeneralSiteInformation.[Site Name], GeneralFacilityInformation.Longitude,
GeneralFacilityInformation.Latitude,

Int([GeneralFacilityInformation].[Longitude]/100000)                          AS  DegreesLong,
[GeneralFacilityInformation].[Longitude]-[DegreesLong]*100000                 AS  IntermediateLong,
Int([IntermediateLong]/1000)                                                  AS  MinutesLong,
([IntermediateLong]-[MinutesLong]*1000)                                       AS  SecondsLong,
[DegreesLong]+[MinutesLong]/60+[SecondsLong]/3600                             AS  LongitudeDecimal,
Int([Forms]![Location]![Input Longitude]/100000)                              AS  DegreesLongInput,
[Forms]![Location]![Input Longitude]-[DegreesLongInput]*100000                AS  IntermediateLongInput,
Int([IntermediateLongInput]/1000)                                             AS  MinutesLongInput,
([IntermediateLongInput]-[MinutesLongInput]*1000)                             AS  SecondsLongInput,
[DegreesLongInput]+[MinutesLongInput]/60+[SecondsLongInput]/3600              AS  LongitudeDecimalInput,
Int([GeneralFacilityInformation].[Latitude]/100000)                           AS   DegreesLat,
[GeneralFacilityInformation].[Latitude]-[DegreesLat]*100000                   AS  IntermediateLat,
Int([IntermediateLat]/1000)                                                   AS  MinutesLat,
 ([IntermediateLat]-[MinutesLat]*1000)                                        AS  SecondsLat,
[DegreesLat]+[MinutesLat]/60+[SecondsLat]/3600                                AS  LatitudeDecimal,
Int([Forms]![Location]![Input Latitude]/100000)                               AS  DegreesLatInput,
[Forms]![Location]![Input Latitude]-[DegreesLat]*100000                       AS  IntermediateLatInput,
Int([IntermediateLatInput]/1000)                                              AS   MinutesLatInput,
([IntermediateLatInput]-[MinutesLatInput]*1000)                               AS  SecondsLatInput,
[DegreesLatInput]+[MinutesLatInput]/60+[SecondsLatInput]/3600                 AS  LatitudeDecimalInput,
[LongitudeDecimalInput]-[LongitudeDecimal]                                    AS  DeltaLong,
[LatitudeDecimalInput]-[LatitudeDecimal]                                      AS  DeltaLat,
Sqr([DeltaLong]*[DeltaLong]+[DeltaLat]*[DeltaLat])                            AS  Miles

FROM GeneralSiteInformation INNER JOIN GeneralFacilityInformation ON GeneralSiteInformation.SiteID =
GeneralFacilityInformation.SiteID
ORDER BY Miles;
```

Figure 4.5.1a "Location Comparison" SQL Statement

Brief Description: "Location Comparison" returns the top five facilities closest to coordinates entered by user.

Tables: GeneralFacilityInformation, GeneralSiteInformation
Data Fields Selected: Fac/SiteID, Facility Name, Longitude, Latitude, SiteName
Link Between Tables: SiteID
User Specified Variables: [Forms]![Location]![Input Longitude], [Forms]![Location]![Input Longitude]

Variable Naming Conventions:
1) The root "Long" refers to longitude, while the suffix "Lat" refers to latitude
2) The suffix "_Input" refers to the coordinates by the user used for comparison
3) The root "Degrees" refers to only the degrees portion of the coordinate, "Minutes" refers only to the minutes portion of the coordinate, and "Seconds" refers t only to the seconds portion of the coordinate
4) The prefix "Delta_" refers to the difference of whatever is contained in the suffix.
5) The variable "Delta" is the final difference between the input coordinates and the coordinates of a facility.
6) The variables [Forms]![Location]![Input Longitude] and [Forms]![Location]![Input Latitude] refers to the coordinates input from the form "Location", but more generally refers to coordinates specified by the user to be used for the comparisons.

Function Definitions:

1) The function "Int(argument)" returns only the integer portion of the argument.

Figure 4.5.1b "Location Comparison" Query Summary

4.5.2 TYPE OF FACILITY COMPARISON SQL

The "Location Query" comparison is relatively more complex than the rest of the comparisons. The Type of Facility query simply lists all those facilities that have been classified as performing the type of operation specified by the variable *[Forms]![Type of Facility]![Operation Choice]* in the SQL statement. Suppose the facility under current investigation has been classified as an "Aircraft maintenance" facility. The "Type Query" will return a list of all those facilities whose operations have been given the same description. The list of valid types of operations is contained in the table TypeofFacility/Operations as shown in Figure 4.5.2a

OpCode	Facility/Operation
1	Aircraft operations
2	Aircraft maintenance
3	Aircraft support
4	Aircraft operations, maintenance, and support
5	Mechanized Army Training
6	Army Maneuvers
7	Mechanized Army Training, maneuvers, and support

Figure 4.5.2a. Sample of TypeofFacility/Operation Table Records

Figure 4.5.2b show the summary description of the "Type Query" and the "Type Query" SQL statement.

```
SELECT DISTINCTROW GeneralFacilityInformation.[Fac/SiteID], GeneralFacilityInformation.[Facility
Name], GeneralSiteInformation.[Site Name], GeneralFacilityInformation.OpCode

FROM GeneralSiteInformation INNER JOIN GeneralFacilityInformation ON
GeneralSiteInformation.SiteID = GeneralFacilityInformation.SiteID
WHERE GeneralFacilityInformation.OpCode=[Forms]![Type of Facility]![Operation Choice];
```

Brief Description: "Type Query" returns a list of all facilities containing the specified type of facility. The list of possible types of facilities is contained in the table TypeofFacility/Operations.

Tables: GeneralFacilityInformation, GeneralSiteInformation
Data Fields Selected: Fac/SiteID, Facility Name, SiteName, GeneralFacilityInformation.OpCode
Link Between Tables: SiteID
User Specified Variables: [Forms]![Type of Facility]![Operation Choice]

Figure 4.5.2b "Type Query" SQL Statement:

4.5.3 HAZARDOUS WASTE COMPARISON SQL

The Hazardous Waste query lists all those facilities containing the hazardous waste specified by the variable *[Forms]![Type of Facility]![Hazardous Choice]*. The list of valid hazardous substances is contained in the table HazardousSubstances. Figure 4.5.3a shows a sample of the hazardous wastes in the table HazardousSubstances.

HazardousID	Name
1	Trichloroethylene
2	Tetrachloroethylene
3	1,2 dicholorethylene
4	Vinyl Chloride
5	Carbon Tetrachloride
6	Benzene
7	Ethylene Dibromide
8	Toluene
9	Ethylbenzene
10	Xylene
11	Other

Figure 4.5.3a Sample of HazardousSubstances Table Records

This list of hazardous substances was taken directly from the 1992 Priority List of Hazardous Substances (http://www.medaccess.com/chemicals/92list.htm, Agency for Toxic Substance and Disease Control) The table GeneralSiteInfoFacilityHazSub lists the hazardous wastes contained in each facility.

Figure 4.5.3b shows the summary description of the "Hazardous Ranking Query" and the "Hazardous Ranking Query" SQL statement.

```
SELECT DISTINCTROW GeneralFacilityInformation.[Fac/SiteID], GeneralFacilityInformation.[Facility
Name], GeneralSiteInformation.[Site Name], HazardousSubstances.HazardousID,
HazardousSubstances.Name
FROM HazardousSubstances INNER JOIN (GeneralSiteInformation INNER JOIN
(GeneralFacilityInformation INNER JOIN GeneralSiteInfoFacilityHazSub ON
GeneralFacilityInformation.[Fac/SiteID] = GeneralSiteInfoFacilityHazSub.[Fac/SiteID]) ON
GeneralSiteInformation.SiteID = GeneralFacilityInformation.SiteID) ON
HazardousSubstances.HazardousID = GeneralSiteInfoFacilityHazSub.HazardousID
WHERE HazardousSubstances.HazardousID=[Forms]![Hazardous Wastes]![Hazardous Choice];
```

Brief Description: "Hazardous Query" returns a list of all facilities containing the specified hazardous substance. The list of possible hazardous substances is contained in the table HazardousSubstances. The list of hazardous wastes contained at each site is listed in the table GeneralSiteInfoFacilityHazSub.

Tables: GeneralFacilityInformation, GeneralSiteInformation, GeneralSiteInfoFacilityHazSub, HazardousSubstances
Data Fields Selected: Fac/SiteID, Facility Name, SiteName, HazardousID, HazardousSubstances.Name
Link Between Tables: SiteID (between GeneralSiteInformation and GeneralFacilityInformation), Fac/SiteID (between GeneralSiteInfoFacilityHazSub and GeneralFacilityInformation), HazardousID (between GeneralSiteInfoFacilityHazSub and HazardousSubstances)
User Specified Variables: [Forms]![Hazardous Wastes]![Hazardous Choice]

Figure 4.5.3b "Hazardous Ranking Query" SQL Statement

4.5.4 WASTE SOURCE TYPE COMPARISON SQL

The Waste Source Type query lists all those facilities containing the type of waste source specified by the *[Forms]![Waste Source]![Source Choice]* variable. The list of valid hazardous substances sources is contained in the table SourceTypes (see Figure 4.5.4a).

Source ID	Source Type
1	Landfill
2	Surface Impoundment
3	Drums
4	Tanks and Non-Drum Containers
5	Contaminated Soil
6	Pile
7	Land Treatment
8	Other

Figure 4.5.4a All Records Contained in the SourceTypes Table

Figures 4.5.4b and 4.5.4c show the summary description of the "Waste Source Comparison" query and the "Waste Source Comparison " SQL statement.

SELECT DISTINCTROW GeneralFacilityInformation.[Fac/SiteID], GeneralFacilityInformation.[Facility Name], GeneralSiteInformation.[Site Name], SourceTypes.[Source Type], SourceTypes.[Source ID] FROM SourceTypes INNER JOIN (GeneralSiteInformation INNER JOIN (GeneralFacilityInformation INNER JOIN SourceEvaluation ON GeneralFacilityInformation.[Fac/SiteID] = SourceEvaluation.[Fac/SiteID]) ON GeneralSiteInformation.SiteID = GeneralFacilityInformation.SiteID) ON SourceTypes.[Source ID] = SourceEvaluation.SourceID
WHERE SourceTypes.[Source ID]=[Forms]![Waste Source]![Source Choice];

Brief Description: "Waste Source Comparison" returns a list of all facilities containing the specified source of hazardous waste. The list of possible hazardous substances is contained in the table SourceTypes. The list of hazardous waste sources at each facility is contained in the table SourceEvaluation.

Tables: GeneralFacilityInformation, GeneralSiteInformation, SourceEvaluation, SourceTypes

Data Fields Selected: Fac/SiteID, Facility Name, SiteName, SourceTypes.[Source Type], SourceTypes.[Source ID]

Link Between Tables: SiteID (between GeneralSiteInformation and GeneralFacilityInformation), Fac/SiteID (between SourceEvaluation and GeneralFacilityInformation), SourceID (between Source Evaluation and Source Types)

User Specified Variables: [Forms]![Waste Source]![Source Choice]

Figure 4.5.4b "Waste Source Comparison" Query Summary

4.5.5 Methodology For Multiple Parameter Comparisons

It would also be useful to be able to compare facilities by choosing multiple parameters. Although the SQL statement for this type of comparison has not been developed here, it would be a rather simple extension of the first four queries. When more than one criterion is used to perform a comparison, the appropriate tables and data fields would be included in the SELECT statement, and the WHERE clause would be extended to only return records containing the desired values. In the case where a location comparison is performed, the result could be ranked by proximity to the input coordinates only for those sites that also match the other criteria.

4.5.6 Methodology For Pathway Comparisons

The criteria list sections of the PA scoresheet consist of a set of questions that require "yes", "no", and "unknown" responses. These responses could be compared to find similarities between facility ground water, surface water, soil exposure, or air characteristics. To implement the pathway comparison, the user would first select a facility and pathway type. The query would tally up the number of similar responses to the questions in the criteria list section of the appropriate pathway. The result would be a ranked list of facilities by the number of questions answered as the same. The user would then have the option to select any facility on this list for further review.

5.0 Implementation of the Preliminary Assessment Scoresheet Database for Cross-Site Comparisons

Now that the framework has been laid for developing the preliminary assessment database, it will be worthwhile to look at an example of how this database is used in a commonly available relational database. Microsoft Access has been selected to show how the preliminary assessment architecture may be implemented, because of its relative ease of utilizing graphical user interfaces or GUI's. Section 5.1 explains what features are involved in developing graphical user interfaces in Microsoft Access, while section 5.2 describes how these features were utilized to complete the final functional database.

5.1 Using Microsoft Access 95

5.1.1 Graphical User Interface and Forms

A graphical user interface or GUI is a set of screen displays through which the user makes use of a computer system. It consists of menus, buttons, selection boxes, text entry input fields, pictures, decorative lines and boxes, and other items. The GUI allows the entering and viewing of data as well as the execution of SQL statements and other high level database operations. The advantage of graphical user interface is that the details of the database structure and the execution of database operations are transparent to the user. Thus, he or she only needs to understand a relatively small set of actions and commands to navigate through the GUI.

In Microsoft Access, the graphical user interface consists of a set of *forms* which may or may not be *bound* to tables and queries. When a form is "bound", it shows the records, usually one at a time, of the table or query on which it is based. These records may be opened just for viewing or they may be edited depending on the properties of the form. Forms that are not bound to any tables or queries may just be used as switchboards which allow access to or categorize other forms.

5.1.2 CONTROLS

Forms consist of *controls* which are text labels, text boxes, list boxes, combo boxes, buttons, and pictures. Figure 5.1.2a shows a screen shot of a sample form in Microsoft Access.

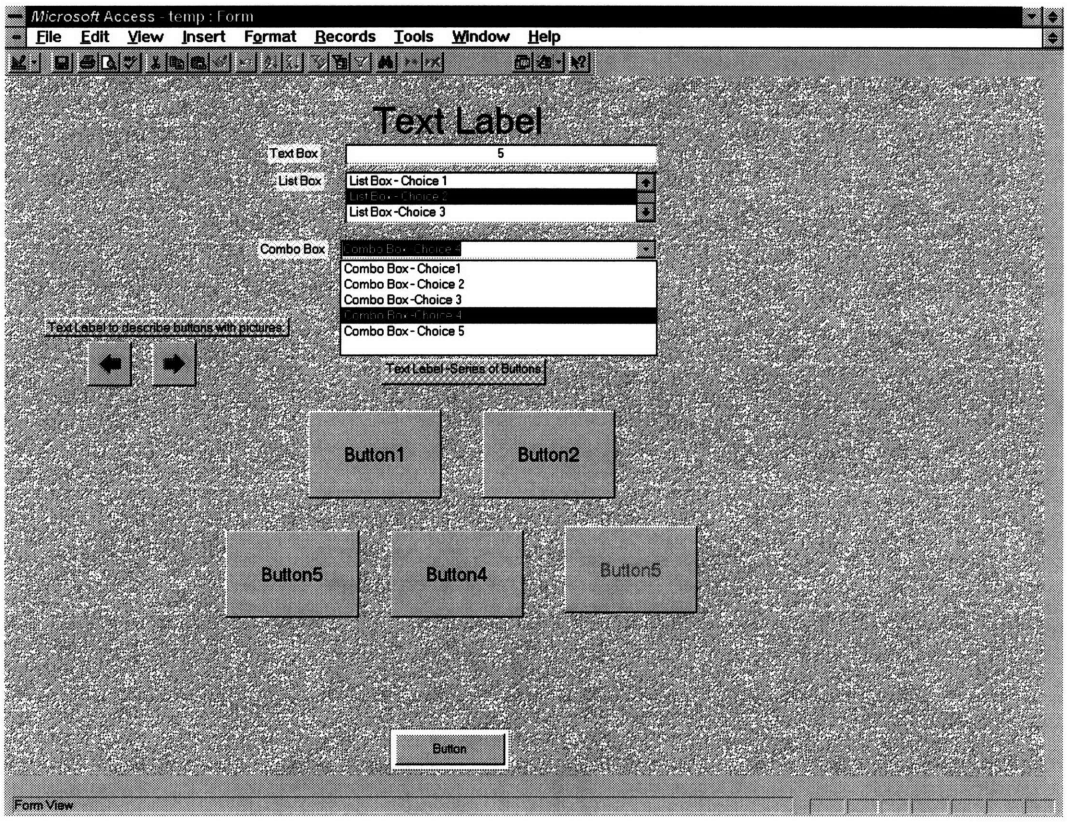


Figure 5.1.2a Controls in a Sample Form

Each control has certain properties that can be adjusted by the GUI developer. A control can be bound to a particular data field of a table or aquery. For example, if a form is bound to a table consisting of data fields Fac/SiteID, and Facility Name, and a text box control is bound to either one of these data fields, then that text box will display the contents of the data field for the current record on display. A text box control that is not bound to a data field may also be bound to a calculation. It may show the result of performing some mathematical operation on some other data fields. Finally, a control text box may not be bound to anything at all. It may just exist for text to be typed in by the user. List box and combo box controls contain lists that are derived from a query or table or have been entered in by the GUI developer. There are other properties associated with controls that affect their appearance such as text size, background color, and

visibility. Lastly, actions may be triggered when *events* occur in relation to individual controls or the form itself.

5.1.3 MACROS AND EVENTS

An event can range from a mouse click, to pressing a key on the keyboard, to moving the mouse over a control. The result of these events can be tied into thexecution of a *macro*. A macro is a sequence of high-level database operations. Macros can do things like open or close a form, find a record which satisfies some criteria, move or resize a window, or synchronize data between two open forms. Each macro is a set of sequential *actions* with a set of parameters. For example the action statement OpenForm requires specification of the name of the Form in order to be executed. The action statement SetValue requires a name of a control text box and a value to be displayed by the control. Figure 5.1.3a shows an example of a macro.

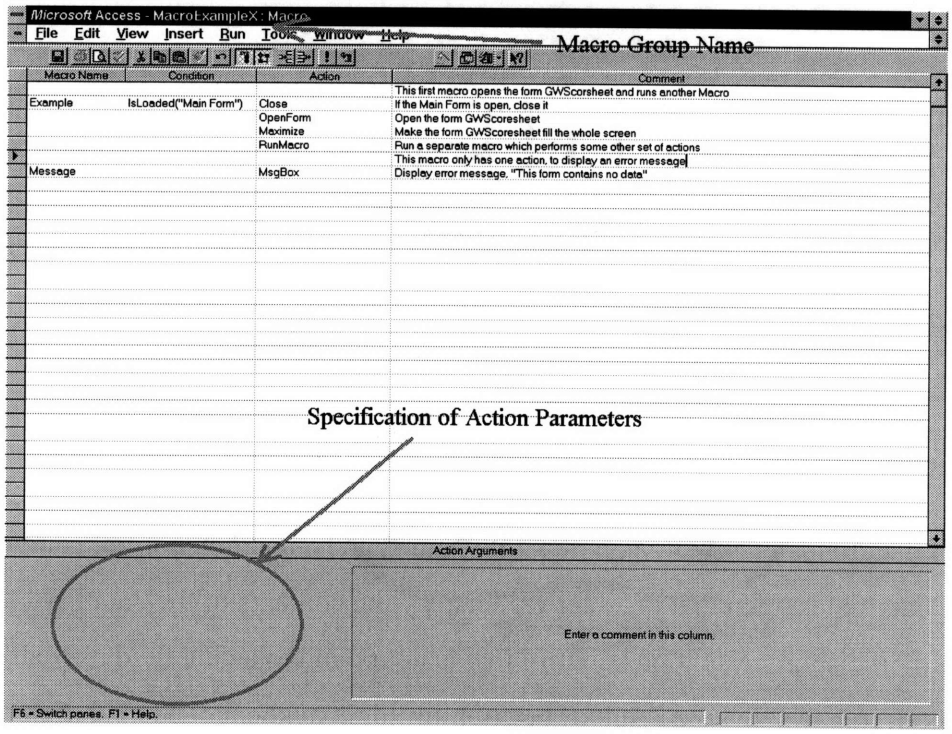


Figure 5.1.3a A Sample Macro

The first two columns in Figure 5.1.3a show a Macro Name column and Condition column. In this case, the macros Example and Message belong to the *macro group* MacroExampleX which

is identified directly after the name "Microsoft Access" on the top of the screen. Microsoft access allows the grouping of related macros into a single macro group. When an event in a form triggers the execution of a macro, the specific macro to be executed is specified by using the *dot notation.* For example, the Example macro would be referenced as MacroExampleX.Example.

The condition column can be used to execute actions only when certain criteria are met. In Figure 5.1.3a, the Condition of the Close action is *IsLoaded("Main Form")* This condition tests to see if the form "Main Form" is open. If it is, then the form is subsequently closed.

The Action column shows a sample of valid macro actions. The "Close" action closes a form while the "OpenForm" action opens a form. "RunMacro" executes a separate macro and the "MsgBox" action displays a message box on the screen.

A Comments column is used to provide further details in the execution of the action. All of these actions require the specification of some parameters, which will appear in the lower left hand corner of the screen in Figure 5.1.3a.

The Event's property of a control is set to associate a macro with an event. A common operation associated with clicking on a button is to set the OnClick event to execute a macro which opens up some form. Figure 5.1.3b shows a screenshot of setting a button's OnClick property to execute a specified macro.

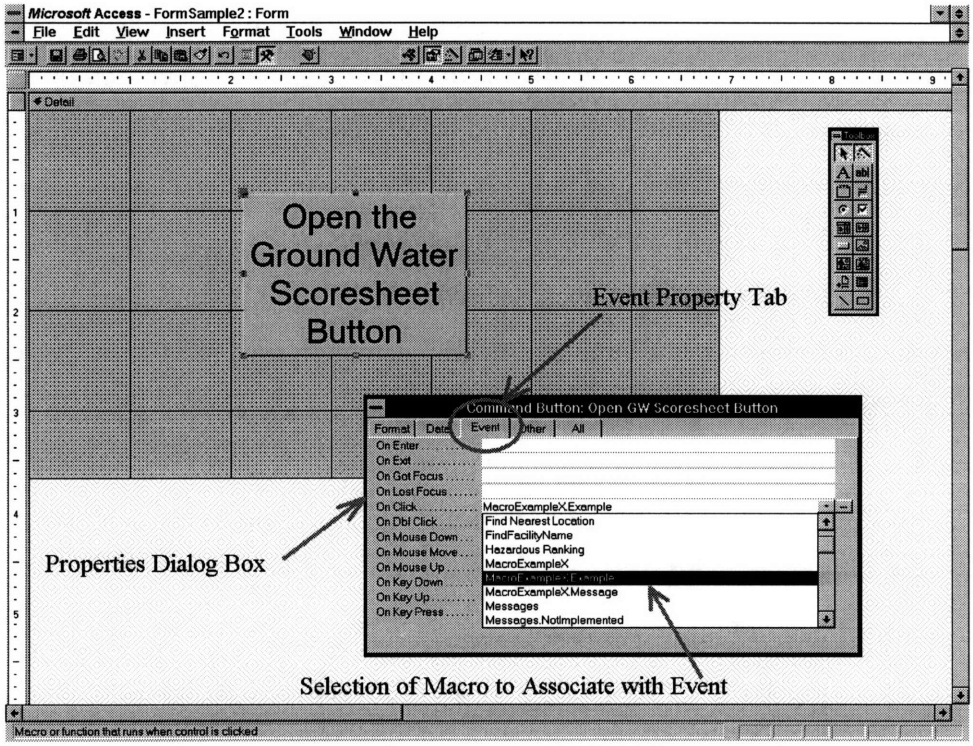


Figure 5.1.3b Linking an Event with a Macro

The gray squares surrounding the Open the Ground Water Scoresheet Button indicate that the Properties Dialog Box refers to this control. The Event Property Tab is selected and the macro that was used in the previous example is selected from the combo box. These are the few simple steps that are necessary to associate an event with a macro.

5.1.4 SUBFORMS

One of the controls that can be included in a form is a *subform*. The term subform just refers to a regular form that is part of another form. Forms and subforms are often used to show one-to-many relationships. For example, a facility contains a number of hazardous wastes. A form that is bound to the GeneralFacilityInformation table may contain a subform that is bound to the table GeneralSiteInfoFacilityHazSub that contains the hazardous wastes contained by the facilities. Figure 5.1.4a shows an example of the Hazardous Substance subform contained within the GeneralFacilityInformation form

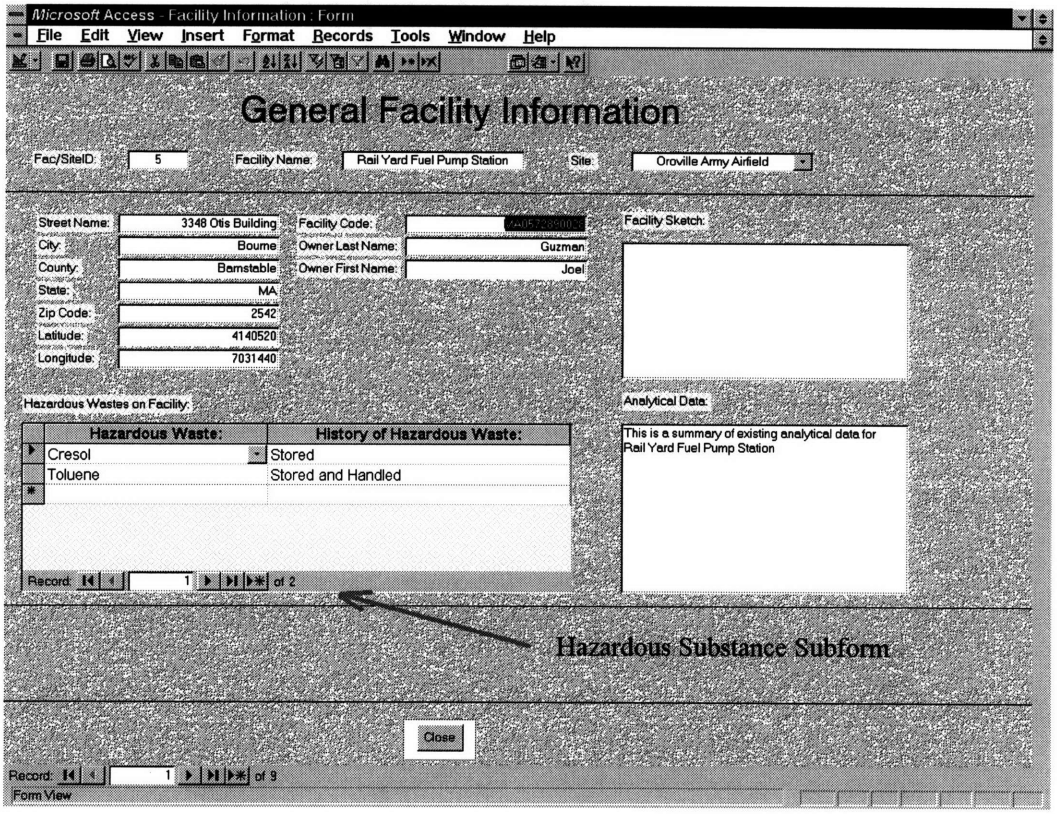


Figure 5.1.4a Example of Subform

The subform is commonly displayed in *datasheet* view where multiple records are displayed all at once rather than only one record at a time. The choice to display the records, either one at a time or in datasheet view, depends on the properties of the form.

These basic features of using forms, events, controls, and macros in Microsoft Access make it possible to develop a graphical user interface for the interactive database of the preliminary assessment scoresheets database.

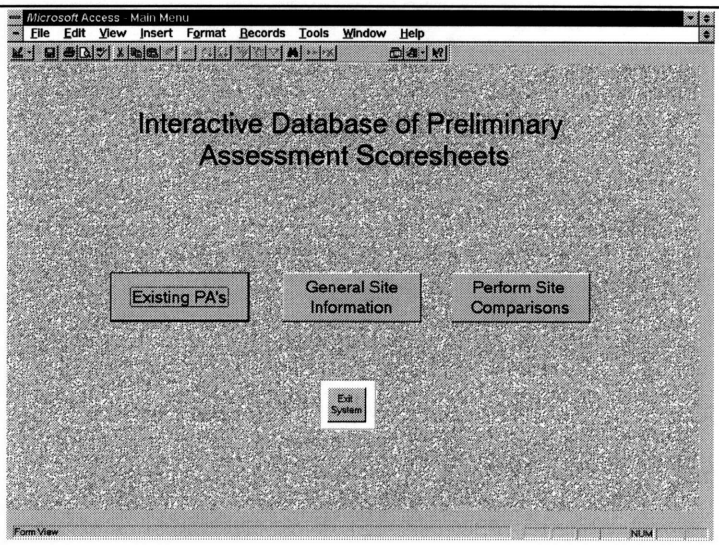
5.2 Implementing the Interactive PA Scoresheet Database

In addition to creating a graphical user interface through forms, Microsoft Access 95 contains modules for developing the underlying database structure and the SQL statements. This section only describes the graphical user interface used to review and edit the PA scoresheet and

perform site comparisons. The user interface may also be used for entering data, although the primary purpose of it is to perform the site comparisons. The interface that is described in this section will make use of the tables and SQL statements that have already been input into the Microsoft Access database

5.2.1 SWITCHBOARDS

A *switchboard* refers to a form whose main purpose is to provide links to other forms by clicking on buttons. Figure 5.2.1a shows the Main Switchboard that is opened when the Microsoft Access PA database file called EIDSS.mdb is opened. It's On Open event is triggered to activate the Maximize macro. This macro simply maximizes the currently active form. The caption property of this form has been set to "Main Menu", so this is what appears on the top of the screen after the *Microsoft Access* title. Each of the button's On Click event properties has been set up to trigger a macro which opens the appropriate form.



Form Name: Main Switchboard

Description: This is the main menu of the PA database

Subforms: none

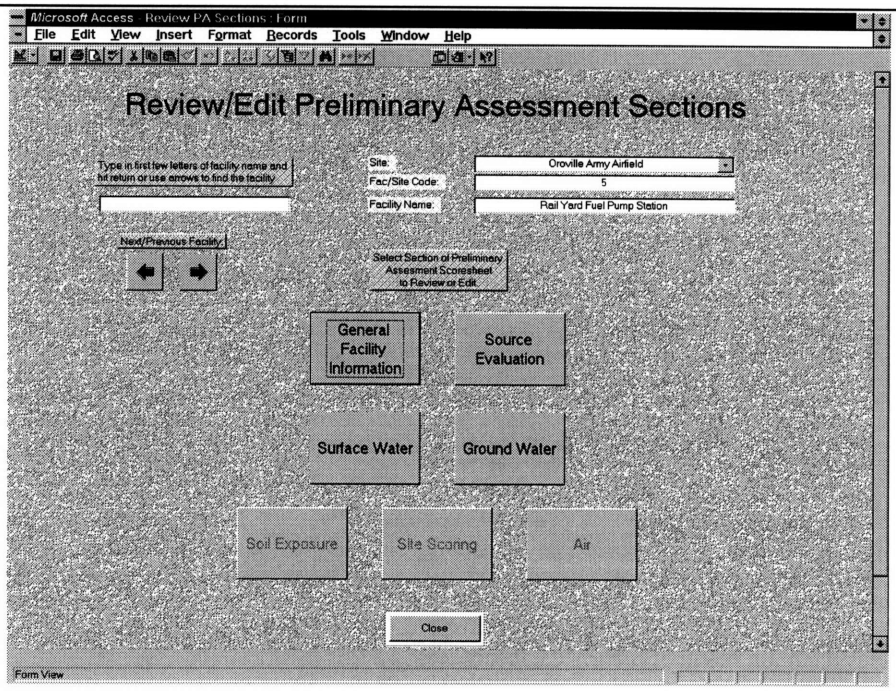
Form Data: none

Events and Macros: The form's On Open event triggers executing of the "Maximize" macro. The "Existing PA's" button's On Click event triggers the OpeningForms.PAReview macro. The "General Site Information" button's On Click event triggers the OpeningForms.SiteInformation macro. Lastly, the "Perform Site Comparisons" button's On Click event triggers the OpeningForms.Criteria macro.

Opened From: Automatically opened when the database file is opened.

Figure 5.2.1a Main Menu of the PA Database

When the user clicks on the Existing PA's button, the Review PA Sections form is opened. This form is bound to the GeneralFacilityInformation table, although the form only contains controls to display the SiteID, Fac/SiteID, and Facility Name. A property of the SiteID has been set to display the SiteName rather than the numerical code. Each of these control's Locked property has been set to "Yes", so that the user may not edit this information. Figure 5.2.1b shows the Review PA Sections form.



Form Name: Review PA Sections

Description: This is the menu for reviewing the various sections of the preliminary assessment scoresheet.

Subforms: none

Form Data: GeneralFacilityInformation table

Events and Macros: The form's On Open event triggers the Maximize macro. The "Close" button's On Click closes the form. Clicking on any of the buttons representing the sections of the PA executes a macro in the group macro OpeningForms. Clicking on the "General Facility Information" button opens the "Facility Information" form. Clicking on the "Source Evaluation" button opens the "Source Evaluations" form. Clicking on the "Surface Water" button opens the "Surface Water Main" form. Clicking the "Ground Water" button opens the form "Ground Water". Hitting return in the text box in which the first few letters of a known facility name are typed, executes the macro FindFacilityName. This macro makes the facility name control active (using the GoToControl action) and finds the record where the name matches the text entered by the user (using the FindRecord action).

Opened From: Main Switchboard

Figure 5.2.1b Main Menu for Reviewing PA Sections.

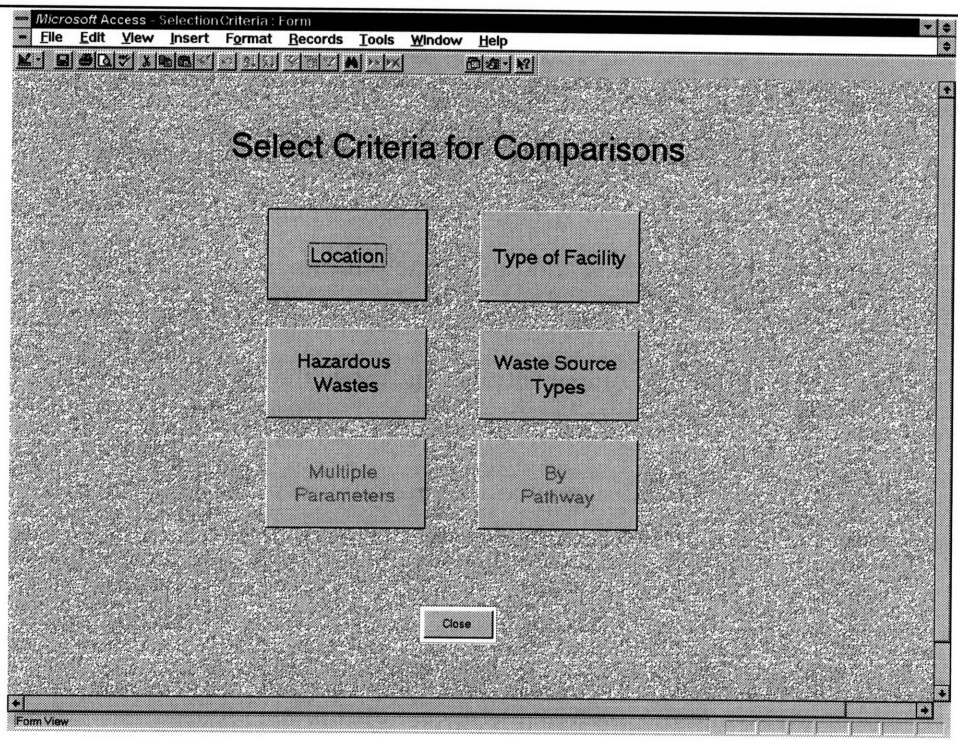
On this form, the user must first select the facility to review before clicking on one of the buttons which will open up the next form showing the section of the PA scoresheet the user wishes to review. To select a facility for review, the user has two options. If the name of the facility is known, the user can type in the first few letters of the facility name, hit return, and a macro will be run that searches for the record containing that facility name. If the user does not know the facility name, he or she may use the buttons with arrow icons to go through the stored facility records one by one.

The text on the Soil Exposure, Site Scoring, and Air buttons is grayed out because they do not actually open any forms. They have been included in the Review PA Sections form to show that these sections should be included if the PA database completely modeled the PA scoresheet. When the user clicks on any of these buttons, the macro Message.NotImplemented is executed, which displays a message box with the message, "This feature is not implemented."

Once a facility has been selected and one of the buttons has been clicked, a macro is executed that opens the appropriate form (Open Form action), gives "control" to the Fac/SiteID text box of the opened form (GoToControl action), and finds the record that matches the Fac/SiteID on the Review PA Sections form (FindRecord action).

The description of the forms associated with reviewing and editing the sections of the PA scoresheet is included in Appendix D.

From the Main Switchboard, if the user clicks on the Perform Site Comparisons button, the SelectionCriteria form is opened shown in Figure 5.2.1c. The buttons "Multiple Parameters" and "By Pathway" only execute the macro Messages.NotImplemented.



Form Name: SelectionCriteria

Description: This is the main menu for performing facility comparisons

Subforms: none

Form Data: none

Events and Macros: The macro group OpeningForms contains all of the macros for opening forms when any of the buttons are closed, except for the "Close" button, which closes the SelectionCriteria form. Clicking on the buttons, "Location", "Type of Facility", "Hazardous Wastes", and "Waste Source Types" execute the appropriate macros to open the forms "Location", "Type of Facility", "Hazardous Wastes", and "Waste Source", respectively.

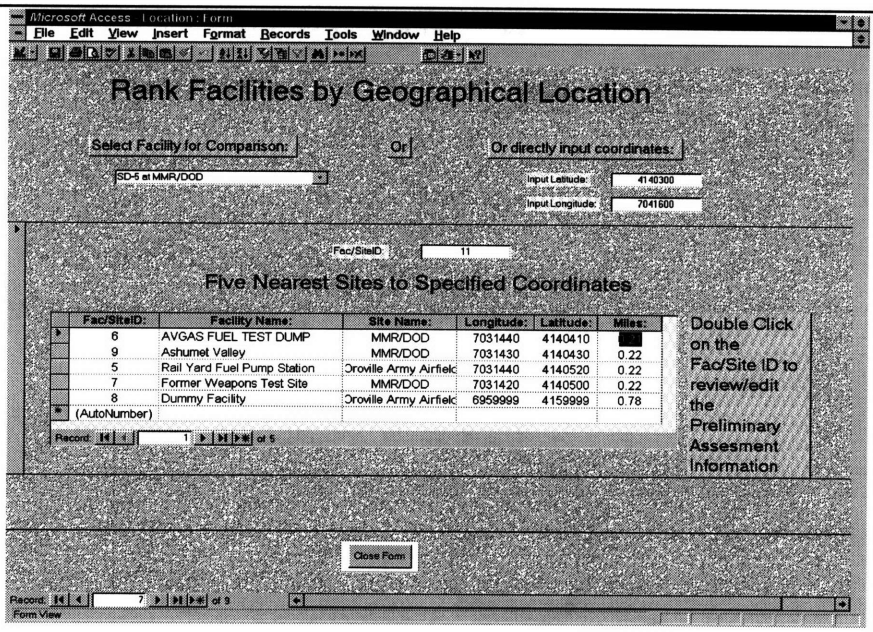
Opened From: Main Switchboard

Figure 5.2.1c Main Menu for Performing Facility Comparisons

5.2.2 LOCATION COMPARISON

The Location form becomes active when the "Location" button is clicked from the main menu for performing facility comparisons (SelectionCriteria form). The user either enters the geographical coordinates or selects a facility from which the comparisons will be based. The Location Subform then appears in the middle of the form showing the result of performing the

Location Comparison query. This query ranks the facilities by their proximity to the specified geographical coordinates. Figure 5.2.1d shows the Location form after geographical coordinates have been selected.



Description: Latitude and longitude coordinates are entered as input or the geographical coordinates of a selected facility are used to rank the top 5 closest facilities to the input coordinates.

Subforms: The "Location Subform" shows the result of the query comparing geographical coordinates with the user's input(Location Comparison query discussed in Section 4.5.1.)

Form Data: Based on the query "Location Form Query"

Events and Macros: When a facility is selected by the user, the Macro "Coordinates" is executed. This places the geographical coordinates of the selected facility into the Input Latitude and Input Longitude controls. Whenever the Input Latitude or Input Longitude coordinates are modified (an AfterUpdate event is generated), the Macro "Find Nearest Location" is executed. This macro updates the list of top 5 closest facilities. Upon the first time the Input Longitude is updated, the Visible properties of the Location Subform, the "Five Nearest Sites...", and "Double Click on the..." text boxes are set to "Yes".Double-clicking on the Fac/SiteID in the Location Subform executes the macro OpenMainPA.Location. This macro opens the form for reviewing the sections of the PA where the double-clicked facility has been selected for review.

Opened From: SelectionCriteria form

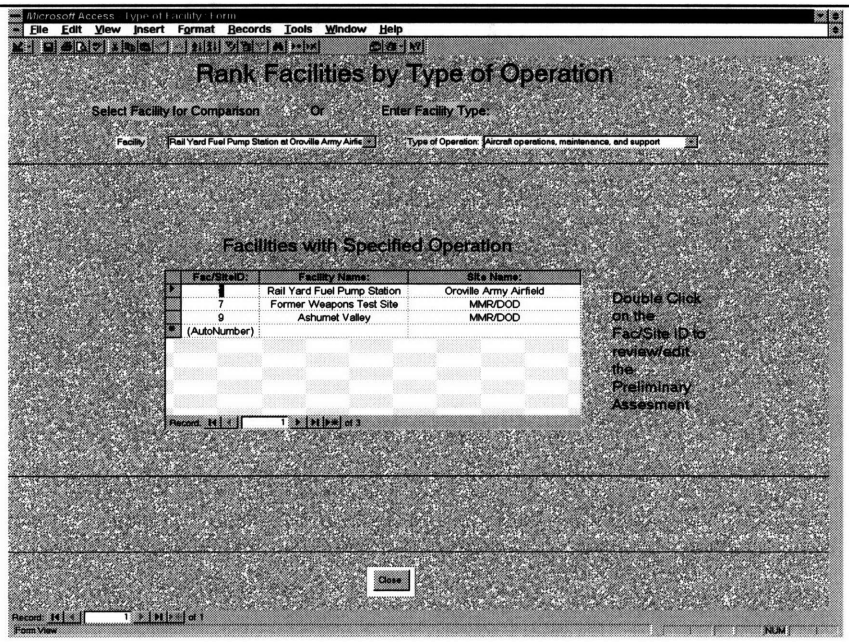
Figure 5.2.1d Location Comparison Form

From the list of ranked facilities, the user can double-click on the Fac/SiteID where the form for reviewing sections of the PA is opened. The macro that accomplishes this also synchronizes the form, so that the facility that was double-clicked is already chosen for review on the PA section

review form.

5.2.3 TYPE OF FACILITY COMPARISON

Comparisons between facilities based on the type of operation is performed in the Type of Facility form. This form is opened from the main menu for performing comparisons by clicking on the "Type of Facility" button (see Figure 5.2.1c). Similar to the form where facilities are ranked by location, the user either selects from a list of valid facility operations or selects a facility from which the comparison is based. The result is a list of facilities that have been characterized by the selected operation. Figure 5.2.3a shows the Type of Facility form.



Description: The type of operation is entered as input or a facility is selected from which the comparison is based.

Subforms: The "Type Subform" (based on TypeQuery query, which returns records from the GeneralFacilityInformation and GeneralSiteInformation tables where the operation code matches what was selected by the user on the Type of Facility form.)

Form Data: Based on the query "Location Form Query"

Events and Macros: When a facility is selected by the user, the macro "Type Update" is executed. This macro places the type of operation in the Type of Operation text box, makes the Type Subform visible (SetValue action) and requeries it (Requery action) The Type Subform displays the facilities that perform the type of operation selected by the user. Double-clicking on the Fac/SiteID in the Type Subform executes the macro OpenMainPA.Type which opens the main menu for reviewing PA sections.

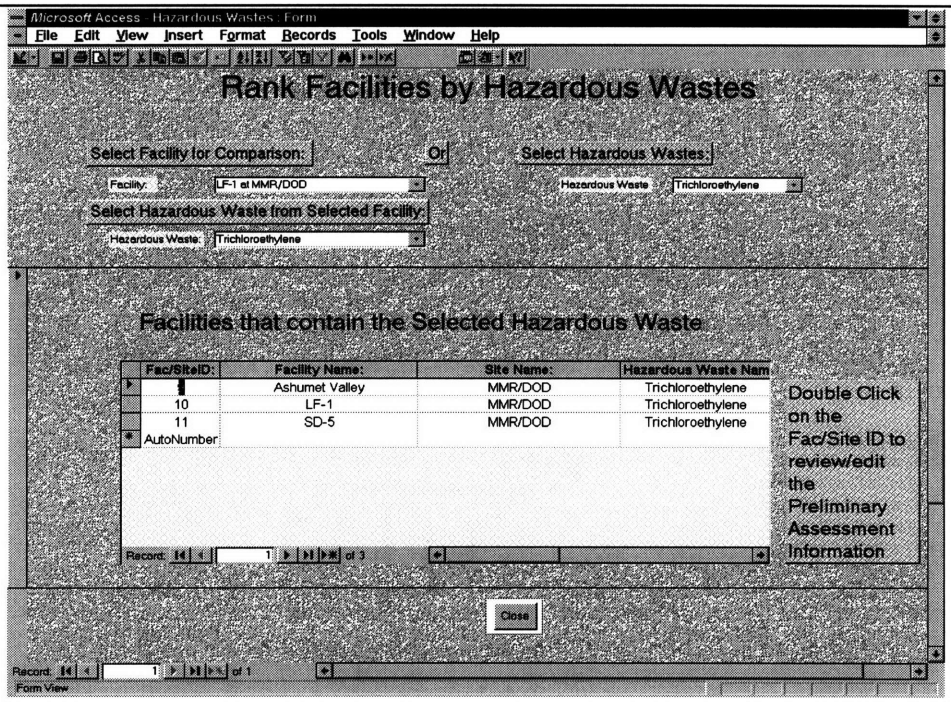
Opened From: SelectionCriteria form

Figure 5.2.3a Type of Facility Comparison Form

From the list of facilities, the user can double-click on the Fac/SiteID to open the form for reviewing sections of the PA. The macro that accomplishes this also synchronizes the form so that the facility that was double-clicked is already chosen for review on the PA section review form.

5.2.4 HAZARDOUS WASTES COMPARISON

Another relevant way to compare facilities is by the hazardous wastes they contain. Clicking on the "Hazardous Waste" button opens the Hazardous Wastes form. If the user selects a facility, he or she must also choose the hazardous waste that is contained on that facility before any comparisons are made. Alternately, a hazardous waste can be selected from a hazardous waste combo box. Once a hazardous substance has been selected, a subform appears listing the facilities that contain the hazardous substance. The user may then double click on a facility to review the desired section of the PA scoresheet of that facility. Figure 5.2.4a shows the Hazardous Wastes form after a particular hazardous waste has been selected.



Form Name: Hazardous Wastes

Description: After a hazardous waste is selected from a facility or a list of valid hazardous wastes, a list of facilities containing that waste is displayed.

Subforms: Hazardous Subform. This subform is based on the Hazardous Ranking Query which lists those facilities containing the specified hazardous waste in the form Hazardous Wastes

Form Data: none

Events and Macros: Selection of a facility triggers execution of the macro "Facility Hazardous" from an After Update event. This macro updates the list of hazardous wastes that may be chosen from the combo box located directly below the Facility combo box. When a hazardous waste is selected from this combo box, the "Set Waste" macro is executed which sets the value of the hazardous waste in the combo box located on the right side of the screen. This macro also displays and requeries the Hazardous Subform. The Hazardous Subform is displayed and requeried also whenever a hazardous substance is chosen from the hazardous waste combo box on the right hand side of the form.

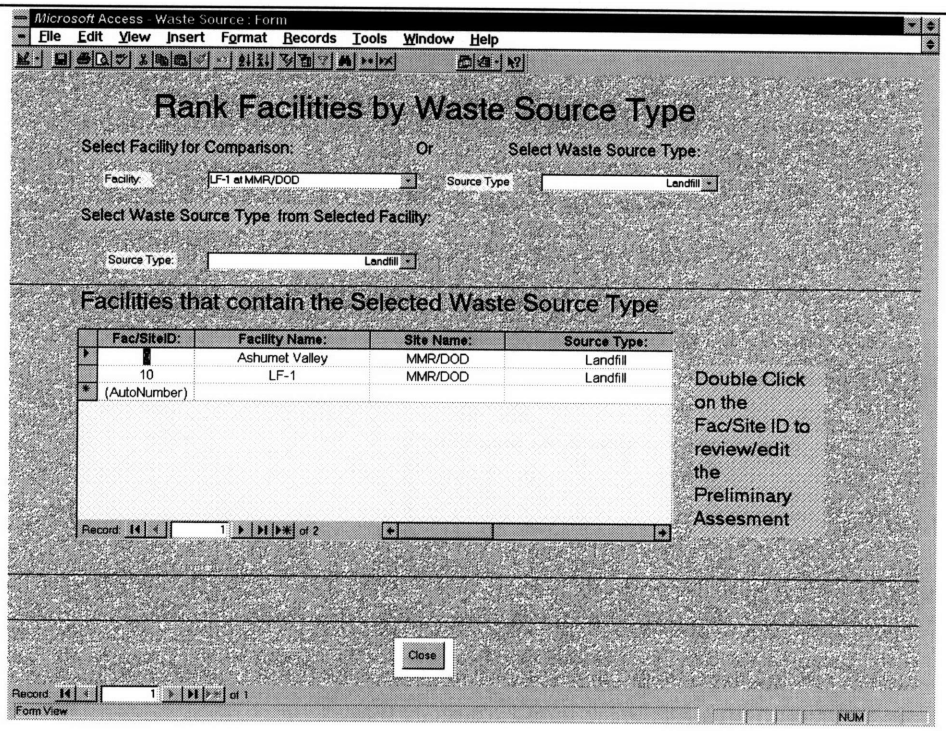
Opened From: SelectionCriteria form.

Figure 5.2.4a Hazardous Waste Comparison Form

5.2.5 WASTE SOURCE TYPE COMPARISON

The final type of comparison that was implemented in Microsoft Access is the comparison of facilities by type of waste source. This comparison mimics the comparison by

hazardous waste, where the user first selects a facility and then selects a waste source from that facility. Similar to the other comparisons, the user may also simply choose a waste source from the combo box on the upper left portion of the screen. Figure 5.2.5a shows the Waste Source form after a waste source has been selected.



Form Name: Waste Source

Description: This form lists the facilities that contain the source of waste specified by the user.

Subforms: Waste Source Subform. This subform is based on the Waste Subform Query which primarily returns the Source Evaluation table.

Form Data: none

Events and Macros: When a facility is selected from the combo box, the macro "Facility Waste Source" is executed which updates the list of waste sources that can be selected from the waste source combo box. When either of the waste source combo boxes is updated, the macro "Source Macro" is executed which makes the list of facilities that contain the specified waste source visible.

Opened From: SelectionCriteria form

Figure 5.2.5a Waste Source Comparison Form

6.0 Conclusions and Recommendations

This thesis focused on the development of an interactive database used for cross-site comparisons, but included the development of a complete Information System to support decision making during the Preliminary Assessment (PA) phase of the Superfund process. There are four aspects of the system which require further resolution; availability of electronic information relevant to the PA Scoresheets, the technology used, the system's expandability and the potential impacts of this system on the Preliminary Assessment phase of the Superfund process.

6.1 Electronic Information

6.1.1 AVAILABILITY

As mentioned in previous sections, the EIDSS was severely limited by the lack of suitably stored information (environmental consultant Kenneth Till was only able to locate one suitable data source). There is currently very little information relevant to the PA Scoresheets in an electronic format, and even less in a standard format. Thus, the scope of both the Internet Search Engine and the Data Store Search Engine is currently limited. Many of the questions on the scoresheets are conducive to being answered with the assistance of electronic media, but the questions of when the information will be standardized, and who will actually set the standards, remain unanswered. Until more of these sources appear, filling out the relevant small sections of the scoresheets electronically is more difficult than filling them out manually.

6.1.2 STANDARDIZATION

It should also be noted that the information required to complete the PA Scoresheets must be from sources reliable enough so that potential user of the system will be comfortable with the data. If the system is to eventually be set up so users pay for accounts to use the system, the sources must conform to industry and countrywide reliability standards. This problem is best illustrated by considering the case where information is duplicated in two or more data sources. Under such circumstances, there must be a method to determine which source will be used.

Possible determination methods include surveying the current clients to find their preference, and performing a thorough investigation on the origins of the data sources to determine which is more reliable.

6.2 Technology

A number of technology-related issues arose during the system's development. These issues restricted some aspects of the current system, but advancements in the tools used are anticipated, so future systems should not face such restrictions.

6.2.1 Data Source Collection

There were many obstacles encountered in trying to connect the scoresheet pages on the web to the data source itself. A large part of this was because the development of the Data Store Search Engine was first attempted completely using Java Database Connectivity (JDBC), a very new tool that is not yet reliable as May, 1997. As mentioned in Section 3.3, the system currently requires that the program executing the query on the data source and the data source itself be on the same server. So when more sources are gathered, either a program must be installed at the location of each data source, or all of the data sources must be brought on to a single server.

Using JDBC for the entire Search Engine, however, would eliminate this requirement. Thus, there could be one query program located on a server that is easily accessed by system administrators, and the data could be stored anywhere, as long as they are in the correct format and their owners give the system access permission. JDBC will almost definitely be more reliable within a year, so the possibility of using a more robust system is conceivable in the future.

6.2.2 Linking the EIS with the DSS

Currently, the system is a separated Executive Information System (EIS) and Decision Support System (DSS). The integration of these two parts was also greatly restricted by available technology, although again, these restrictions will likely disappear in the near future, as

the technology matures.

The basis of the EIS is a local database, which reşides on the same machine as the web pages that the user accesses. Each user's decisions are recorded on this database for later examination. The DSS links data sources in three different forms: an Internet Search Engine locating possible relevant data locations in non-standard form; an Data Store Search Engine retrieving data that is known to exist from standardized data sources; and a Database for Cross-Site Comparison, which allows retrieval of data from previously filled out PA Scoresheets. More database types, however, are becoming easier to connect to the Internet through the use of tools that easily display query results or table values on a web page. This facility makes the inclusion of EIS access from a DSS link very easy to implement. The versatility and reliability of these web connections, however, is not yet reliable, though is expected to be so in the near future.

6.3 Improvements to the Interactive Database of PA Scoresheets for Cross-Site Comparison

The current database is limited to performing several specific comparisons. The system could be greatly enhanced by performing comparisons based on additional sets of criteria. For example, it would be useful to search all scoresheets by answers to individual questions or combinations of questions. Using SQL, these types of searches are not difficult, but it would be beneficial to set up an interface such as those demonstrated with Microsoft Access in order to make it easier to perform these types of comparisons.

6.4 Expandability - Beyond the Preliminary Assessment Phase

Due to the scope of the project, the information system is currently limited to the Preliminary Assessment phase. Using the EIDSS principle of combining a document-like user-interface, a search engine and a database management system, however, a similar information system could be designed and used in other steps of the Superfund process. This expanded system would then allow users to easily refer to, and use, information from various Superfund

process steps throughout their work on a particular site.

6.5 Impact of the System on the Preliminary Assessment Phase

This system provides numerous advantages to its users at all levels. First of all, the documentation needed to keep track of the PA process will be greatly reduced. The questions that were filled out electronically will be stored electronically, along with the bibliographical information. In addition, if there is more than one person scoring a certain site, much of the confusion accompanying trading the papers and other documentation will be eliminated by the accessibility of the forms over the Internet. Anyone with permission who needs to access the forms can do so from any office using a desktop computer, or from the field using a laptop. Furthermore, the electronic format of the forms eliminates the need to remember to bring the proper forms or to give them to the proper person before they leave for the site, etc. It is also much faster to answer questions using the EIDSS, than to manually search out, and go through, data sources stored in many different locations.

On a management level, this system provides many advantages to executives as well. The EIS portion provides greater accountability both to the people filling out the scoresheets and the sources being used for the information entered into these scoresheets. Having the exact source, date and time of a particular decision can greatly simplify the assessment process if particular decisions or sources need to be questioned. Indeed, the system can even be used to identify the more efficient engineers filling out the scoresheets and reward or promote them accordingly.

Although this system is currently limited by certain constraints, overall it has great potential to assist in the Preliminary Assessment process. Perhaps with backing from a few Environmental Engineering firms, data standards can be set and reliable data sources compiled so that future preliminary assessments will be far less error-prone, time-consuming and costly than at present.

7.0 References

Agency for Toxic Substance and Disease Control. *1992 Priority List of Hazardous Substances,* April 15, 1997 (http://www.medaccess.com/chemicals/92list.htm)

Applegate, Lynda M."Frito-Lay Inc.: A Strategic Transition—C," *Harvard Business School Case Study 9-190-071*, 1989.

Bilicki, Ed. *Environmental Decision Support System*, May 28, 1996 (http://www.iceis.mcnc.org/EDSS/EDSSPage.html).

Environmental Protection Agency. *CERCLIS Entity & Attribute Information*, April 15, 1997 (http://envirofacts.sql.epa.gov/enviro/html/cerclis/site_lat.html).

Bronsema, Gloria and Shoshana Zuboff. "The Expense Tracking System at Tiger Creek," *Harvard Business School Case Study 485-057*, 1984.

Dickson, John N. *Environmental Search Engine*, March 10, 1997 (http://www.webdirectory.com/)

Filo, David and Jerry Yang. *YAHOO*, March 10, 1997 (http://www.yahoo.com/Society_and_Culture/Environment_and_Nature/)

Garvey, Tom. *EnviroPhantom*, March 10, 1997 (http://phantom.rec.org/)

Gifford, Fred. *WWW Geographic Information Sources*, March 6, 1997 (http://nris.msl.mt.gov/gis/other.html)

Hawryszkiewycz, Igor T. *An Introduction to Relational Database Design*, Sydney: Prentice Hall, 1990.

Hoffman, James. *Introduction to the Structured Query Language,* April 15, 1997 (http://w3.one.net/%7Ejhoffman/sqltut.htm)

Kuo, Kai. *Web-based Database Enabled Executive Information System for Preliminary Site Assessment under CERCLA*, Master of Engineering Thesis, Department of Civil and Environemntal Engineering, Massachusetts Institute of Technology: Cambridge, 1997.

Johnson, Dr. Lynn E. *Colorado River Decision Support System*, Feb 14, 1996 (http://cando.dwr.co.gov/overview/bigoverview/crdsscov.html)

Lukasiak, Anna. *Internet-Enabled Integrated Presentation System for Site Remediation Preliminary Assessment*, Master of Engineering Thesis, Department of Civil and Environemntal Engineering, Massachusetts Institute of Technology: Cambridge, 1997.

Mukhopadhyay, S. Rony. *Development of Data Store Search Engine for Surface Water Pathway Criteria List*, Master of Engineering Thesis, Department of Civil and Environemntal Engineering, Massachusetts Institute of Technology: Cambridge, 1997.

Nolan, Richard L. and Donna B. Stoddard. "CIGNA Property and Casualty Reengineering (A)," *Harvard Business School Case Study 9-196-059*, 1995

Leidner, Dorothy E. and Joyce J. Elam, "Executive Information System: Their Impact on Executive Decision Making," *Journal of Management Information Systems*, Winter 1993-94, 10:3, page139-155

Patel, Pratik and Karl Moss. *Java Database Programming with JDBC*. Scottsdale: Arizona, Coriolis Group, 1996.

Rittner, Don. *Ecolinking*, February 7, 1997
(http://www.themesh.com/elinks.html)

Somlyódy, László. *Decision Support System for Evaluation River Basin Strategies*, June 26, 1996
(http://www.iiasa.ac.at/Research/WAT/docs/desert.html)

Stanczyk, Stefan. *Theory and Practice of Relational Databases*, London: Pitman Publishing, 1990.

Appendix A: Preliminary Assessment Scoresheets

OMB Approval Number: 2050-0095
Approved for Use Through: 1/92

PA Scoresheets

Site Name: _____
CERCLIS ID No.: _____
Street Address: _____
City/State/Zip: _____

Investigator: _____
Agency Organization: _____
Street Address: _____
City/State/Zip: _____
Date: _____

INSTRUCTIONS FOR SCORESHEETS

Introduction

This scoresheets package functions as a self-contained workbook providing all of the basic tools to apply collected data and calculate a PA score. Note that a computerized scoring tool, "PA-Score," is also available from EPA (Office of Solid Waste and Emergency Response, Directive 9345.1-11). The scoresheets provide space to:

- Record information collected during the PA
- Indicate references to support information
- Select and assign values ("scores") for factors
- Calculate pathway scores
- Calculate the site score

Do not enter values or scores in shaded areas of the scoresheets. You are encouraged to write notes on the scoresheets and especially on the Criteria Lists. On scoresheets with a reference column, indicate a number corresponding to attached sources of information or pages containing rationale for hypotheses; attach to the scoresheets a numbered list of these references. Evaluate all four pathways. Complete all Criteria Lists, scoresheets, and tables. Show calculations, as appropriate. If scoresheets are photocopy reproduced, copy and submit the numbered pages (right-side pages) only.

GENERAL INFORMATION

Site Description and Operational History: Briefly describe the site and its operating history. Provide the site name, owner/operator, type of facility and operations, size of property, active or inactive status, and years of waste generation. Summarize waste treatment, storage, or disposal activities that have or may have occurred at the site; note also if these activities are documented or alleged. Identify probable source types and prior spills. Summarize highlights of previous investigations.

Probable Substances of Concern: List hazardous substances that have or may have been stored, handled, or disposed at the site, based on your knowledge of site operations. Identify the sources to which the substances may be related. Summarize any existing analytical data concerning hazardous substances detected onsite, in releases from the site, or at targets.

GENERAL INFORMATION

Site Description and Operational History:

Probable Substances of Concern:
(Previous investigations, analytical data)

A-3

GENERAL INFORMATION (continued)

Site Sketch: Prepare a sketch of the site (freehand is acceptable). Indicate all pertinent features of the site and nearby environs, including: waste sources, buildings, residences, access roads, parking areas, drainage patterns, water bodies, vegetation, wells, sensitive environments, etc.

GENERAL INFORMATION (continued)

Site Sketch:
(Show all pertinent features, indicate sources and closest targets, indicate north)

A-5

SOURCE EVALUATION

- Number and name each source (e.g., 1. East Drum Storage Area, 2. Sludge Lagoon, 3. Battery Pile).
- Identify source type according to the list below.
- Describe the physical character of each source (e.g., dimensions, contents, waste types, containment, operating history).
- Show waste quantity (WQ) calculations for each source for appropriate tiers. Refer to instructions opposite page 5 and PA Tables 1a and 1b. Identify waste quantity tier and waste characteristics (WC) factor category score (for a site with a single source, according to PA Table 1a). Determine WC from PA Table 1b for the sum of source WQs for a multiple-source site.
- Attach additional sheets if necessary.
- Determine the site WC factor category score and record at the bottom of the page.

Source Type Descriptions

Landfill: an engineered (by excavation or construction) or natural hole in the ground into which wastes have been disposed by backfilling, or by contemporaneous soil deposition with waste disposal, covering wastes from view.

Surface Impoundment: a topographic depression, excavation, or diked area, primarily formed from earthen materials (lined or unlined) and designed to hold accumulated liquid wastes, wastes containing free liquids, or sludges that were not backfilled or otherwise covered during periods of deposition; depression may be dry if deposited liquid has evaporated, volatilized or leached, or wet with exposed liquid; structures that may be more specifically described as lagoon pond, aeration pit, settling pond, tailings pond, sludge pit, etc.; also a surface impoundment that has been covered with soil after the final deposition of waste materials (i.e., buried or backfilled).

Drums: portable containers designed to hold a standard 55-gallon volume of wastes.

Tanks and Non-Drum Containers: any stationary device, designed to contain accumulated wastes, constructed primarily of fabricated materials (such as wood, concrete, steel, or plastic) that provide structural support; any portable or mobile device in which waste is stored or otherwise handled.

Contaminated Soil: soil onto which available evidence indicates that a hazardous substance was spilled, spread, disposed, or deposited.

Pile: any non-containerized accumulation above the ground surface of solid, non-flowing wastes; includes open dumps. Some types of piles are: Chemical Waste Pile – consists primarily of discarded chemical products, by-products, radioactive wastes, or used or unused feedstocks; Scrap Metal or Junk Pile – consists primarily of scrap metal or discarded durable goods such as appliances, automobiles, auto parts, or batteries, composed of materials suspected to contain or have contained a hazardous substance; Tailings Pile – consists primarily of any combination of overburden from a mining operation and tailings from a mineral mining, beneficiation, or processing operation; Trash Pile – consists primarily of paper, garbage, or discarded non-durable goods which are suspected to contain or have contained a hazardous substance.

Land Treatment: landfarming or other land treatment method of waste management in which liquid wastes or sludges are spread over land and tilled, or liquids are injected at shallow depths into soils.

Other: a source that does not fit any of the descriptions above; examples include contaminated building, ground water plume with no identifiable source, storm drain, dry well, and injection well.

SOURCE EVALUATION

Source No.:	Source Name:	Source Waste Quantity (WQ) Calculations:
Source Description:		

Source No.:	Source Name:	Source Waste Quantity (WQ) Calculations:
Source Description:		

Source No.:	Source Name:	Source Waste Quantity (WQ) Calculations:
Source Description:		

Site WC:

A-7

WASTE CHARACTERISTICS (WC) SCORES

WC, based on waste quantity, may be determined by one or all of four measures called "tiers": constituent quantity, wastestream quantity, source volume, and source area. PA Table 1a (page 5) is divided into these four tiers. The amount and detail of information available determine which tier(s) to use for each source. For each source, evaluate waste quantity by as many of the tiers as you have information to support, and select the result that gives you the highest WC score. If minimal, incomplete, or no information is available regarding waste quantity, assign a WC score of 18 (minimum).

PA Table 1a has 6 columns: column 1 indicates the quantity tier; column 2 lists source types for the four tiers; columns 3, 4, and 5 provide ranges of waste amount for sites with only one source, which correspond to WC scores at the top of the columns (18, 32, or 100); column 6 provides formulas to obtain source waste quantity (WQ) values at sites with multiple sources.

To determine WC for sites with only one source:

1. *Identify source type (see descriptions opposite page 4).*
2. *Examine all waste quantity data available.*
3. *Estimate the mass and/or dimensions of the source.*
4. *Determine which quantity tiers to use based on available source information.*
5. *Convert source measurements to appropriate units for each tier you can evaluate for the source.*
6. *Identify the range into which the total quantity falls for each tier evaluated (PA Table 1a).*
7. *Determine the highest WC score obtained for any tier (18, 32, or 100, at top of PA Table 1a columns 3, 4, and 5, respectively).*
8. *Use this WC score for all pathways.**

To determine WC for sites with multiple sources:

1. *Identify each source type (see descriptions opposite page 4).*
2. *Examine all waste quantity data available for each source.*
3. *Estimate the mass and/or dimensions of each source.*
4. *Determine which quantity tiers to use for each source based on the available information.*
5. *Convert source measurements to appropriate units for each tier you can evaluate for each source.*
6. *For each source, use the formulas in column 6 of PA Table 1a to determine the WQ value for each tier that can be evaluated. The highest WQ value obtained for any tier is the WQ value for the source.*
7. *Sum the WQ values for all sources to get the site WQ total.*
8. *Use the site WQ total from step 7 to assign the WC score from PA Table 1b.*
9. *Use this WC score for all pathways.**

---

* The WC score is considered in all four pathways. However, if a primary target is identified for the groundwater, surface water, or air migration pathway, assign the determined WC or a score of 32, whichever is greater, as the WC score for that pathway.

PA TABLE 1: WASTE CHARACTERISTICS (WC) SCORES

PA Table 1a: WC Scores for Single Source Sites and Formulas for Multiple Source Sites

TIER	SOURCE TYPE	SINGLE SOURCE SITES (assigned WC scores)			MULTIPLE SOURCE SITES
		WC = 18	WC = 32	WC = 100	Formula for Assigning Source WQ Values
CONSTITUENT	N/A	≤100 lb	>100 to 10,000 lb	>10,000 lb	lb ÷ 1
WASTESTREAM	N/A	≤500,000 lb	>500,000 to 50 million lb	>50 million lb	lb ÷ 5,000
VOLUME	Landfill	≤6.75 million ft³ ≤250,000 yd³	>6.75 million to 675 million ft³ >250,000 to 25 million yd³	>675 million ft³ >25 million yd³	ft³ ÷ 67,500 yd³ ÷ 2,500
	Surface impoundment	≤6,750 ft³ ≤250 yd³	>6,750 to 675,000 ft³ >250 to 25,000 yd³	>675,000 ft³ >25,000 yd³	ft³ ÷ 67.5 yd³ ÷ 2.5
	Drums	≤1,000 drums	>1,000 to 100,000 drums	>100,000 drums	drums ÷ 10
	Tanks and non-drum containers	≤50,000 gallons	>50,000 to 5 million gallons	>5 million gallons	gallons ÷ 500
	Contaminated soil	≤6.75 million ft² ≤250,000 yd²	>6.75 million to 675 million ft² >250,000 to 25 million yd²	>675 million ft² >25 million yd²	ft² ÷ 67,500 yd² ÷ 2,500
	Pile	≤6,750 ft² ≤250 yd²	>6,750 to 675,000 ft² >250 to 25,000 yd²	>675,000 ft² >25,000 yd²	ft² ÷ 67.5 yd² ÷ 2.5
	Other	≤6,750 ft² ≤250 yd²	>6,750 to 675,000 ft² >250 to 25,000 yd²	>675,000 ft² >25,000 yd²	ft² ÷ 67.5 yd² ÷ 2.5
AREA	Landfill	≤340,000 ft² ≤7.8 acres	>340,000 to 34 million ft² >7.8 to 780 acres	>34 million ft² >780 acres	ft² ÷ 3,400 acres ÷ 0.078
	Surface impoundment	≤1,300 ft² ≤0.028 acres	>1,300 to 130,000 ft² >0.028 to 2.9 acres	>130,000 ft² >2.9 acres	ft² ÷ 13 acres ÷ 0.00029
	Contaminated soil	≤3.4 million ft² ≤78 acres	>3.4 million to 340 million ft² >78 to 7,800 acres	>340 million ft² >7,800 acres	ft² ÷ 34,000 acres ÷ 0.78
	Pile*	≤1,300 ft² ≤0.028 acres	>1,300 to 130,000 ft² >0.028 to 2.9 acres	>130,000 ft² >2.9 acres	ft² ÷ 13 acres ÷ 0.00029
	Land treatment	≤27,000 ft² ≤0.62 acres	>27,000 to 2.7 million ft² >0.62 to 62 acres	>2.7 million ft² >62 acres	ft² ÷ 270 acres ÷ 0.0062

1 ton = 2,000 lb = 1 yd³ = 4 drums = 200 gallons

* Use area of land surface under pile, not surface area of pile.

PA Table 1b: WC Scores for Multiple Source Sites

WQ Total	WC Score
>0 to 100	18
>100 to 10,000	32
>10,000	100

GROUND WATER PATHWAY

Ground Water Use Description: Provide information on ground water use in the vicinity. Present the general stratigraphy, aquifers used, and distribution of private and municipal wells.

Calculations for Drinking Water Populations Served by Ground Water: Provide populations from private wells and municipal supply systems in each distance category. Show apportionment calculations for blended supply systems.

GROUND WATER PATHWAY
GROUND WATER USE DESCRIPTION

Describe Ground Water Use Within 4-miles of the Site:
(Describe stratigraphy, information on aquifers, municipal and/or private wells)

Calculations for Drinking Water Populations Served by Ground Water:

A-11

GROUND WATER PATHWAY CRITERIA LIST

This "Criteria List" helps guide the process of developing hypotheses concerning the occurrence of a suspected release and the exposure of specific targets to a hazardous substance. The check-boxes record your professional judgment in evaluating these factors. Answers to all of the listed questions may not be available during the PA. Also, the list is not all-inclusive; if other criteria help shape your hypotheses, list them at the bottom of the page or attach an additional page.

The "Suspected Release" section identifies several site, source, and pathway conditions that could provide insight as to whether a release from the site is likely to have occurred. If a release is suspected, use the "Primary Targets" section to evaluate conditions that may help identify targets likely to be exposed to a hazardous substance. Record responses for the well that you feel has the highest probability of being exposed to a hazardous substance. You may use this section of the chart more than once, depending on the number of targets you feel may be considered "primary."

Check the boxes to indicate a "yes," "no," or "unknown" answer to each question. If you check the "Suspected Release" box as "yes," make sure you assign a Likelihood of Release value of 550 for the pathway.

GROUND WATER PATHWAY CRITERIA LIST

SUSPECTED RELEASE

Yes	No	Unk	
☐	☐	☐	Are sources poorly contained?
☐	☐	☐	Is the source a type likely to contribute to ground water contamination (e.g., wet lagoon)?
☐	☐	☐	Is waste quantity particularly large?
☐	☐	☐	Is precipitation heavy?
☐	☐	☐	Is the infiltration rate high?
☐	☐		Is the site located in an area of karst terrain?
☐	☐	☐	Is the subsurface highly permeable or conductive?
☐	☐	☐	Is drinking water drawn from a shallow aquifer?
☐	☐	☐	Are suspected contaminants highly mobile in ground water?
☐	☐	☐	Does analytical or circumstantial evidence suggest ground water contamination?
☐	☐		Other criteria?_____
☐	☐		**SUSPECTED RELEASE?**

Summarize the rationale for Suspected Release (attach an additional page if necessary):

PRIMARY TARGETS

Yes	No	Unk	
☐	☐	☐	Is any drinking water well nearby?
☐	☐	☐	Has any nearby drinking water well been closed?
☐	☐	☐	Has any nearby drinking water user reported foul-tasting or foul-smelling water?
☐	☐	☐	Does any nearby well have a large drawdown or high production rate?
☐	☐	☐	Is any drinking water well located between the site and other wells that are suspected to be exposed to a hazardous substance?
☐	☐	☐	Does analytical or circumstantial evidence suggest contamination at a drinking water well?
☐	☐	☐	Does any drinking water well warrant sampling?
☐	☐		Other criteria?_____
☐	☐		**PRIMARY TARGET(S) IDENTIFIED?**

Summarize the rationale for Primary Targets (attach an additional page if necessary):

A-13

GROUND WATER PATHWAY SCORESHEET

Pathway Characteristics
Answer the questions at the top of the page. Refer to the Ground Water Pathway Criteria List (page 7) to hypothesize whether you suspect that a hazardous substance associated with the site has been released to ground water. Record depth to aquifer (in feet): the difference between the deepest occurrence of a hazardous substance and the depth of the top of the shallowest aquifer at (or as near as possible) to the site. Note whether the site is in karst terrain (characterized by abrupt ridges, sink holes, caverns, springs, disappearing streams). Record the distance (in feet) from any source to the nearest well used for drinking water.

Likelihood of Release (LR)

1. **Suspected Release:** Hypothesize based on professional judgment guided by the Ground Water Pathway Criteria List (page 7). If you suspect a release to ground water, use only Column A for this pathway and do not evaluate factor 2.

2. **No Suspected Release:** If you do not suspect a release, determine score based on depth to aquifer or whether the site is in an area of karst terrain. If you do not suspect a release to ground water, use only Column B to score this pathway.

Targets (T)

This factor category evaluates the threat to populations obtaining drinking water from ground water. To apportion populations served by blended drinking water supply systems, determine the percentage of population served by each well based on its production.

3. **Primary Target Population:** Evaluate populations served by all drinking water wells that you suspect have been exposed to a hazardous substance released from the site. Use professional judgment guided by the Ground Water Pathway Criteria List (page 7) to make this determination. In the space provided, enter the population served by any wells you suspect have been exposed to a hazardous substance from the site. If only the number of residences is known, use the average county residents per household (rounded up to the next integer) determine population served. Multiply the population by 10 to determine the Primary Target Population scor Note that if you do not suspect a release, there can be no primary target population.

4. **Secondary Target Population:** Evaluate populations served by all drinking water wells within 4 miles that you do not suspect have been exposed to a hazardous substance. Use PA Table 2a or 2b (for wells drawing from non-karst and karst aquifers, respectfully) (page 9). If only the number of residences is known, use the average county residents per household (rounded to the nearest integer) to determine population served. Circle the assigned value for the population in each distance category and enter it in the column on the far-right side of the table. Sum the far-right column and enter the total as the Secondary Target Population factor score.

5. **Nearest Well** represents the threat posed to the drinking water well that is most likely to be exposed to a hazardous substance. If you have identified a primary target population, enter 50. Otherwise, assign the score from PA Table 2a or 2b for the closest distance category with a drinking water well population.

6. **Wellhead Protection Area (WHPA):** WHPAs are special areas designated by States for protection under Section 1428 of the Safe Drinking Water Act. Local/State and EPA Regional water officials can provide information regarding the location of WHPAs.

7. **Resources:** A score of 5 can generally be assigned as a default measure. Assign zero only if ground water within 4 miles has no resource use.

Sum the target scores in Column A (Suspected Release) or Column B (No Suspected Release).

Waste Characteristics (WC)

8. **Waste Characteristics:** Score is assigned from page 4. However, if you have identified any primary target for ground water, assign either the score calculated on page 4 or a score of 32, whichever is greater.

Ground Water Pathway Score: Multiply the scores for LR, T, and WC. Divide the product by 82,500. Round the result to the nearest integer. If the result is greater than 100, assign 100.

GROUND WATER PATHWAY SCORESHEET

Pathway Characteristics	
Do you suspect a release (see Ground Water Pathway Criteria List, page 71)?	Yes ___ No ___
Is the site located in karst terrain?	Yes ___ No ___
Depth to aquifer:	_____ ft
Distance to the nearest drinking water well:	_____ ft

	A *Suspected Release*	B *No Suspected Release*	*Reference*
LIKELIHOOD OF RELEASE			
1. SUSPECTED RELEASE: If you suspect a release to ground water (see page 7), assign a score of 550. Use only column A for this pathway.			
2. NO SUSPECTED RELEASE: If you do not suspect a release to ground water, and the site is in karst terrain or the depth to aquifer is 70 feet or less, assign a score of 500; otherwise, assign a score of 340. Use only column B for this pathway.			
LR =			
TARGETS			
3. PRIMARY TARGET POPULATION: Determine the number of people served by drinking water wells that you suspect have been exposed to a hazardous substance from the site (see Ground Water Pathway Criteria List, page 7). _____ people x 10 =			
4. SECONDARY TARGET POPULATION: Determine the number of people served by drinking water wells that you do NOT suspect have been exposed to a hazardous substance from the site, and assign the total population score from PA Table 2. Are any wells part of a blended system? Yes ___ No ___ If yes, attach a page to show apportionment calculations.			
5. NEAREST WELL: If you have identified a primary target population for ground water, assign a score of 50; otherwise, assign the Nearest Well score from PA Table 2. If no drinking water wells exist within 4 miles, assign a score of zero.			
6. WELLHEAD PROTECTION AREA (WHPA): If any source lies within or above a WHPA, or if you have identified any primary target well within a WHPA, assign a score of 20; assign 5 if neither condition holds but a WHPA is present within 4 miles; otherwise assign zero.			
7. RESOURCES			
T =			
WASTE CHARACTERISTICS			
8. A. If you have identified any primary target for ground water, assign the waste characteristics score calculated on page 4, or a score of 32, whichever is GREATER; do not evaluate part B of this factor. B. If you have NOT identified any primary target for ground water, assign the waste characteristics score calculated on page 4.			
WC =			
GROUND WATER PATHWAY SCORE: $\frac{LR \times T \times WC}{82,500}$			

PA TABLE 2: VALUES FOR SECONDARY GROUND WATER TARGET POPULATIONS

PA Table 2a: Non-Karst Aquifers

Distance from Site	Population	Nearest Well (choose highest)	Population Served by Wells Within Distance Category: 1 to 10	11 to 30	31 to 100	101 to 300	301 to 1,000	1,001 to 3,000	3,001 to 10,000	10,001 to 30,000	30,001 to 100,000	Greater than 100,000	Population Value
0 to ¼ mile	_____	20	1	2	5	16	52	163	521	1,633	5,214	16,325	_____
> ¼ to ½ mile	_____	10	1	1	3	10	32	101	323	1,012	3,233	10,121	_____
> ½ to 1 mile	_____	9	1	1	2	5	17	52	167	522	1,668	5,224	_____
> 1 to 2 miles	_____	5	1	1	1	3	9	29	94	294	939	2,938	_____
> 2 to 3 miles	_____	3	1	1	1	2	7	21	68	212	678	2,122	_____
> 3 to 4 miles	_____	2	1	1	1	1	4	13	42	131	417	1,308	_____
Nearest Well =												Score =	

PA Table 2b: Karst Aquifers

Distance from Site	Population	Nearest Well (use 20 for karst)	Population Served by Wells Within Distance Category: 1 to 10	11 to 30	31 to 100	101 to 300	301 to 1,000	1,001 to 3,000	3,001 to 10,000	10,001 to 30,000	30,001 to 100,000	Greater than 100,000	Population Value
0 to ¼ mile	_____	20	1	2	5	16	52	163	521	1,633	5,214	16,325	_____
> ¼ to ½ mile	_____	20	1	1	3	10	32	101	323	1,012	3,233	10,121	_____
> ½ to 1 mile	_____	20	1	1	3	8	26	82	261	816	2,607	8,162	_____
> 1 to 2 miles	_____	20	1	1	3	8	26	82	261	816	2,607	8,162	_____
> 2 to 3 miles	_____	20	1	1	3	8	26	82	261	816	2,607	8,162	_____
> 3 to 4 miles	_____	20	1	1	3	8	26	82	261	816	2,607	8,162	_____
Nearest Well =												Score =	

SURFACE WATER PATHWAY

Migration Route Sketch: Sketch the surface water migration pathway (freehand is acceptable) illustrating the drainage route and identifying water bodies, probable point of entry, flows, and targets.

A-18

SURFACE WATER PATHWAY
MIGRATION ROUTE SKETCH

Suface Water Migration Route Sketch:
(include runoff route, probable point of entry, 15-mile target distance limit, intakes, fisheries, and sensitive environments)

A-19

SURFACE WATER PATHWAY CRITERIA LIST

This "Criteria List" helps guide the process of developing hypotheses concerning the occurrence of suspected release and the exposure of specific targets to a hazardous substance. The check-boxes record your professional judgment in evaluating these factors. Answers to all of the listed questions may not be available during the PA. Also, the list is not all-inclusive; if other criteria help shape your hypotheses, list them at the bottom of the page or attach an additional page.

The "Suspected Release" section identifies several site, source, and pathway conditions that could provide insight as to whether a release from the site is likely to have occurred. If a release is suspected, use the "Primary Targets" section to guide you through evaluation of some conditions that may help identify targets likely to be exposed to a hazardous substance. Record responses for the target that you feel has the highest probability of being exposed to a hazardous substance. You may use this section of the chart more than once, depending on the number of targets you feel may be considered "primary."

Check the boxes to indicate a "yes," "no," or "unknown" answer to each question. If you check the "Suspected Release" box as "yes," make sure you assign a Likelihood of Release value of 550 for the pathway.

If the distance to surface water is greater than 2 miles, do not evaluate the surface water migration pathway. Document the source of information in the text boxes below the surface water criteria list.

SURFACE WATER PATHWAY CRITERIA LIST

SUSPECTED RELEASE

Yes	No	Unk	
☐	☐	☐	Is surface water nearby?
☐	☐	☐	Is waste quantity particularly large?
☐	☐	☐	Is the drainage area large?
☐	☐	☐	Is rainfall heavy?
☐	☐	☐	Is the infiltration rate low?
☐	☐	☐	Are sources poorly contained or prone to runoff or flooding?
☐	☐	☐	Is a runoff route well defined (e.g., ditch or channel leading to surface water)?
☐	☐	☐	Is vegetation stressed along the probable runoff route?
☐	☐	☐	Are sediments or water unnaturally discolored?
☐	☐	☐	Is wildlife unnaturally absent?
☐	☐	☐	Has deposition of waste into surface water been observed?
☐	☐	☐	Is ground water discharge to surface water likely?
☐	☐	☐	Does analytical or circumstantial evidence suggest surface water contamination?
☐	☐		Other criteria?_____
☐	☐		SUSPECTED RELEASE?

PRIMARY TARGETS

Yes	No	Unk	
☐	☐	☐	Is any target nearby? If yes:
			☐ Drinking water intake ☐ Fishery ☐ Sensitive environment
☐	☐	☐	Has any intake, fishery, or recreational area been closed?
☐	☐	☐	Does analytical or circumstantial evidence suggest surface water contamination at or downstream of a target?
☐	☐	☐	Does any target warrant sampling? If yes:
			☐ Drinking water intake ☐ Fishery ☐ Sensitive environment
☐	☐		Other criteria?_____
☐	☐		PRIMARY INTAKE(S) IDENTIFIED?
☐	☐		PRIMARY FISHERY(IES) IDENTIFIED?
☐	☐		PRIMARY SENSITIVE ENVIRONMENT(S) IDENTIFIED?

Summarize the rationale for Suspected Release (attach an additional page if necessary):

Summarize the rationale for Primary Targets (attach an additional page if necessary):

A-21

SURFACE WATER PATHWAY LIKELIHOOD OF RELEASE AND DRINKING WATER THREAT SCORESHEET

Pathway Characteristics

The surface water pathway includes three threats: Drinking Water Threat, Human Food Chain Threat, and Environmental Threat. Answer the questions at the top of the page. Refer to the Surface Water Pathway Criteria List (page 11) to hypothesize whether you suspect that a hazardous substance associated with the site has been released to surface water. Record the distance to surface water (the shortest overland drainage distance from a source to a surface water body). Record the flood frequency at the site (e.g., 100-yr, 200-yr). If the site is located in more than one floodplain, use the most frequent flooding event. Identify surface water use(s) along the surface water migration path and their distance(s) from the site.

Likelihood of Release (LR)

1. **Suspected Release:** Hypothesize based on professional judgment guided by the Surface Water Pathway Criteria List (page 11). If you suspect a release to surface water, use only Column A for this pathway and do not evaluate factor 2.

2. **No Suspected Release:** If you do not suspect a release, determine score based on the shortest overland drainage distance from a source to a surface water body. If distance to surface water is 2,500 feet or less, assign a score of 500. If distance to surface water is greater than 2,500 feet, determine score based on flood frequency. If you do not suspect a release to surface water, use only Column B to score this pathway.

Drinking Water Threat Targets (T)

3. List all drinking water intakes on downstream surface water bodies along the surface water migration path. Record the intake name, the type of water body on which the intake is located, the flow of the water body, and the number of people served by the intake (apportion the population if part of a blended system).

4. **Primary Target Population:** Evaluate populations served by all drinking water intakes that you suspect have been exposed to a hazardous substance released from the site. Use professional judgment guided by the Surface Water Pathway Criteria List (page 11) to make this determination. In the space provided, enter the population served by all intakes you suspect have been exposed to a hazardous substance from the site. If only the number of residences is known, use the average county residents per household (rounded up to the next integer) to determine population served. Multiply by 10 to determine the Primary Target Population score. Remember, if you do not suspect a release, there can be no primary target population.

5. **Secondary Target Population:** Evaluate populations served by all drinking water intakes within the target distance limit that you do not suspect have been exposed to a hazardous substance. Use PA Table 3 (page 13) and enter the population served by intakes for each flow category. If only the number of residences is known, use the average county residents per household (rounded to the nearest integer) to determine population served. Circle the assigned value for the population in each flow category and enter it in the column on the far-right side of the table. Sum the far-right column and enter the total as the Secondary Target Population factor score.

Gauging station data for many surface water bodies are available from USGS or other sources. In the absence of gauging station data, estimate flow using the list of surface water body types and associated flow categories in PA Table 4 (page 13). The flow for lakes is determined by the sum of flows of streams entering or leaving the lake. Note that the flow category "mixing zone of quiet flowing rivers" is limited to 3 miles from the probable point of entry.

6. **Nearest Intake** represents the threat posed to the drinking water intake that is most likely to be exposed to a hazardous substance. If you have identified a primary target population, enter 50. Otherwise, assign the score from PA Table 3 (page 13) for the lowest-flowing water body on which there is an intake.

7. **Resources:** A score of 5 can generally be assigned as a default measure. Assign zero only if surface water within the target distance limit has no resource use.

Sum the target scores in Column A (Suspected Release) or Column B (No Suspected Release).

SURFACE WATER PATHWAY
LIKELIHOOD OF RELEASE AND DRINKING WATER THREAT SCORESHEET

Pathway Characteristics	
Do you suspect a release (see Surface Water Pathway Criteria List, page 111?	Yes ___ No ___
Distance to surface water:	___ ft
Flood frequency:	___ yrs
What is the downstream distance to the nearest drinking water intake?	___ miles
Nearest fishery? ___ miles	Nearest sensitive environment? ___ miles

LIKELIHOOD OF RELEASE

1. **SUSPECTED RELEASE:** If you suspect a release to surface water (see page 11), assign a score of 550. Use only column A for this pathway.

2. **NO SUSPECTED RELEASE:** If you do not suspect a release to surface water, use the table below to assign a score based on distance to surface water and flood frequency. Use only column B for this pathway.

Distance to surface water ≤ 2,500 feet	500
Distance to surface water > 2,500 feet, and	
Site in annual or 10-year floodplain	500
Site in 100-year floodplain	400
Site in 500-year floodplain	300
Site outside 500-year floodplain	100

	A (Suspected Release)	B (No Suspected Release)	References
LR =			

DRINKING WATER THREAT TARGETS

3. Record the water body type, flow (if applicable), and number of people served by each drinking water intake within the target distance limit. If there is no drinking water intake within the target distance limit, factors 4, 5, and 6 each receive zero scores.

Intake Name	Water Body Type	Flow	People Served
		___ cfs	
		___ cfs	
		___ cfs	

4. **PRIMARY TARGET POPULATION:** If you suspect any drinking water intake listed above has been exposed to a hazardous substance from the site (see Surface Water Pathway Criteria List, page 11), list the intake name(s) and calculate the factor score based on the total population served.

___, ___

___ people x 10 =

5. **SECONDARY TARGET POPULATION:** Determine the number of people served by drinking water intakes that you do NOT suspect have been exposed to a hazardous substance from the site, and assign the total population score from PA Table 3.

Are any intakes part of a blended system? Yes ___ No ___
If yes, attach a page to show apportionment calculations.

6. **NEAREST INTAKE:** If you have identified a primary target population for the drinking water threat (factor 4), assign a score of 50; otherwise, assign the Nearest Intake score from PA Table 3. If no drinking water intake exists within the target distance limit, assign a score of zero.

7. **RESOURCES**

T =		

PA TABLE 3: VALUES FOR SECONDARY SURFACE WATER TARGET POPULATIONS

Surface Water Body Flow (see PA Table 4)	Population	Nearest Intake (choose highest)	Population Served by Intakes Within Flow Category: 1 to 30	31 to 100	101 to 300	301 to 1,000	1,001 to 3,000	3,001 to 10,000	10,001 to 30,000	30,001 to 100,000	100,001 to 300,000	300,001 to 1,000,000	Greater than 1,000,000	Population Value
< 10 cfs	_____	20	2	5	16	52	163	521	1,633	5,214	16,325	52,138	163,248	_____
10 to 10 cfs	_____	2	1	1	2	5	16	52	163	521	1,633	5,214	16,325	_____
> 100 to ,000 cfs	_____	1	0	0	1	1	2	5	16	52	163	521	1,633	_____
> 1,000 10,000 cfs	_____	0	0	0	0	0	1	1	2	5	16	52	163	_____
> 10,00 cfs or Great La s	_____	0	0	0	0	0	0	0	1	1	2	5	16	_____
3-mile Mi ng Zone	_____	10	1	3	8	26	82	261	816	2,607	8,162	26,068	81,863	_____
Nearest Intake =													Score =	

PA TABLE 4: SURFACE WATER TYPE / FLOW CHARACTERISTICS WITH DILUTION WEIGHTS FOR SECONDARY SURFACE WATER SENSITIVE ENVIRONMENTS

Type of Surface Water Body: Water Body Type	OR Flow	Dilution Weight
minimal stream	< 10 cfs	1
small to moderate stream	10 to 100 cfs	0.1
moderate to large stream	> 100 to 1,000 cfs	N/A
large stream to river	> 1,000 to 10,000 cfs	N/A
large river	> 10,000 cfs	N/A
3-mile mixing zone of quiet flowing streams or rivers	10 cfs or greater	N/A
coastal tidal water (harbors, sounds, bays, etc.), ocean, or Great Lakes	N/A	N/A

SURFACE WATER PATHWAY HUMAN FOOD CHAIN THREAT SCORESHEET

Likelihood of Release (LR)

LR is the same for all surface water pathway threats. Enter LR score from page 12.

Human Food Chain Threat Targets (T)

8. The only human food chain targets are fisheries. A fishery is an area of a surface water body from which food chain organisms are taken or could be taken for human consumption on a subsistence, sporting, or commercial basis. Food chain organisms include fish, shellfish, crustaceans, amphibians, and amphibious reptiles. Fisheries are delineated by changes in surface water body type (i.e., streams and rivers, lakes, coastal tidal waters, and oceans/Great Lakes) and whenever the flow characteristics of a stream or river change.

In the space provided, identify all fisheries within the target distance limit. Indicate the surface water body type and flow for each fishery. Gauging station flow data are available for many surface water bodies from USGS or other sources. In the absence of gauging station data, estimate flow using the list of surface water body types and associated flow categories in PA Table 4 (page 13). The flow for lakes is determined by the sum of flows of streams entering or leaving the lake. Note that, if there are no fisheries within the target distance limit, the Human Food Chain Threat Targets score is zero.

9. Primary fisheries are any fisheries within the target distance limit that you suspect have been exposed to a hazardous substance released from the site. Use professional judgment guided by the Surface Water Pathway Criteria List (page 11) to make this determination. If you identify any primary fisheries, list them in the space provided, enter 300 as the Primary Fisheries factor score, and do not evaluate Secondary Fisheries. Note that if you do not suspect a release, there can be no primary fisheries.

10. Secondary fisheries are fisheries that you do not suspect have been exposed to a hazardous substance. Evaluate this factor only if fisheries are present within the target distance limit, but none is considered a primary fishery.

A. If you suspect a release to surface water and have identified a secondary fishery but no primary fishery, assign a score of 210.

B. If you do not suspect a release, evaluate this factor based on flow. In the absence of gauging station flow data, estimate flow using the list of surface water body types and associated flow categories in PA Table 4 (page 13). Assign a Secondary Fisheries score from the table on the scoresheet using the lowest flow at any fishery within the target distance limit. (Dilution weight multiplier does not apply to PA evaluation of this factor.)

Sum the target scores in Column A (Suspected Release) or Column B (No Suspected Release).

SURFACE WATER PATHWAY (continued)
HUMAN FOOD CHAIN THREAT SCORESHEET

		A	B	
LIKELIHOOD OF RELEASE		*Suspected Release*	*No Suspected Release*	*Reference*
Enter Surface Water Likelihood of Release score from page 12.	LR =			

HUMAN FOOD CHAIN THREAT TARGETS

8. Record the water body type and flow (if applicable) for each fishery within the target distance limit. If there is no fishery within the target distance limit, assign a Targets score of 0 at the bottom of the page.

Fishery Name	*Water Body Type*	*Flow*
		cfs
		cfs
		cfs
		cfs
		cfs

9. PRIMARY FISHERIES: If you suspect any fishery listed above has been exposed to a hazardous substance from the site (see Surface Water Criteria List, page 11), assign a score of 300 and do not evaluate Factor 10. List the primary fisheries:

_____, _____.

_____, _____.

10. SECONDARY FISHERIES

A. If you suspect a release to surface water and have identified a secondary fishery but no primary fishery, assign a score of 210.

B. If you do not suspect a release, assign a Secondary Fisheries score from the table below using the lowest flow at any fishery within the target distance limit.

Lowest Flow	*Secondary Fisheries Score*
< 10 cfs	210
10 to 100 cfs	30
> 100 cfs, coastal tidal waters, oceans, or Great Lakes	12

T =

SURFACE WATER PATHWAY ENVIRONMENTAL THREAT SCORESHEET

Likelihood of Release (LR)

LR is the same for all surface water pathway threats. Enter LR score from page 12.

Environmental Threat Targets (T)

11. PA Table 5 (page 16) lists sensitive environments for the Surface Water Pathway Environmental Threat. In the space provided, identify all sensitive environments located within the target distance limit. Indicate the surface water body type and flow at each sensitive environment. Gauging station flow data for many surface water bodies are available from USGS or other sources. In the absence of gauging station data, estimate flow using the list of surface water body types and associated flow categories in PA Table 4 (page 13). The flow for lakes is determined by the sum of flows of streams entering or leaving the lake. Note that if there are no sensitive environments within the target distance limit, the Environmental Threat Targets score is zero.

12. Primary sensitive environments are surface water sensitive environments within the target distance limit that you suspect have been exposed to a hazardous substance released from the site. Use professional judgment guided by the Surface Water Pathway Criteria List (page 11) to make this determination. If you identify any primary sensitive environments, list them in the space provided, enter 300 as the Primary Sensitive Environments factor score, and do not evaluate Secondary Sensitive Environments. Note that if you do not suspect a release, there can be no primary sensitive environments.

13. Secondary sensitive environments are surface water sensitive environments that you do not suspect have been exposed to a hazardous substance. Evaluate this factor only if surface water sensitive environments are present within the target distance limit, but none is considered a primary sensitive environment. Evaluate secondary sensitive environments based on flow.

- In the table provided, list all secondary sensitive environments on surface water bodies with flow of 100 cfs or less.

  1) Use PA Table 4 (page 13) to determine the appropriate dilution weight for each.

  2) Use PA Tables 5 and 6 (page 16) to determine the appropriate value for each sensitive environment type and for wetlands frontage.

  3) For a sensitive environment that falls into more than one of the categories in PA Table 5, sum the values for each type to determine the environment value (e.g., a wetland with 1.5 miles frontage (value of 50) that is also a critical habitat for a Federally designated endangered species (value of 100) would receive a total value of 150).

  4) For each sensitive environment, multiply the dilution weight by the environment type (or length of wetlands) value and record the product in the far-right column.

  5) Sum the values in the far-right column and enter the total as the Secondary Sensitive Environments score. Do not evaluate part B of this factor.

- If all secondary sensitive environments are on surface water bodies with flows greater than 100 cfs, assign 10 as the Secondary Sensitive Environments score.

Sum the target scores in Column A (Suspected Release) or Column B (No Suspected Release).

SURFACE WATER PATHWAY (continued)
ENVIRONMENTAL THREAT SCORESHEET

LIKELIHOOD OF RELEASE		A *Suspected Release*	B *No Suspected Release*	References
Enter Surface Water Likelihood of Release score from page 12.	LR =			

ENVIRONMENTAL THREAT TARGETS

11. Record the water body type and flow (if applicable) for each surface water sensitive environment within the target distance limit (see PA Tables 4 and 5). If there is no sensitive environment within the target distance limit, assign a Targets score of 0 at the bottom of the page.

Environment Name	*Water Body Type*	*Flow*
		cfs
		cfs
		cfs
		cfs
		cfs

12. PRIMARY SENSITIVE ENVIRONMENTS: If you suspect any sensitive environment listed above has been exposed to a hazardous substance from the site (see Surface Water Criteria List, page 11), assign a score of 300 and do not evaluate factor 13. List the primary sensitive environments:

_____, _____.

_____, _____.

13. SECONDARY SENSITIVE ENVIRONMENTS: If sensitive environments are present, but none is a primary sensitive environment, evaluate Secondary Sensitive Environments based on flow.

A. For secondary sensitive environments on surface water bodies with flows of 100 cfs or less, assign scores as follows, and do not evaluate part B of this factor:

Flow	*Dilution Weight (PA Table 4)*	*Environment Type and Value (PA Tables 5 and 6)*	*Total*
cfs	x	=	
cfs	x	=	
cfs	x	=	
cfs	x	=	
cfs	x	=	

Subtotal =

B. If all secondary sensitive environments are located on surface water bodies with flows > 100 cfs, assign a score of 10.

T =

PA TABLE 5: SURFACE WATER AND AIR PATHWAY SENSITIVE ENVIRONMENTS VALUES

Sensitive Environment	*Assigned Value*
Critical habitat for Federally designated endangered or threatened species Marine Sanctuary National Park Designated Federal Wilderness Area Ecologically important areas identified under the Coastal Zone Wilderness Act Sensitive Areas identified under the National Estuary Program or Near Coastal Water Program of the Clean Water Act Critical Areas identified under the Clean Lakes Program of the Clean Water Act (subareas in lakes or entire small lakes) National Monument (air pathway only) National Seashore Recreation Area National Lakeshore Recreation Area	100
Habitat known to be used by Federally designated or proposed endangered or threatened species National Preserve National or State Wildlife Refuge Unit of Coastal Barrier Resources System Federal land designated for the protection of natural ecosystems Administratively Proposed Federal Wilderness Area Spawning areas critical for the maintenance of fish/shellfish species within a river system, bay, or estuary Migratory pathways and feeding areas critical for the maintenance of anadromous fish species in a river system Terrestrial areas utilized for breeding by large or dense aggregations of vertebrate animals (air pathway) or semi-aquatic foragers (surface water pathway) National river reach designated as Recreational	75
Habitat known to be used by State designated endangered or threatened species Habitat known to be used by a species under review as to its Federal endangered or threatened status Coastal Barrier (partially developed) Federally designated Scenic or Wild River	50
State land designated for wildlife or game management State designated Scenic or Wild River State designated Natural Area Particular areas, relatively small in size, important to maintenance of unique biotic communities	25
State designated areas for protection/maintenance of aquatic life under the Clean Water Act	5
Wetlands	See PA Table 6 (Surface Water Pathway) or PA Table 9 (Air Pathway)

PA TABLE 6: SURFACE WATER PATHWAY WETLANDS FRONTAGE VALUES

Total Length of Wetlands	*Assigned Value*
Less than 0.1 mile	0
0.1 to 1 mile	25
Greater than 1 to 2 miles	50
Greater than 2 to 3 miles	75
Greater than 3 to 4 miles	100
Greater than 4 to 8 miles	150
Greater than 8 to 12 miles	250
Greater than 12 to 16 miles	350
Greater than 16 to 20 miles	450
Greater than 20 miles	500

SURFACE WATER PATHWAY WASTE CHARACTERISTICS, THREAT, AND PATHWAY SCORES

Waste Characteristics (WC)

14. **Waste Characteristics**: Score is assigned from page 4. However, if a primary target has been identified for any surface water threat, assign either the score calculated on page 4 or a score of 32, whichever is greater.

Surface Water Pathway Threat Scores

Fill in the matrix with the appropriate scores from the previous pages. To calculate the score for each threat: multiply the scores for LR, T, and WC; divide the product by 82,500; and round the result to the nearest integer. The Drinking Water Threat and Human Food Chain Threat are each subject to a maximum of 100. The Environmental Threat is subject to a maximum of 60. Enter the rounded threat scores in the far-right column.

Surface Water Pathway Score

Sum the individual threat scores to determine the Surface Water Pathway Score. If the sum is greater than 100, assign 100.

SURFACE WATER PATHWAY (concluded)
WASTE CHARACTERISTICS, THREAT, AND PATHWAY SCORE SUMMARY

WASTE CHARACTERISTICS	A *Suspected Release*	B *No Suspected Release*
14. A. If you have identified any primary target for surface water (pages 12, 14, or 15), assign the waste characteristics score calculated on page 4, or a score of 32, whichever is GREATER; do not evaluate part B of this factor.		
B. If you have NOT identified any primary target for surface water, assign the waste characteristics score calculated on page 4.		
WC =		

SURFACE WATER PATHWAY THREAT SCORES

Threat	*Likelihood of Release (LR) Score from page 12)*	*Targets (T) Score (pages 12, 14, 15)*	*Pathway Waste Characteristics (WC) Score (determined above)*	*Threat Score LR x T x WC / 82,500*
Drinking Water				
Human Food Chain				
Environmental				

SURFACE WATER PATHWAY SCORE
(Drinking Water Threat + Human Food Chain Threat + Environmental Threat)

SOIL EXPOSURE PATHWAY CRITERIA LIST

Areas of surficial contamination can generally be assumed. This "Criteria List" helps guide the process of developing a hypothesis concerning the exposure of specific targets to a hazardous substance at the site. Use the "Resident Population" section to evaluate site and source conditions that may help identify targets likely to be exposed to a hazardous substance. The check-boxes record your professional judgment. Answers to all of the listed questions may not be available during the PA. Also, the list is not all-inclusive; if other criteria help shape your hypothesis, list them at the bottom of the page or attach an additional page.

Check the boxes to indicate a "yes," "no," or "unknown" answer to each question.

SOIL EXPOSURE PATHWAY CRITERIA LIST

SUSPECTED CONTAMINATION	*RESIDENT POPULATION*
Surficial contamination can generally be assumed.	Yes ☐ No ☐ Unk ☐ Is any residence, school, or daycare facility on or within 200 feet of an area of suspected contamination? ☐ ☐ ☐ Is any residence, school, or daycare facility located on adjacent land previously owned or leased by the site owner/operator? ☐ ☐ ☐ Is there a migration route that might spread hazardous substances near residences, schools, or daycare facilities? ☐ ☐ ☐ Have onsite or adjacent residents or students reported adverse health effects, exclusive of apparent drinking water or air contamination problems? ☐ ☐ ☐ Does any neighboring property warrant sampling? ☐ ☐ Other criteria?_____ ☐ ☐ RESIDENT POPULATION IDENTIFIED?

Summarize the rationale for Resident Population (attach an additional page if necessary):

SOIL EXPOSURE PATHWAY SCORESHEET

Pathway Characteristics

Answer the questions at the top of the page. Identify people who may be exposed to a hazardous substance because they work at the facility, or reside or attend school or daycare on or within 200 feet of an area of suspected contamination. If the site is active, estimate the number of full and part-time workers. Note that evaluation of targets is based on current site conditions.

Likelihood of Exposure (LE)

1. Suspected Contamination: Areas of surficial contamination are present at most sites, and a score of 550 can generally be assigned as a default measure. Assign zero, which effectively eliminates the pathway from further consideration, only if there is no surficial contamination; reliable analytical data are generally necessary to make this determination.

Resident Population Threat Targets (T)

2. Resident Population corresponds to "primary targets" for the migration pathways. Use professional judgment guided by the Soil Exposure Pathway Criteria List (page 18) to determine if there are people living or attending school or daycare on or within 200 feet of areas of suspected contamination. Record the number of people identified as resident population and multiply by 10 to determine the Resident Population factor score.

3. Resident Individual: Assign 50 if you have identified a resident population; otherwise, assign zero.

4. Workers: Estimate the number of full and part-time workers at this facility and adjacent facilities where contamination is also suspected. Assign a score for the Workers factor from the table.

5. Terrestrial Sensitive Environments: In the table provided, list each terrestrial sensitive environment located on an area of suspected contamination. Use PA Table 7 (page 20) to assign a value for each. Sum the values and assign the total as the factor score.

6. Resources: A score of 5 can generally be assigned as a default measure. Assign zero only if there is no land resource use on an area of suspected contamination.

Sum the target scores.

Waste Characteristics (WC)

7. Enter the WC score determined on page 4.

Resident Population Threat Score: Multiply the scores for LE, T, and WC. Divide the product by 82,500. Round the result to the nearest integer. If the result is greater than 100, assign 100.

Nearby Population Threat Score: Do not evaluate this threat if you gave a zero score to Likelihood of Exposure. Otherwise, assign a score based on the population within a 1-mile radius (use the same 1-mile radius population you evaluate for air pathway population targets):

Population Within One Mile	Nearby Population Threat Score
< 10,000	1
10,000 to 50,000	2
>50,000	4

Soil Exposure Pathway Score: Sum the Resident Population Threat score and the Nearby Population Threat score, subject to a maximum of 100.

SOIL EXPOSURE PATHWAY SCORESHEET

Pathway Characteristics	
Do any people live on or within 200 ft of areas of suspected contamination?	Yes ___ No ___
Do any people attend school or daycare on or within 200 ft of areas of suspected contamination?	Yes ___ No ___
Is the facility active? Yes ___ No ___ If yes, estimate the number of workers: ___	

LIKELIHOOD OF EXPOSURE	Suspected Contamination	References
1. SUSPECTED CONTAMINATION: Surficial contamination can generally be assumed, and a score of 550 assigned. Assign zero only if the absence of surficial contamination can be confidently demonstrated. LE =		

RESIDENT POPULATION THREAT TARGETS

2. RESIDENT POPULATION: Determine the number of people occupying residences or attending school or daycare on or within 200 feet of areas of suspected contamination (see Soil Exposure Pathway Criteria List, page 18).
___ people x 10 =

3. RESIDENT INDIVIDUAL: If you have identified a resident population (factor 2), assign a score of 50; otherwise, assign a score of 0.

4. WORKERS: Use the following table to assign a score based on the total number of workers at the facility and nearby facilities with suspected contamination:

Number of Workers	Score
0	0
1 to 100	5
101 to 1,000	10
> 1,000	15

5. TERRESTRIAL SENSITIVE ENVIRONMENTS: Use PA Table 7 to assign a value for each terrestrial sensitive environment on an area of suspected contamination:

Terrestrial Sensitive Environment Type	Value
	Sum =

6. RESOURCES

T =

WASTE CHARACTERISTICS

7. Assign the waste characteristics score calculated on page 4. WC =

RESIDENT POPULATION THREAT SCORE: $\frac{LE \times T \times WC}{82,500.}$

NEARBY POPULATION THREAT SCORE:

SOIL EXPOSURE PATHWAY SCORE:
Resident Population Threat + Nearby Population Threat

PA TABLE 7: SOIL EXPOSURE PATHWAY
TERRESTRIAL SENSITIVE ENVIRONMENT VALUES

Terrestrial Sensitive Environment	*Assigned Value*
Terrestrial critical habitat for Federally designated endangered or threatened species National Park Designated Federal Wilderness Area National Monument	100
Terrestrial habitat known to be used by Federally designated or proposed threatened or endangered species National Preserve (terrestrial) National or State terrestrial Wildlife Refuge Federal land designated for protection of natural ecosystems Administratively proposed Federal Wilderness Area Terrestrial areas utilized by large or dense aggregations of animals (vertebrate species) for breeding	75
Terrestrial habitat used by State designated endangered or threatened species Terrestrial habitat used by species under review for Federal designated endangered or threatened status	50
State lands designated for wildlife or game management State designated Natural Areas Particular areas, relatively small in size, important to maintenance of unique biotic communities	25

AIR PATHWAY CRITERIA LIST

This "Criteria List" helps guide the process of developing a hypothesis as to whether a release to the air is likely to be detected. The check-boxes record your professional judgment. Answers to all of the listed questions may not be available during the PA. Also, the list is not all-inclusive; if other criteria help shape your hypothesis, list them at the bottom of the page or attach an additional page.

The "Suspected Release" section identifies several conditions that could provide insight as to whether a release from the site is likely to be detected. If a release is suspected, primary targets are any residents, workers, students, and sensitive environments on or within ¼ mile of the site.

Check the boxes to indicate a "yes," "no," or "unknown" answer to each question. If you check the "Suspected Release" box as "yes," make sure you assign a Likelihood of Release value of 550 for the pathway.

AIR PATHWAY CRITERIA LIST

SUSPECTED RELEASE				*PRIMARY TARGETS*
Yes	No	Unk		
☐	☐	☐	Are odors currently reported?	If you suspect a release to air, evaluate all populations and sensitive environments within 1/4 mile (including those onsite) as primary targets.
☐	☐	☐	Has release of a hazardous substance to the air been directly observed?	
☐	☐	☐	Are there reports of adverse health effects (e.g., headaches, nausea, dizziness) potentially resulting from migration of hazardous substances through the air?	
☐	☐	☐	Does analytical or circumstantial evidence suggest a release to the air?	
☐	☐		Other criteria?_____	
☐	☐		**SUSPECTED RELEASE?**	

Summarize the rationale for Suspected Release (attach an additional page if necessary):

AIR PATHWAY SCORESHEET

Pathway Characteristics

Answer the questions at the top of the page. Refer to the Air Pathway Criteria List (page 21) to hypothesize whether you suspect that a hazardous substance release to the air could be detected. Due to dispersion, releases to air are not as persistent as releases to water migration pathways and are much more difficult to detect. Develop your hypothesis concerning the release of hazardous substances to air based on "real time" considerations. Record the distance (in feet) from any source to the nearest regularly occupied building.

Likelihood of Release (LR)

1. Suspected Release: Hypothesize based on professional judgment guided by the Air Pathway Criteria List (page 21). If you suspect a release to air, use only Column A for this pathway and do not evaluate factor 2.

2. No Suspected Release: If you do not suspect a release, enter 500 and use only Column B for this pathway.

Targets (T)

3. Primary Target Population: Evaluate populations subject to exposure from release of a hazardous substance from the site. If you suspect a release, the resident, student, and worker populations on and within ¼ mile of the site are considered primary target population. If only the number of residences is known, use the average county residents per household (rounded up to the next integer) to determine the population. In the space provided, enter this population. Multiply the population by 10 to determine the Primary Target Population score. Note that if you do not suspect a release, there can be no primary target population.

4. Secondary Target Population: Evaluate populations in distance categories not suspected to be subject to exposure from release of a hazardous substance from the site. If you suspect a release, residents, students, and workers in the ¼- to 4-mile distance categories are secondary target population. If you do not suspect a release, all residents, students, and workers onsite and within 4 miles are considered secondary target population.

Use PA Table 8 (page 23). Enter the population in each secondary target population distance category, circle the assigned value, and record it on the far-right side of the table. Sum the far-right column and enter the total as the Secondary Target Population factor score.

5. Nearest Individual represents the threat posed to the person most likely to be exposed to a hazardous substance release from the site. If you have identified a primary target population, enter 50. Otherwise, assign the score from PA Table 8 (page 23) for the closest distance category in which you have identified a secondary target population.

6. Primary Sensitive Environments: If a release is suspected, all sensitive environments on or within ¼ mile of the site are considered primary targets. List them and assign values for sensitive environment type (from PA Table 5, page 16) and/or wetland acreage (from PA Table 9, page 23). Sum the values and enter the total as the factor score.

7. Secondary Sensitive Environments: If a release is suspected, sensitive environments in the ¼- to ½-mile distance category are secondary targets; greater distances need not be evaluated because distance weighting greatly diminishes the impact on site score. If you do not suspect a release, all sensitive environments on and within ½ mile of the site are considered secondary targets. List each secondary sensitive environment on PA Table 10 (page 23) and assign a value to each using PA Tables 5 and 9. Multiply each value by the indicated distance weight and record the product in the far-right column. Sum the products and enter the total as the factor score.

8. Resources: A score of 5 can generally be assigned as a default measure. Assign zero only if there is no land resource use within ½ mile.

Sum the target scores in Column A (Suspected Release) or Column B (No Suspected Release).

Waste Characteristics (WC)

9. Waste Characteristics: Score is assigned from page 4. However, if you have identified any primary target for the air pathway, assign either the score calculated on page 4 or a score of 32, whichever is greater.

Air Pathway Score: Multiply the scores for LR, T, and WC. Divide the product by 82,500. Round the result to the nearest integer. If the result is greater than 100, assign 100.

A-42

AIR PATHWAY SCORESHEET

Pathway Characteristics	
Do you suspect a release (see Air Pathway Criteria List, page 211)?	Yes ___ No ___
Distance to the nearest individual:	_____ ft

	A (Suspected Release)	B (No Suspected Release)	Reference
LIKELIHOOD OF RELEASE			
1. SUSPECTED RELEASE: If you suspect a release to air (see page 211), assign a score of 550. Use only column A for this pathway.			
2. NO SUSPECTED RELEASE: If you do not suspect a release to air, assign a score of 500. Use only column B for this pathway.			
LR =			
TARGETS			
3. PRIMARY TARGET POPULATION: Determine the number of people subject to exposure from a suspected release of hazardous substances to the air. _____ people x 10 =			
4. SECONDARY TARGET POPULATION: Determine the number of people not suspected to be exposed to a release to air, and assign the total population score using PA Table 6.			
5. NEAREST INDIVIDUAL: If you have identified any Primary Target Population for the air pathway, assign a score of 80; otherwise, assign the Nearest Individual score from PA Table 6.			
6. PRIMARY SENSITIVE ENVIRONMENTS: Sum the sensitive environment values (PA Table 5) and wetland acreage values (PA Table 9) for environments subject to exposure from a suspected release to the air.			
Sensitive Environment Type / Value: _____ / _____ ; _____ / _____			
Sum =			
7. SECONDARY SENSITIVE ENVIRONMENTS: Use PA Table 10 to determine the score for secondary sensitive environments.			
8. RESOURCES			
T =			
WASTE CHARACTERISTICS			
9. A. If you have identified any Primary Target for the air pathway, assign the waste characteristics score calculated on page 4, or a score of 32, whichever is GREATER; do not evaluate part B of this factor.			
B. If you have NOT identified any Primary Target for the air pathway, assign the waste characteristics score calculated on page 4.			
WC =			

AIR PATHWAY SCORE: $\dfrac{LR \times T \times WC}{82{,}500}$ =

PA TABLE 8: VALUES FOR SECONDARY AIR TARGET POPULATIONS

Distance from Site	Population	Nearest Individual (choose highest)	Population Within Distance Category											Population Value
			1 to 10	11 to 30	31 to 100	101 to 300	301 to 1,000	1,001 to 3,000	3,001 to 10,000	10,001 to 30,000	30,001 to 100,000	100,001 to 300,000	300,001 to 1,000,000	Greater than 1,000,000	
Onsite	_____	20	1	2	8	16	52	163	521	1,633	5,214	16,325	52,136	163,246	_____
>0 to ¼ mile	_____	20	1	1	1	4	12	41	130	408	1,303	4,081	13,034	40,811	_____
> ¼ to ½ mile	_____	2	0	0	1	1	3	9	28	88	282	882	2,815	8,815	_____
> ½ to 1 mile	_____	1	0	0	0	1	1	3	8	28	83	261	834	2,612	_____
> 1 to 2 miles	_____	0	0	0	0	0	1	1	3	8	27	83	266	833	_____
> 2 to 3 miles	_____	0	0	0	0	0	1	1	1	4	12	38	120	376	_____
> 3 to 4 miles	_____	0	0	0	0	0	0	1	1	2	7	23	73	229	_____
Nearest Individual =														Score =	

PA TABLE 9: AIR PATHWAY VALUES FOR WETLAND AREA

Wetland Area	Assigned Value
Less than 1 acre	0
1 to 50 acres	25
Greater than 50 to 100 acres	75
Greater than 100 to 150 acres	125
Greater than 150 to 200 acres	175
Greater than 200 to 300 acres	250
Greater than 300 to 400 acres	350
Greater than 400 to 500 acres	450
Greater than 500 acres	500

PA TABLE 10: DISTANCE WEIGHTS AND CALCULATIONS FOR AIR PATHWAY SECONDARY SENSITIVE ENVIRONMENTS

Distance	Distance Weight	Sensitive Environment Type and Value (from PA Table 6 or 9)	Product
Onsite	0.10	a. _____	
		b. _____	
0-1/4 mi	0.025	a. _____	
		b. _____	
		c. _____	
1/4-1/2 mi	0.0054	a. _____	
		b. _____	
		c. _____	
		d. _____	
Total Environments Score =			

SITE SCORE CALCULATION

In the column labeled S, record the Ground Water Pathway score, the Surface Water Pathway score, the Soil Exposure Pathway score, and the Air Pathway score. Square each pathway score and record the result in the S^2 column. Sum the squared pathway scores. Divide the sum by 4, and take the square root of the result to obtain the Site Score.

SUMMARY

Answer the summary questions, which ask for a qualitative evaluation of the relative risk of targets being exposed to a hazardous substance from the site. You may find your responses to these questions a good cross-check against the way you scored the individual pathways. For example, if you scored the ground water pathway on the basis of no suspected release and secondary targets only, yet your response to question #1 is "yes," this presents apparently conflicting conclusions that you need to reconsider and resolve. Your answers to the questions on page 24 should be consistent with your evaluations elsewhere in the PA scoresheets package.

SITE SCORE CALCULATION

	S	S^2
GROUND WATER PATHWAY SCORE (S_{gw}):		
SURFACE WATER PATHWAY SCORE (S_{sw}):		
SOIL EXPOSURE PATHWAY SCORE (S_s):		
AIR PATHWAY SCORE (S_a):		
SITE SCORE: $\sqrt{\frac{S_{gw}^2+S_{sw}^2+S_s^2+S_a^2}{4}}$		

SUMMARY

	YES	NO
1. Is there a high possibility of a threat to any nearby drinking water well(s) by migration of a hazardous substance in ground water?	☐	☐
A. If yes, identify the well(s). _____		
B. If yes, how many people are served by the threatened well(s)? _____		
2. Is there a high possibility of a threat to any of the following by hazardous substance migration in surface water?		
A. Drinking water intake	☐	☐
B. Fishery	☐	☐
C. Sensitive environment (wetland, critical habitat, others)	☐	☐
D. If yes, identify the target(s). _____		
3. Is there a high possibility of an area of surficial contamination within 200 feet of any residence, school, or daycare facility?	☐	☐
If yes, identify the property(ies) and estimate the associated population(s). _____		
4. Are there public health concerns at this site that are not addressed by PA scoring considerations? If yes, explain: _____	☐	☐

Appendix B: USGS Data Files

All of the fields in the USGS data files were reformatted in May 1996. They are now in a 'spreadsheet' format, in which all fields are separated by a TAB character. This method should allow the data to be easily imported into a spreadsheet program on your computer.

The first line in each data file contains the 'headers', which is an 8 (or so) letter code describing each data element. For example, the code for 'Commercial withdrawals, ground-water, fresh' is 'co-wgwfr'. Each line (record) in the data files begins with some descriptive information, such as the state, year, and county or water-resources region code. There are headers for these items at the beginning of the first line (the header line). These first few header codes vary by file type.

NOTE: The initial header information (first 4 or 5 items in each line) varies by file. The actual code headers for the data elements are described in Appendix B.

The data elements with a "Entered by user" under the "HOW ELEMENT IS COMPUTED" heading indicate data that has been compiled and entered manually into the files. A data element that is not "Entered by user" was calculated by using other data values. For example, 'ps-popgw' (index #3) and 'ps-popsw' (index #4) are summed to create 'ps-popto'. So, total public supply population served is computed by summing the population served by ground water and the population served by surface water.

A table of data elements, element codes and method of computation is in Appendix B.

Source: http://h2o.usgs.gov/public/watuse/data/wudict.txt

Appendix C : Complete List of Water-Use Data Elements

WATER-USE DATA ELEMENT		CODE	INDEX	HOW ELEMENT IS COMPUTED
TOTAL POPULATION (in AREA, in thousands)		po-total	2	Entered by user
PUBLIC SUPPLY:				
Population served (thousands):	Ground Water	ps-popgw	3	Entered by user
" "	Surface Water	ps-popsw	4	Entered by user
" "	Total	ps-popto	5	3+4
Water withdrawals, fresh	Ground Water	ps-wgwfr	6	Entered by user
" " "	Surface Water	ps-wswfr	7	Entered by user
" " "	Total, Fresh	ps-wtofr	8	6+7
Water withdrawals, saline	Ground Water	ps-wgwsa	9	Entered by user
" " "	Surface Water	ps-wswsa	10	Entered by user
" " "	Total Saline	ps-wsato	11	9+10
Water withdrawals, total	Total Total	ps-total	12	6+7+9+10
Water deliveries:	Public Use and Losses	ps-loss	13	12-21-30-44-71-93
" "	Total Deliveries	ps-deliv	14	13+21+30+44+71+93
Per capita use:	Total, Gal/d	ps-prcap	15	12*1000/5
Number of facilities:	In Area	ps-facil	16	Entered by user
" " "	Water-Use Database	ps-facdb	17	Entered by user
COMMERCIAL:				
Self-supplied withdrawals:	Ground Water	co-wgwfr	18	Entered by user
" "	Surface Water	co-wswfr	19	Entered by user
" "	Total	co-wtotl	20	18+19
Deliveries from water supply:	Fresh	co-psdel	21	Entered by user
Total:	Withdr. + Deliveries	co-total	22	18+19+21
Consumptive use:		co-cuse	23	Entered by user
DOMESTIC:				
SELF SUPPLIED:				

Population:	Thousands	do-sspop	24	2-5
Water withdrawals:	Ground Water	do-ssgwf	25	Entered by user
"	Surface Water	do-ssswf	26	Entered by user
"	Total	do-sstot	27	25+26
Per capita use:	Gal/d	do-sspcp	28	(25+26)*1000/24
PUBLIC SUPPLIED:				
Population:	Thousands	do-pspop	29	5
Deliveries from water supply	Fresh	do-psdel	30	Entered by user
Per capita use:	Gal/d	do-pspcp	31	30*1000/29
TOTAL:				
Withdrawals + deliveries:		do-total	32	25+26+30
Consumptive use:		do-cuse	33	Entered by user
INDUSTRIAL:				
Self-supplied withdrawals:	Ground Water, fresh	in-wgwfr	34	Entered by user
" "	Ground Water, saline	in-wgwsa	35	Entered by user
" "	Ground Water, total	in-wgwto	36	34+35
" "	Surface Water, fresh	in-wswfr	37	Entered by user
" "	Surface Water, saline	in-wswsa	38	Entered by user
" "	Surface Water, total	in-wswto	39	37+38
" "	Total, Fresh	in-wtofr	40	34+37
" "	Total, Saline	in-wtosa	41	35+38
" "	Total, Total	in-wtotl	42	34+35+37+38
" "	Reclaimed Sewage	in-recww	43	Entered by user
Deliveries from water supply	Fresh	in-psdel	44	Entered by user
Total, withdrawal+deliveries	Fresh	in-total	45	34+37+44
Consumptive use	Fresh	in-cufr	46	Entered by user
" "	Saline	in-cusal	47	Entered by user
" "	Total	in-cuse	48	46+47
Number of facilities:	In Area	in-facil	49	Entered by user
" " "	In Water-Use Database	in-facdb	50	Entered by user

THERMOELECTRIC POWER

(ELECTRIC):

All thermoelectric water use:	Ground Water, fresh	pt-wgwfr	51	65+79+87
" " " "	Surface Water, fresh	pt-wswfr	52	66+88
" " " "	Surface Water, saline	pt-wswsa	53	67+89
" " " "	Surface Water, total	pt-wswto	54	66+67+88+89
" " " "	Total, Fresh	pt-frtot	55	65+66+79+87+88
" " " "	Total, Total	pt-wtotl	56	65+66+67+79+80+87+88+89
Deliveries from water supply	Fresh	pt-psdel	57	71+93
Total: Withdrawal + deliv	Fresh	pt-total	58	65+66+71+79+87+88+93
Consumptive use	Fresh	pt-cufr	59	73+82+95
" "	Saline	pt-cusal	60	74+83+96
" "	Total	pt-cuse	61	73+74+82+83+95+96
Total power generated:	Gigawatthours/year	pt-power	62	76+84+98
Number of facilities:	In Area	pt-facil	63	77+85+99
" " "	In Water-Use Database	pt-facdb	64	78+86+100

THERMOELECTRIC POWER

(ELECTRIC), FOSSIL FUEL

Fossil fuel: Withdrawals	Ground Water, fresh	pf-wgwfr	65	Entered by user
" " "	Surface Water, fresh	pf-wswfr	66	Entered by user
" " "	Surface Water, saline	pf-wswsa	67	Entered by user
" " "	Surface Water, total	pf-wswto	68	66+67
" " "	Total, Fresh	pf-frtot	69	65+66
" " "	Total, Total	pf-wtotl	70	65+66+67
Deliveries from water supply	Fresh	pf-psdel	71	Entered by user
Total: Withdrawal + deliver	Fresh	pf-total	72	65+66+71
Consumptive use	Fresh	pf-cufr	73	Entered by user
" "	Saline	pf-cusal	74	Entered by user
" "	Total	pf-cuse	75	73+74
Power generation	Gigawatthours/year	pf-power	76	Entered by user
Number of facilities:	In Area	pf-facil	77	Entered by user

" " "	In Water-Use Database	pf-facdb	78	Entered by user
THERMOELECTRIC POWER (ELECTRIC), GEOTHERMAL				
Geothermal: Withdrawals	Ground Water, fresh	pg-wgwfr	79	Entered by user
" "	Ground Water, saline	pg-wgwsa	80	Entered by user
" "	Total	pg-wtotl	81	79+80
Consumptive use	Fresh	pg-cufr	82	Entered by user
" "	Saline	pg-cusal	83	Entered by user
Power generation	Gigawatthours/year	pg-power	84	Entered by user
Number of facilities:	In Area	pg-facil	85	Entered by user
" " "	In Water-Use Database	pg-facdb	86	Entered by user
THERMOELECTRIC POWER (ELECTRIC), NUCLEAR:				
Nuclear: Withdrawals	Ground Water, fresh	pn-wgwfr	87	Entered by user
" "	Surface Water, fresh	pn-wswfr	88	Entered by user
" "	Surface Water, saline	pn-wswsa	89	Entered by user
" "	Surface Water, total	pn-wswto	90	88+89
" "	Total, Fresh	pn-frtot	91	87+88
" "	Total, Total	pn-wtotl	92	87+88+89
Deliveries from water supply	Fresh	pn-psdel	93	Entered by user
Total: Withdrawal + deliveries	Fresh	pn-total	94	87+88+93
Consumptive use	Fresh	pn-cufr	95	Entered by user
" "	Saline	pn-cusal	96	Entered by user
" "	Total	pn-cuse	97	95+96
Power generation	Gigawatthours/year	pn-power	98	Entered by user
Number of facilities:	In Area	pn-facil	99	Entered by user
" " "	In Water-Use Database	pn-facdb	100	Entered by user
MINING:				
Withdrawals:	Ground Water, fresh	mi-wgwfr	101	Entered by user

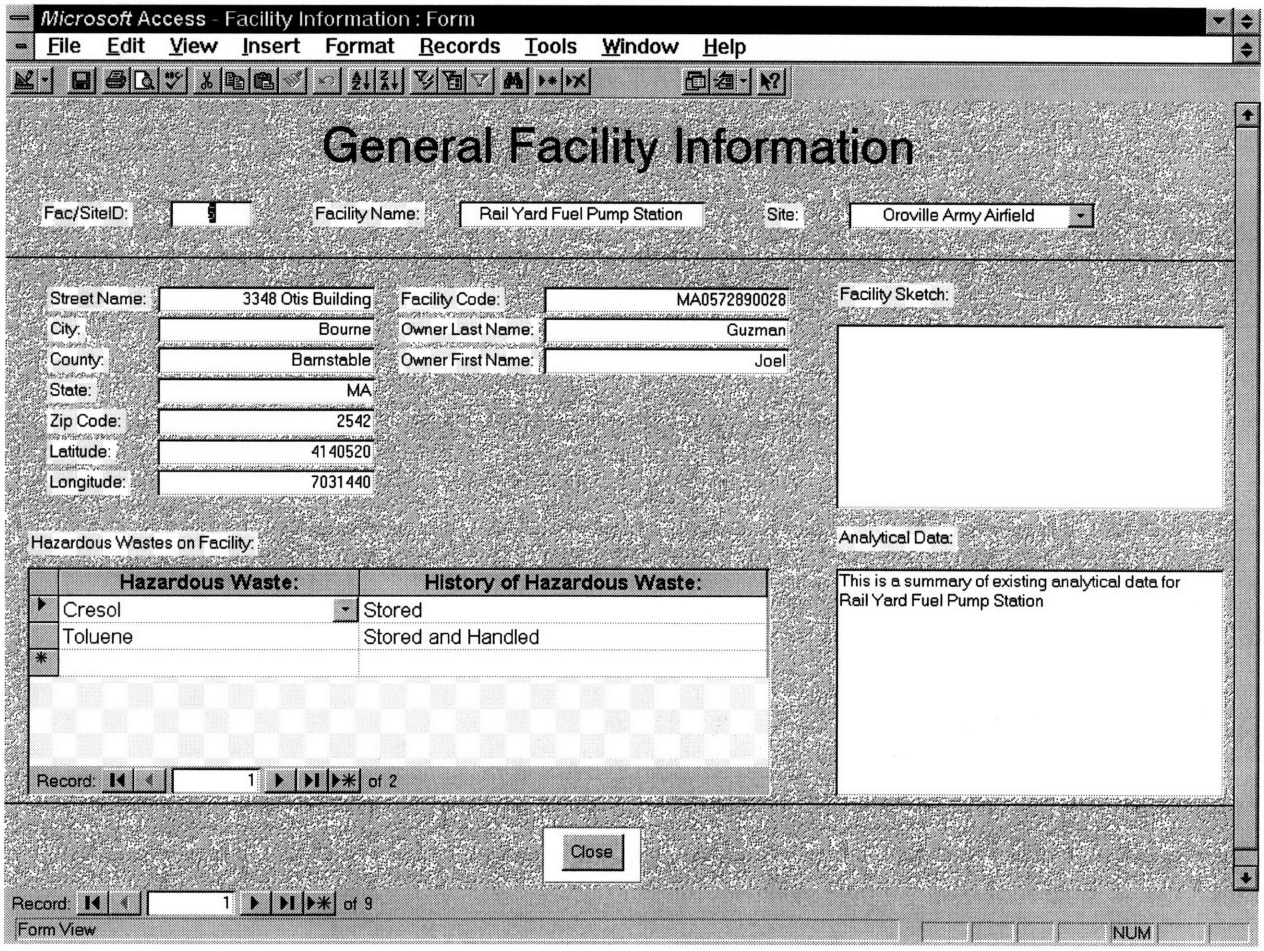
"	Ground Water, saline	mi-wgwsa	102	Entered by user
"	Ground Water, total	mi-gwtot	103	101+102
"	Surface Water, fresh	mi-wswfr	104	Entered by user
"	Surface Water, saline	mi-wswsa	105	Entered by user
"	Surface Water, total	mi-swtot	106	104+105
Withdrawals, total:	Total, Fresh	mi-frtot	107	101+104
" "	Total, Saline	mi-satot	108	102+105
" "	Total, Total	mi-total	109	101+102+104+105
Consumptive Use:	Fresh	mi-cufr	110	Entered by user
" "	Saline	mi-cusal	111	Entered by user
" "	Total	mi-cuse	112	110+111
LIVESTOCK:				
Stock: Withdrawals:	Ground Water	ls-gwtot	113	Entered by user
" "	Surface Water	ls-swtot	114	Entered by user
" "	Total	ls-total	115	113+114
Stock, consumptive use:		ls-cuse	116	Entered by user
Animal specialties, withdrawals:	Ground Water	la-gwtot	117	Entered by user
" " "	Surface Water	la-swtot	118	Entered by user
" " "	Total	la-total	119	117+118
Animal specialties, consumptive use:		la-cuse	120	Entered by user
Total livestock: Withdrawals:	Ground Water	lv-gwtot	121	113+117
" " "	Surface Water	lv-swtot	122	114+118
" " "	Total	lv-total	123	113+114+117+118
Total livestock, consumptive use:		lv-cuse	124	116+120
IRRIGATION				
Withdrawals, fresh	Ground Water	ir-wgwfr	125	Entered by user
" "	Surface Water	ir-wswfr	126	Entered by user
"	Reclaimed Sewage	ir-recww	127	Entered by user
" "	Total, Fresh	ir-frtot	128	125+126
Irrigated land, in 1000 acres:	Sprayed	ir-spray	129	Entered by user

" " " "	Flooded	ir-flood	130	Entered by user
" " " "	Total	ir-irrig	131	129+130
Conveyance losses:	Total	ir-convy	132	Entered by user
Consumptive use:		ir-cuse	133	Entered by user
HYDROELECTRIC POWER:				
Water use:	Total	hy-total	134	Entered by user
Power generation:	Gigawatthours/year	hy-power	135	Entered by user
Number of facilities:	In Area	hy-facil	136	Entered by user
" " "	In Water-Use Database	hy-facdb	137	Entered by user
SEWAGE TREATMENT:				
Number of facilities:	Public	ww-facpu	138	Entered by user
" " "	Industrial + Other	ww-facot	139	Entered by user
" " "	Total Number	ww-facil	140	138+139
Municipal system returns:		ww-retrn	141	Entered by user
Number of facilities:	In Water-Use Database	ww-facdb	142	Entered by user
Reclaimed waste water (WW) from pub. WW facilities		ww-recww	143	Entered by user
RESERVOIR EVAPORATION				
Amt evaporated, (1000 acre ft):	Fresh	re-evap	144	Entered by user
Surface area, in 1000 acres:		re-area	145	Entered by user
TOTAL WATER USE:				
	Ground water, fresh	to-gwfr	146	6+18+25+34+65+79+87+101+113+117+125
	Ground water, saline	to-gwsal	147	9+35+80+102
	Ground water, total	to-totgw	148	6+9+18+25+34+35+65+79+80+87+101+ 102+113+117+125
	Surface water, fresh	to-swfr	149	7+19+26+37+66+88+104+114+118+126
	Surface water, saline	to-swsal	150	10+38+67+89+105
	Surface water, total	to-totsw	151	7+10+19+26+37+38+66+67+88+89+104+

			105+114+118+126
Total, fresh	to-totfr	152	6+7+18+19+25+26+34+37+65+66+79+87+
Total, saline	to-totsa	153	9+10+35+38+67+80+89+102+105
Total, total	to-total	154	6+7+9+10+18+19+25+26+34+35+37+38+65+ 66+67+79+80+87+88+89+101+102+104+105+ 113+114+117+118+125+126
Reclaimed sewage	to-recww	155	43+127
Consumptive use, fresh	to-cufr	156	23+33+46+73+82+95+110+116+120+133
Consumptive use, saline	to-cusal	157	47+74+83+96+111
Consumptive use, total	to-cuse	158	23+33+46+47+73+74+82+83+95+96+110+ 111+116+120+133
Conveyance losses	to-convy	159	132

Source: http://h2o.usgs.gov/public/watuse/data/wudict.txt

Appendix D: Microsoft Access Preliminary Assessment Scoresheet Forms



Form Name: Facility Information

Description: This form contains all data recorded in the General Information Section of the PA scoresheet.

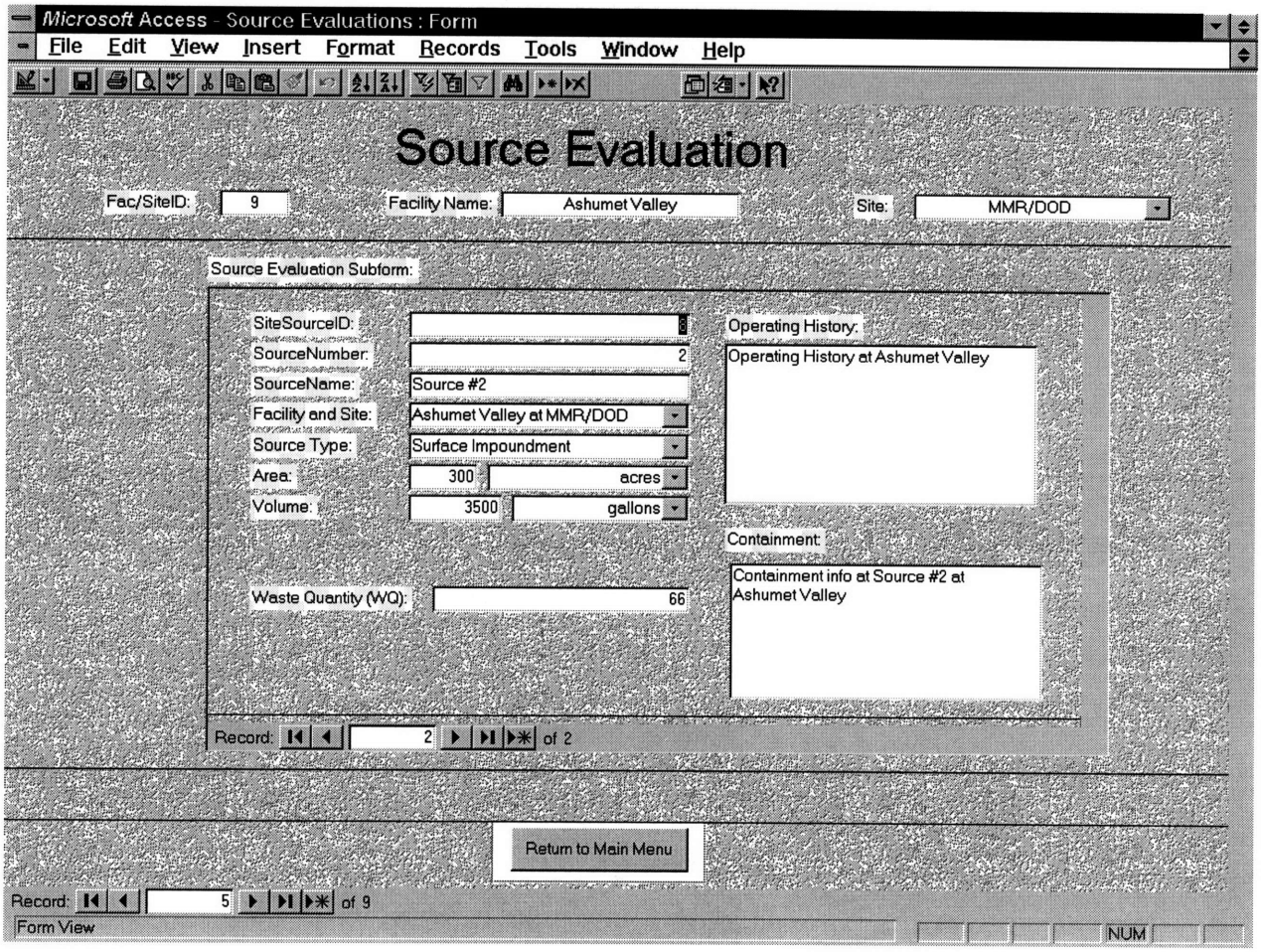
Subforms: Facility Haz Subform (data based on GeneralSiteInfoFacilityHazSub table. This tables lists the hazardous wastes at each facility).

Form Data: GeneralFacilityInformation table.

Events and Macros: The form is closed when the Close button is clicked.

Opened From: Review PA Sections form

Figure D1. General Facility Information Form



Form Name: Source Evaluations

Description: This forms contains the data recorded in the source Evaluation section of the PA scoresheet.

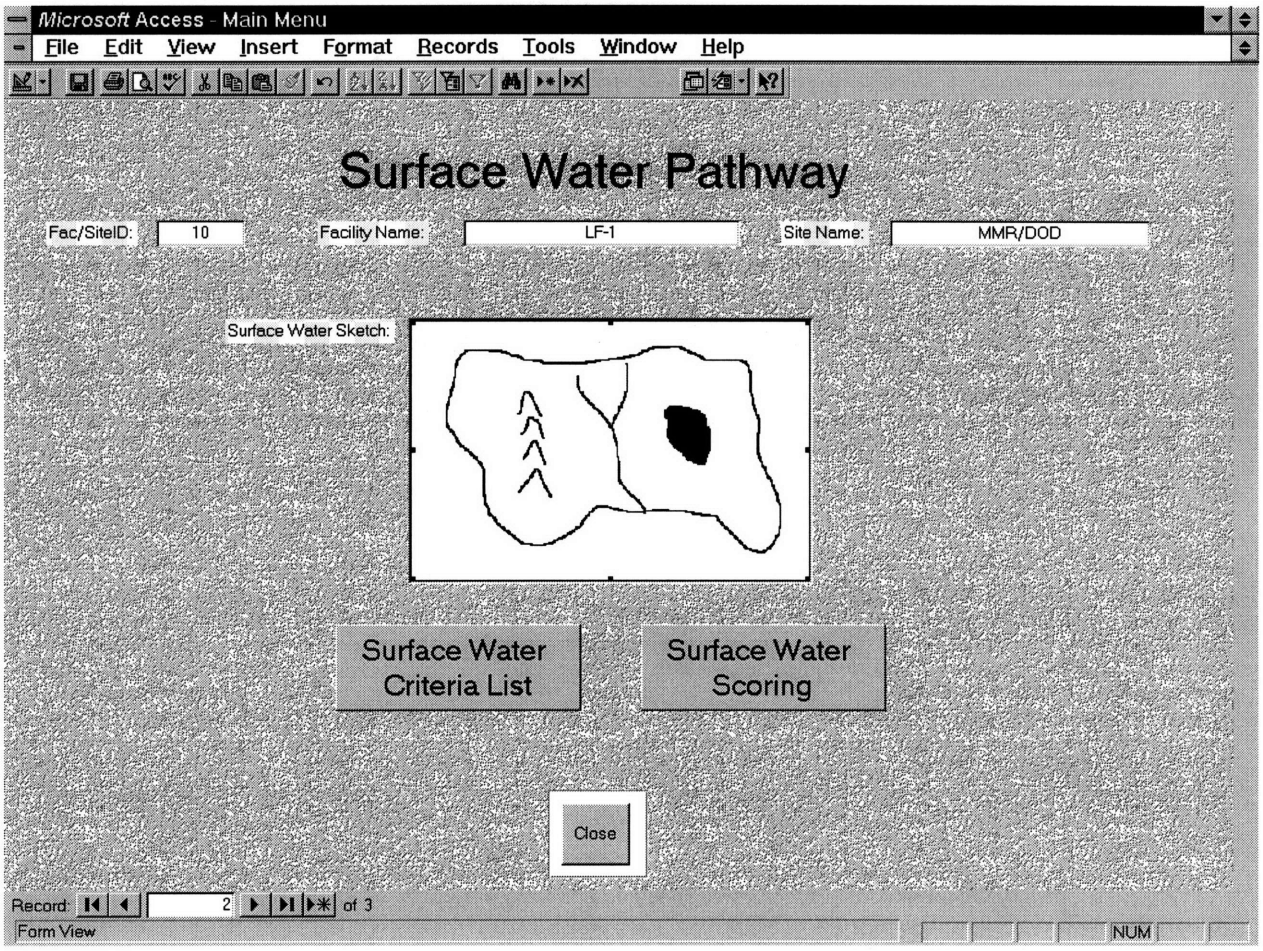
Subforms: Source Subform

Form Data: GeneralFacilityInformation table.

Events and Macros: The form is closed when the Close button is clicked.

Opened From: Review PA Sections form

Figure D2. Source Evaluation Form



Form Name: Surface Water Main

Description: This form contains just the surface water sketch for the facility and two buttons for opening the Surface Water Criteria List and Surface Water Scoring.

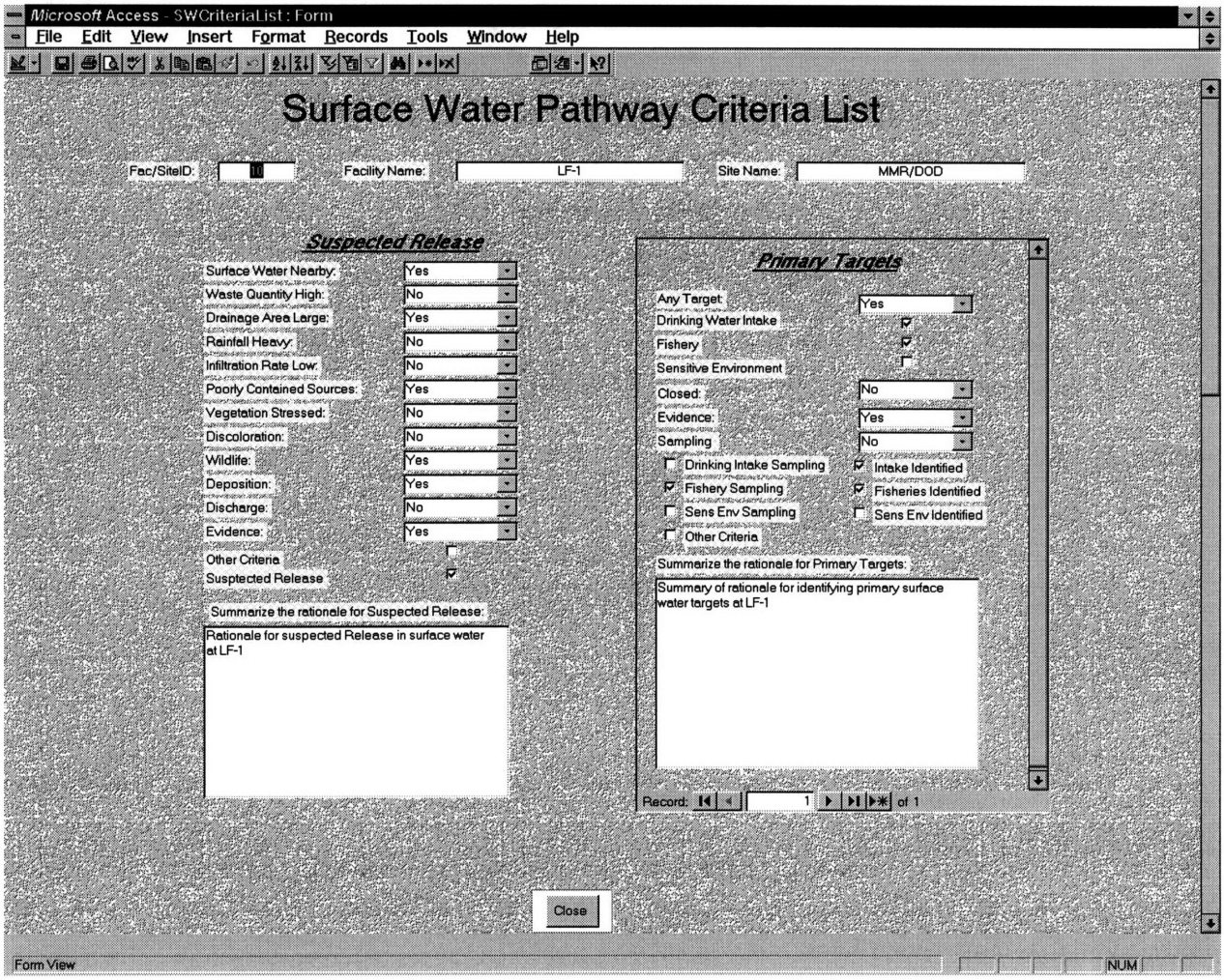
Subforms: none

Form Data: GeneralFacilityInformation table.

Events and Macros: Clicking on the "Surface Water Criteria List" and "Surface Water Scoring" buttons opens the SWCriteriaList and SWSS forms respectively. The corresponding macros are OpeningForms.SWCriteria and OpeningForms.SWSS. These macros also synchronize the newly opened forms with the Surface Water Main form so that the record pertaining to the same facility and site is viewed. The form is closed when the Close button is clicked.

Opened From: Review PA Sections form

Figure D3. Main Surface Water Pathway Form



Form Name: SWCriteriaList

Description: This form contains the information contained on the Surface Water Pathway Criteria List section of the PA scoresheet

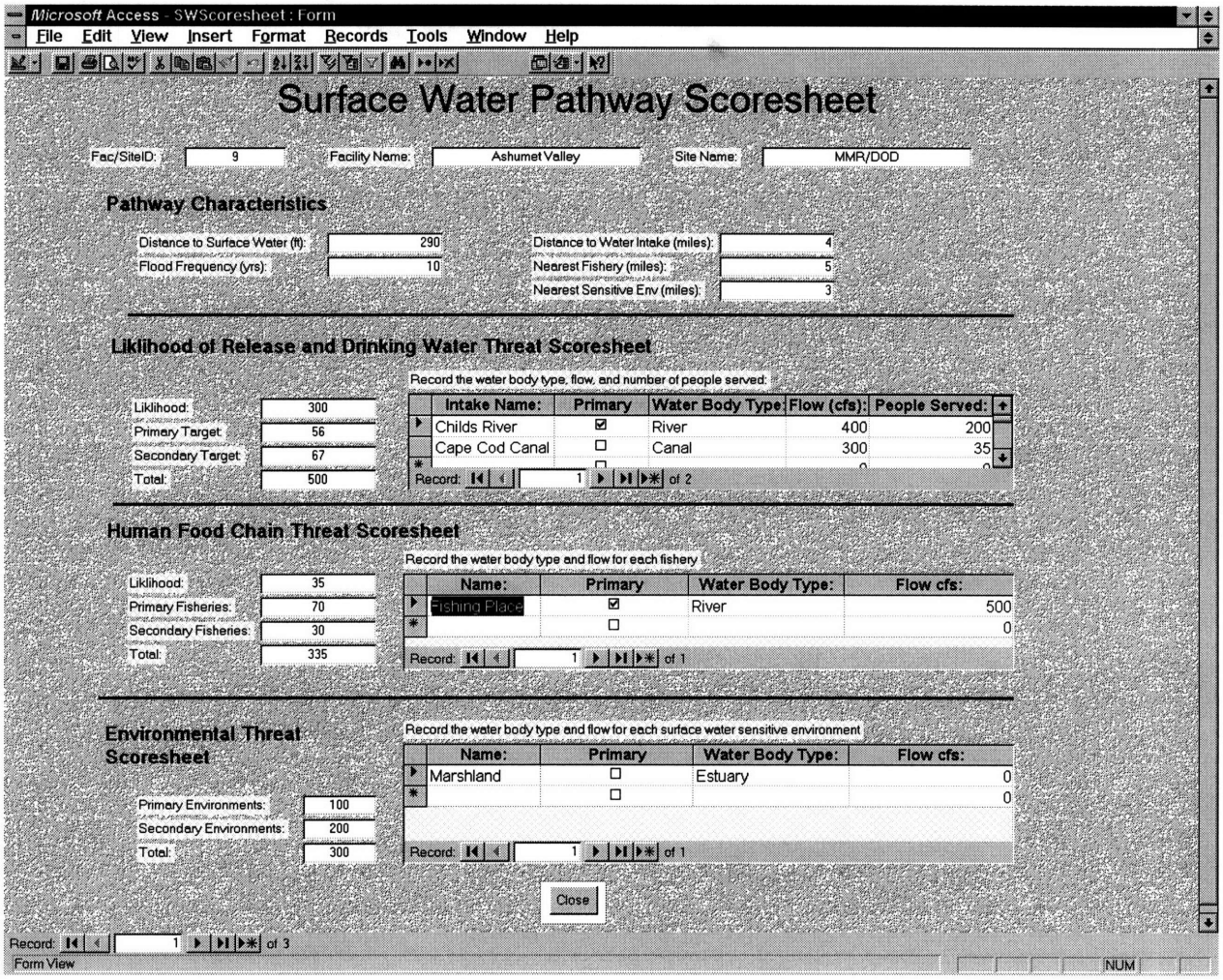
Subforms: SWPrimary Subform (data based on SWPrimaryTargetQuery query which is based on the SurfaceWater-Primary Targets table and the Fac/SiteID from the GeneralInformation table).

Form Data: SurfaceWaterQuery query. This query essentially returns all data fields related to the Surface Water Pathway.

Events and Macros: The form is closed when the Close button is clicked.

Opened From: Surface Water Main form

Figure D4. Surface Water Pathway Criteria List Form



Form Name: SWScoresheet

Description: This form contains the data recorded in the Surface Water Pathway scoresheet.

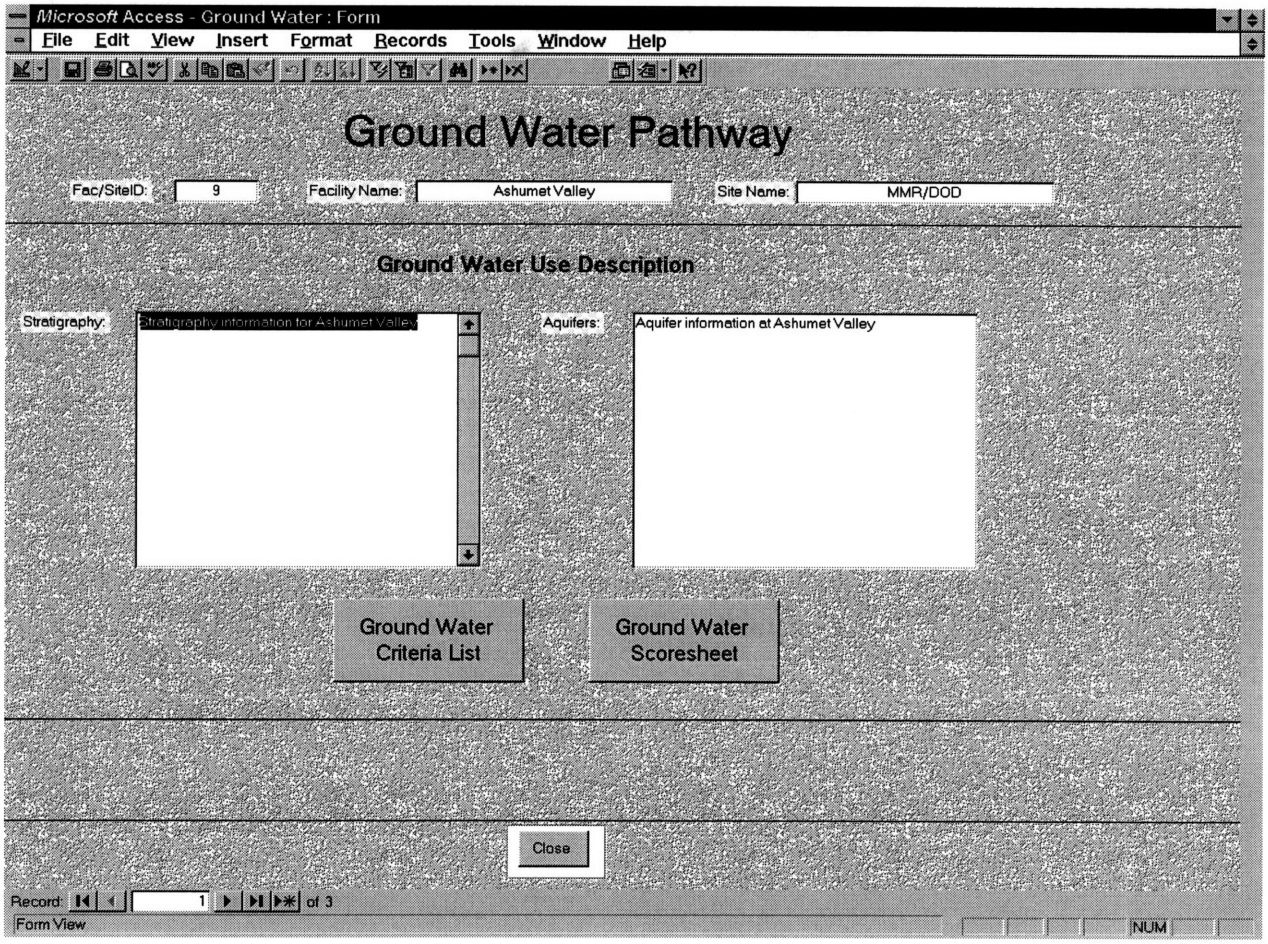
Subforms: SWDrikingWater Subform (data based on SW-WaterBodies-DrinkingWaterThreat), SWHumanFood Subform (data based on SW-WaterBodies-HumanFd/EnvThreat table), SWEnvironmental Subform (data based on SW-WaterBodies-HumanFd/EnvThreat table)

Form Data: SurfaceWaterQuery query. This query essentially returns all data fields related to the Surface Water Pathway.

Events and Macros: The form is closed when the Close button is clicked.

Opened From: Surface Water Main form

Figure D5. Surface Water Pathway Scoresheet Form



Form Name: Ground Water

Description: This form contains information about the stratigraphy and aquifers on the facility and two buttons that open the Ground Water Criteria List and Ground Water Scoresheet Forms

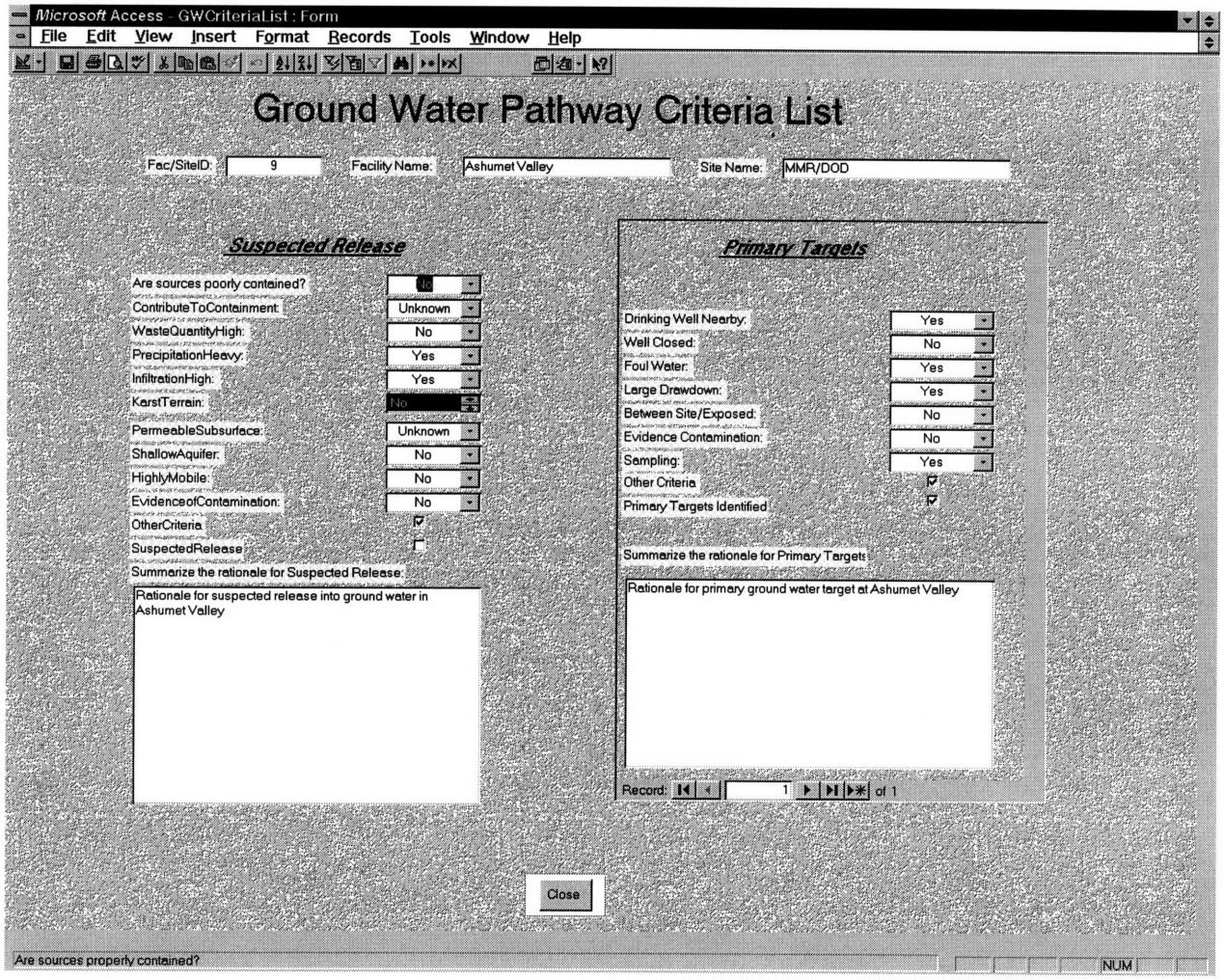
Subforms: none

Form Data: GroundWaterQuery query. This query essentially returns all data fields related to the Ground Water Pathway.

Events and Macros: Clicking on the "Ground Water Criteria List" and "Ground Water Scoresheet" buttons opens the GWCriteriaList and GWSS forms respectively. The corresponding macros are OpeningForms.GWCriteria and OpeningForms.GWSS. These macros also synchronize the newly opened forms with the Ground Water form so that the record pertaining to the same facility and site is viewed. The form is closed when the Close button is clicked.

Opened From: Review PA Sections form

Figure D6. Main Ground Water Pathway Form



Form Name: GWCriteriaList

Description: This form contains the information contained on the GroundWater Pathway Criteria List section of the PA scoresheet.

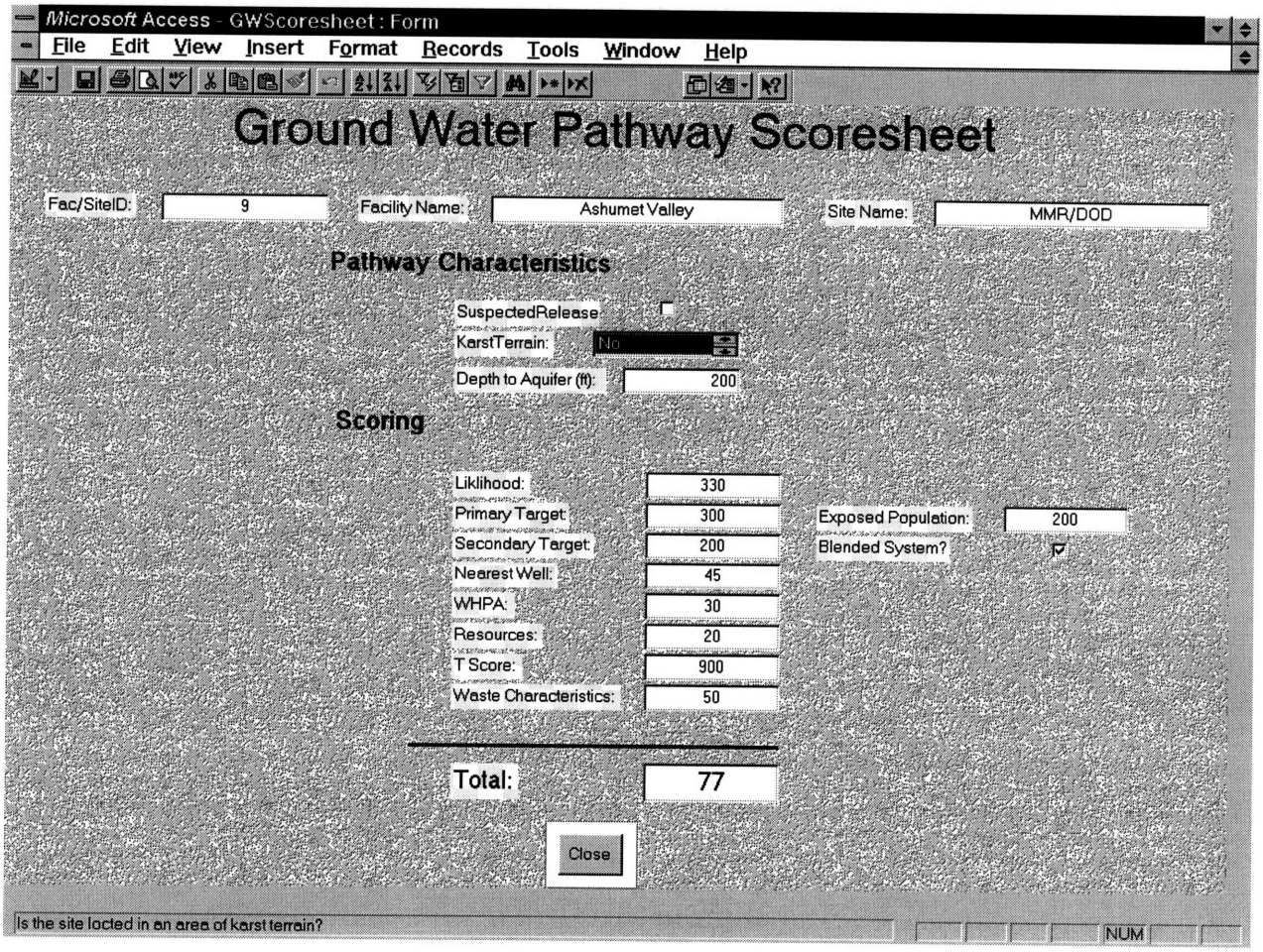
Subforms: GWPrimary Subform (data based on GWPrimary Target query which used the GroundWater-Primary Targets table and the Fac/SiteID from the GeneralFacilityInformation table)

Form Data: GroundWaterQuery query. This query essentially returns all data fields related to the Ground Water Pathway.

Events and Macros: The form is closed when the Close button is clicked.

Opened From: Ground Water

Figure D7. Ground Water Pathway Criteria List Form



Form Name: GWScoresheet

Description: This form contains the data recorded in the Ground Water Pathway scoresheet.

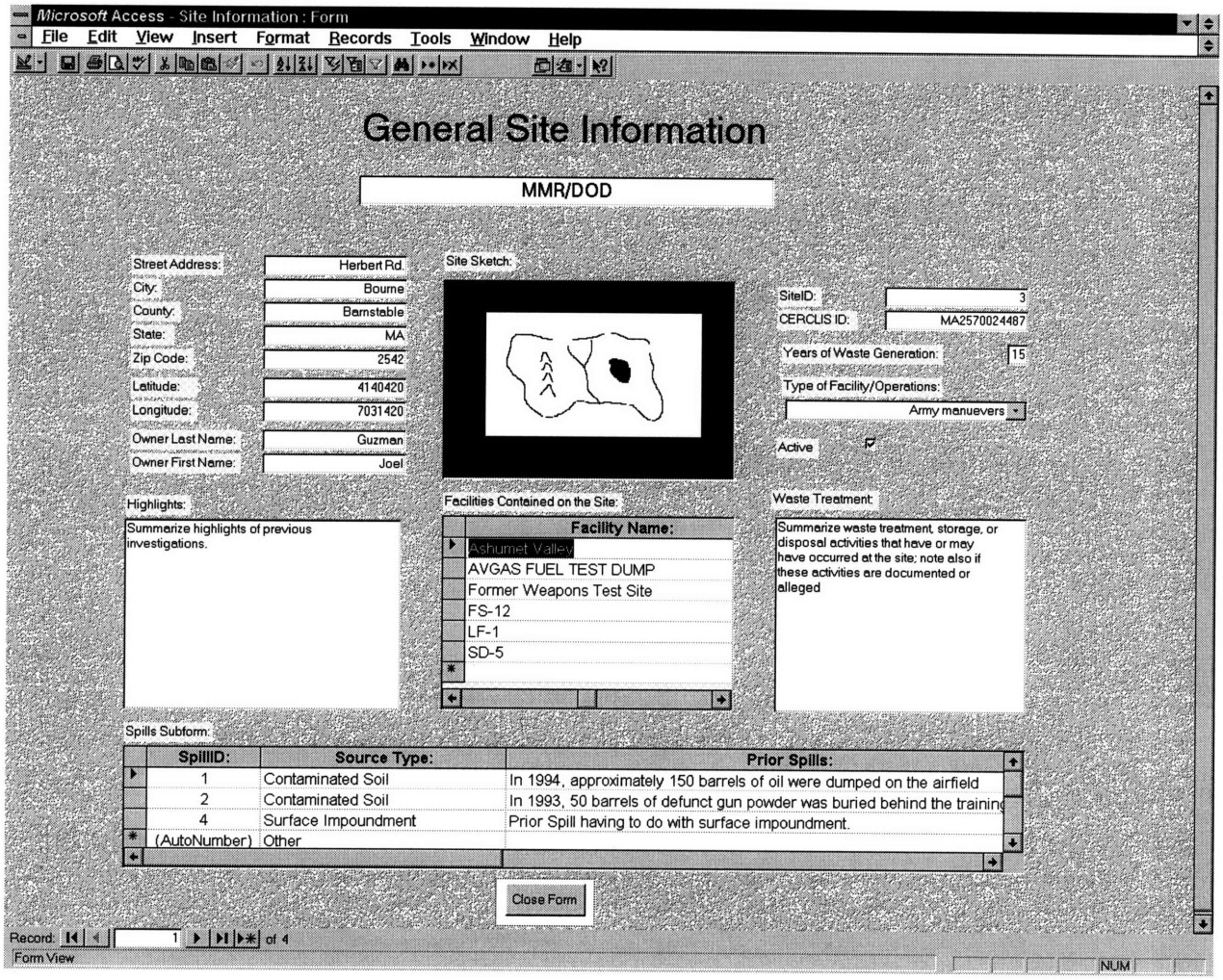
Subforms: none

Form Data: GroundWaterQuery query. This query essentially returns all data fields related to the Ground Water Pathway.

Events and Macros: The form is closed when the Close button is clicked.

Opened From: GroundWater

Figure D8. Ground Water Pathway Scoresheet Form



Form Name: Site Information

Description: This form contains all the data stored in the GeneralSiteInformation table plus a list of facilities on that site.

Subforms: Facilities Subform (data based on ListofSitesandFacilities query. This query consists of the data fields Fac/SiteID and Facility Name from the GeneralFacilityInformation table and the Site Name data field from the GeneralSiteInformation table.)

Form Data: GeneralSiteInformation table

Events and Macros: none

Opened From: Main Switchboard (from Figure 5.2.1a)

Figure D9. General Site Information Form