

**PB88-192398**

**Public Report for Options to Make the  
Toxic Release Inventory (TRI) Data Base  
Accessible to the Public**

**CRC Systems, Inc., Fairfax, VA**

**Prepared for:**

**Environmental Protection Agency, Washington, DC**

**4 Mar 88**

## BIBLIOGRAPHIC INFORMATION

PB88-192398

Public Report for Options to Make the Toxic Release  
Inventory (TRI) Data Base Accessible to the Public,

4 Mar 88

PERFORMER: CRC Systems, Inc., Fairfax, VA.

SPONSOR: Environmental Protection Agency, Washington, DC.  
Office of Toxic Substances.

Sponsored by Environmental Protection Agency, Washington,  
DC. Office of Toxic Substances.

The purpose of the paper is to present appropriate options for implementation of a publicly accessible Toxic Release Inventory (TRI) data base, to analyze and evaluate the costs and benefits of those options, and to recommend one or more of the best alternatives for making the TRI data base available to the public. The analysis addresses only the data base options. It does not address 'other means' of public access to TRI data, e.g., printed versions of the data or Freedom of Information Act (FOIA) requests, other than to note the potential effect of 'other means' on usage of the data base.

KEYWORDS: \*Toxic substances, \*TRI database, \*Public access.

Available from the National Technical Information Service,  
SPRINGFIELD, VA. 22161

PRICE CODE: PC A04/MF A01

**PUBLIC REPORT FOR  
OPTIONS TO MAKE THE TOXIC  
RELEASE INVENTORY (TRI) DATA  
BASE ACCESSIBLE TO THE PUBLIC**

**Prepared for:**

**Office of Toxic Substances  
Environmental Protection Agency  
401 M Street, S.W.  
Washington, D.C. 20460**

**Prepared by:**

**CRC Systems, Inc.  
11242 Waples Mill Road  
Fairfax, VA 22030  
(703) 359-9400**

**March 4, 1988**

## TABLE OF CONTENTS

| Section  | Description                              | Page        |
|----------|--|-------------|
|          | <b>EXECUTIVE SUMMARY</b>                 | <b>ES-1</b> |
| <b>1</b> | <b>INTRODUCTION</b>                      | <b>1-1</b>  |
|          | 1.1 BACKGROUND                           | 1-1         |
|          | 1.2 PURPOSE OF PAPER                     | 1-2         |
|          | 1.3 USER COMMUNITY                       | 1-2         |
|          | 1.4 BIBLIOGRAPHY                         | 1-3         |
|          | 1.5 GLOSSARY OF TERMS                    | 1-3         |
| <b>2</b> | <b>APPROACH</b>                          | <b>2-1</b>  |
|          | 2.1 INFORMATION GATHERING                | 2-1         |
|          | 2.1.1 Requirements Analysis              | 2-1         |
|          | 2.1.2 FIRMR                              | 2-1         |
|          | 2.1.3 ETD Demand Analysis                | 2-2         |
|          | 2.1.4 Commerce Business Daily            | 2-4         |
|          | 2.1.5 Analysis of Additional Information | 2-5         |
|          | 2.2 ASSUMPTIONS                          | 2-5         |
|          | 2.3 POTENTIAL SYSTEM IMPACTS             | 2-7         |
|          | 2.4 EVALUATION CRITERIA                  | 2-8         |
|          | 2.4.1 Minimum Requirements               | 2-8         |
|          | 2.4.2 Optional Characteristics           | 2-9         |
| <b>3</b> | <b>DATA BASE OPTIONS</b>                 | <b>3-1</b>  |
|          | 3.1 BACKGROUND                           | 3-1         |
|          | 3.1.1 Overall System Description         | 3-1         |
|          | 3.1.2 System Costs                       | 3-2         |
|          | 3.1.3 User Costs                         | 3-3         |
|          | 3.2 EPA CLEARINGHOUSE                    | 3-3         |
|          | 3.2.1 System Description                 | 3-3         |
|          | 3.2.2 Costs                              | 3-6         |

|            |  |             |
|------------|--|-------------|
| <b>3.3</b> | <b>COMMERCIAL VENDORS/INFORMATION PROVIDERS</b>                    | <b>3-8</b>  |
| 3.3.1      | System Description   | 3-8         |
| 3.3.2      | Costs  | 3-9         |
| <b>3.4</b> | <b>IAG WITH GOVERNMENT AGENCY PROVIDER</b>                         | <b>3-11</b> |
| 3.4.1      | System Description   | 3-11        |
| 3.4.2      | Costs  | 3-11        |
| <b>3.5</b> | <b>AGREEMENT WITH NON-PROFIT ORGANIZATION INFORMATION PROVIDER</b> | <b>3-12</b> |
| 3.5.1      | System Description   | 3-12        |
| 3.5.2      | Costs  | 3-12        |
| <b>3.6</b> | <b>EVALUATION OF REQUIREMENTS</b>                                  | <b>3-13</b> |
| 3.6.1      | Tier A   | 3-16        |
| 3.6.2      | Tier B   | 3-20        |
| 3.6.3      | Tier C   | 3-20        |
| <b>3.7</b> | <b>COMPARISON OF POTENTIAL RISKS</b>                               | <b>3-22</b> |
| 3.7.1      | Timely Negotiation   | 3-22        |
| 3.7.2      | Timely System Implementation                                       | 3-22        |
| 3.7.3      | Long Term Availability   | 3-24        |
| <b>3.8</b> | <b>COMPARISON OF SYSTEM COSTS</b>                                  | <b>3-24</b> |
|            | <b>SUMMARY</b>   | <b>4-1</b>  |
| <b>4.1</b> | <b>COMPARISON ANALYSIS</b>   | <b>4-1</b>  |
| 4.1.1      | Comparison of System Costs   | 4-1         |
| 4.1.2      | Comparison of User Costs   | 4-1         |
| 4.1.3      | Comparison of Capability to Meet Requirements                      | 4-3         |
| 4.1.4      | Comparison of Potential Risks                                      | 4-3         |
| 4.1.5      | Summary of Comparisons   | 4-4         |
| <b>4.2</b> | <b>STRATEGIC APPROACHES</b>  | <b>4-6</b>  |
| 4.2.1      | Unembellished Approach   | 4-7         |
| 4.2.2      | Enhanced System Approach   | 4-7         |
| <b>4.3</b> | <b>EPA PLANNED APPROACH</b>  | <b>4-8</b>  |

## **APPENDICES**

**APPENDIX A**

**APPENDIX B**

**APPENDIX C**

**BIBLIOGRAPHY**

**GLOSSARY**

**OPTION FEATURES**

**A-1**

**B-1**

**C-1**

## **LIST OF EXHIBITS**

| <b>Exhibit</b> | <b>Description</b>                                 | <b>Page</b> |
|----------------|--|-------------|
| 2-1            | Application of FIRMR to Selection of Options       | 2-3         |
| 3-1            | EPA Clearinghouse                                  | 3-5         |
| 3-2            | Summary of Cost to Government                      | 3-7         |
| 3-3            | Ability to Meet Requirements                       | 3-14        |
| 3-4            | Average Annual User Access Cost                    | 3-17        |
| 3-5            | Comparison of Potential Risk                       | 3-23        |
| 3-6            | Comparison of System Costs                         | 3-25        |
| 4-1            | Comparison of Options                              | 4-2         |
| 4-2            | Cumulative Benefits for Tier B at First Year Costs | 4-5         |

## **EXECUTIVE SUMMARY**

### **Purpose**

The purpose of this paper is to present appropriate options for implementation of a publicly accessible Toxic Release Inventory (TRI) data base, to analyze and evaluate the costs and benefits of those options, and to recommend one or more of the best alternatives for making the TRI data base available to the public. This analysis addresses only the data base options. It does not address "other means" of public access to TRI data, e.g., printed versions of the data or Freedom of Information Act (FOIA) requests, other than to note the potential effect of "other means" on usage of the data base.

### **Analysis**

The four options presented and analyzed in this paper are:

- o An EPA clearinghouse, where the EPA would implement and support a public access data base from an existing Federal Data Processing Center.
- o A commercial vendor, using the existing environment and support structure of a commercial data base vendor specializing in chemical information systems.
- o An interagency agreement (IAG) with a government agency that has the experience and ability to support the data processing and user support aspects.
- o An agreement with a non-profit organization information provider, such as a university, which will provide processing and user support through an existing structure.

The cost of each of the options has been determined for hardware/software, maintenance, system development and ongoing application maintenance, systems operations and maintenance staff, telecommunications, marketing, user support, account management, and additional facility support costs, where applicable. Net costs to the Government were calculated for the first year costs and an average of five years of annual cost with maximum estimated system usage. In addition,



average annual costs to the users were computed to enable evaluation of each option to provide the data base to the public at a reasonable cost.

The options were analyzed to determine how well they meet the requirements for the public access data base. System requirements were considered in three tiers, i.e., Tiers A, B, and C, where Tier A requirements represent the minimum requirements mandated by legislation and/or specified by the EPA, and Tier B and C requirements represent optional enhanced characteristics which are desirable for the TRI public data base to enhance data utility and analysis and display characteristics. The requirements for each tier are provided in the following chart:

| REQUIREMENTS                  |        |        |        |
|-------------------------------|--------|--------|--------|
|                               | TIER A | TIER B | TIER C |
| HARDWARE AND SOFTWARE         |        |        |        |
| ACCESSIBLE                    |        |        |        |
| REASONABLE COST               |        |        |        |
| AGGREGATION CAPABILITIES      |        |        |        |
| ACCUMULATION                  |        |        |        |
| SEARCH AND RETRIEVAL          |        |        |        |
| DOWNLOADING CAPABILITIES      |        |        |        |
| ACCOUNT MANAGEMENT            |        |        |        |
| SYSTEM AVAILABILITY           |        |        |        |
| USER SUPPORT                  |        |        |        |
| STATISTICAL ANALYSIS SOFTWARE |        |        |        |
| COMPLEMENTARY FILES           |        |        |        |
| MENU DRIVEN SCREENS           |        |        |        |
| MAPPING CAPABILITIES          |        |        |        |
| CROSSLINKS TO EPA DATA BASES  |        |        |        |

Potential risks for the options in terms of timeliness and long term availability are also discussed.

## Findings

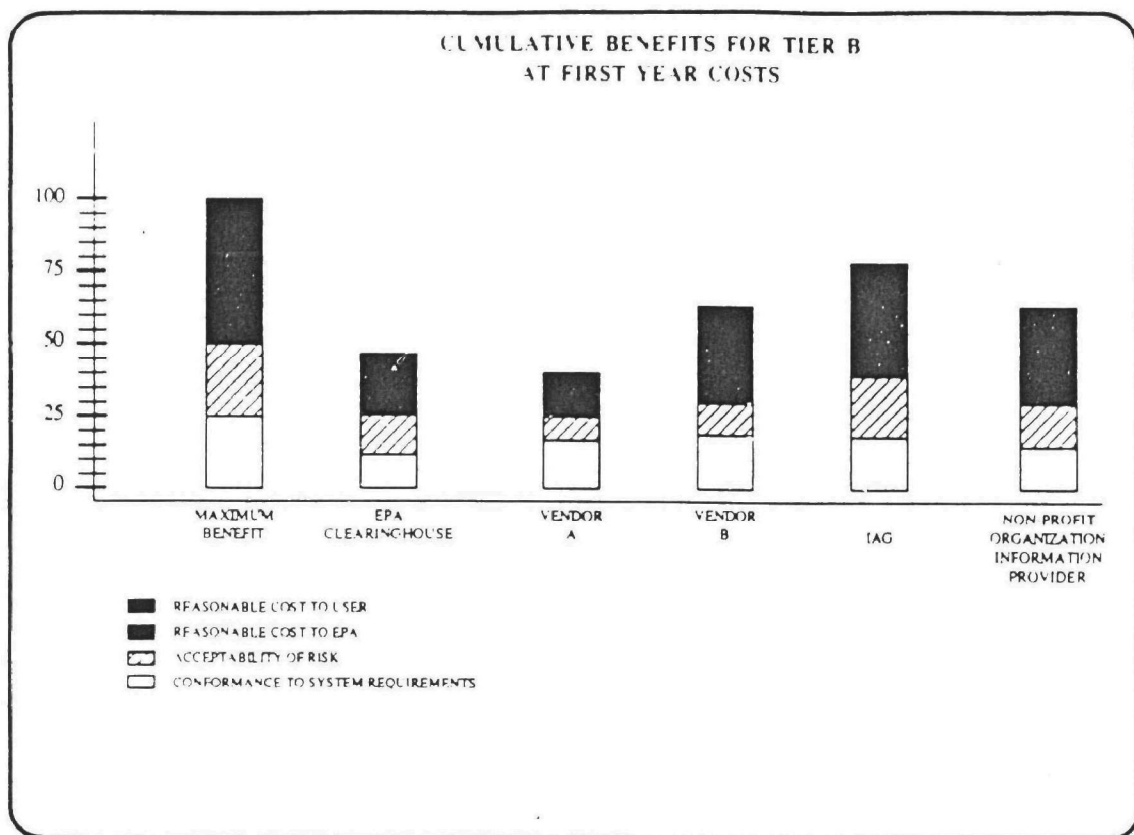
The EPA Clearinghouse does not appear to be a viable option because of its high initial cost, the high level of risk for timely implementation and its minimal level of benefit.

The Commercial vendor option has the potential for meeting all requirements, at reasonable cost, and with a satisfactory or greater level of benefits. However, the desirability for a vendor to maintain the TRI on a long term basis is dependent on the user community and the demand for access to the data which affects the profitability of the system. System usage will be affected by the data quality, the volume of data received, the usage of "other means" of TRI access, and other factors. In addition, competitive procurement from commercial vendors cannot be accomplished within as short a timeframe as an interagency agreement, increasing the risk that the system might not be available when needed.

The IAG option has the potential for providing the best benefit for the user requirements with the best complementary file access for health and environmental effects data. The option represents moderate implementation cost to the EPA and reasonable cost to the user.

The agreement with a non-profit organization information provider represents adequate support of the TRI public access data base within a short implementation timeframe with reasonable cost to the EPA and to the user. However, the availability of crosslinked and complementary files is limited. The organization presented in the document supports a very narrow market, which would need to be expanded to meet the needs of the TRI. In addition, procurement of a contract vehicle may not be accomplished within as short a timeframe as an interagency agreement, increasing the risk that the system might not be available when needed.

The following bar chart illustrates the potential of each of the options to meet maximum benefit, when consideration is given to first year costs and Tier B level of



implementation. Because system costs, favorable user fees, conformance with system requirements, and least risk for timely, long-term implementation are given equal consideration on the chart, this chart can best be viewed as providing a relative comparison of the various options.

#### EPA Planned Approach

The recommended strategic approach is to select an implementation option with low risk, low to moderate initial system cost, and favorable benefits, including favorable user costs and system capabilities, at the Tier A level. Tier A has been selected because of funding constraints. The EPA planned approach has been to select the IAG option which has been determined to be the option which most closely provides the best balance of all evaluation criteria established for this analysis. The IAG option has the potential of providing an "enhanced system" at the partial Tier B level for the initial implementation, while being funded at the Tier A level. Overall, the IAG provides the best balance of the evaluation criteria at the Tier A level while providing

many complementary and environmental effects files (partial Tier B). EPA believes these files will provide enhanced utility of the data collected under the Community-Right-to-Know legislation.

The EPA will periodically evaluate and reassess the need to continue support or to enhance the system capabilities as the system usage and availability through other sources are established.

Commercial vendors will also be encouraged to explore the market potential for a TRI public data base within their own systems, using magnetic tapes of TRI data available through the National Technical Information Service (NTIS).

## **SECTION 1 INTRODUCTION**

### **1.1 BACKGROUND**

Title III, Section 313 of the Superfund Amendments Reauthorization Act (SARA) of 1986, requires that the Environmental Protection Agency (EPA) collect data for the annual releases of certain toxic chemicals to the environment, according to the specific thresholds as defined in Section 313(f) of the law. The list of the chemicals to be regulated is included in Senate Environment and Public Works Committee Print Number 99-169, as it may be revised by EPA. The initial list contains 309 specific chemicals and 20 additional categories of chemicals. The law further requires that the data collected under this legislation be made accessible to the public by computer through telecommunication access and by other means, on a cost reimbursable basis.

The Toxic Chemical Release Reporting; Community Right-to-Know proposed rule was published in the Federal Register, Volume 52, Number 107, June 4, 1987.(1) The final rule was published February 16, 1988.

Section 322 of Title III permits chemical identity to be claimed as a trade secret, where substantiation for the trade secret claim is provided to the EPA. Where claims for trade secret prevent making the identity of the released chemicals known, information concerning the health and environmental effects of the chemicals must be provided to the public. A proposed rule for handling of the Title III trade secret claims was published in the Federal Register, Volume 52, Number 199, October 1987.(2) The final rule is scheduled for publication June 1988.

The Office of Toxic Substances (OTS) has been assigned responsibility for implementation of the Toxic Release Inventory (TRI). The Information Management Division (IMD) of OTS has worked closely with other OTS divisions, other offices within EPA Headquarters, EPA regions, states, and the public to determine the requirements for implementation of the TRI data base.

The first reporting cycle requires that releases to the environment during the 1987 calendar year must be reported to EPA and the states between January 1 and July 1 of 1988. It is expected that the data base for these data will be available for public access in 1989.

## **1.2            *PURPOSE OF PAPER***

The purpose of this paper is to present appropriate options for implementation of a publicly accessible TRI data base, to analyze and evaluate the costs and benefits of those options, and to recommend one or more of the best alternatives for making the TRI data base available to the public. This analysis addresses only the data base options. It does not address the "other means" of public access to TRI data, e.g., printed versions of the data or reports or Freedom of Information Act requests, other than to note the potential effect of "other means" on usage of the data base.

## **1.3            *USER COMMUNITY***

The potential user community for the TRI public data base is expected to be comprised primarily of the private sector, including users in the following areas: manufacturing, information research, chemistry, pharmaceuticals, law, medicine, public libraries, academics, media, and non-profit areas. EPA Headquarters staff, other government agencies, states, and EPA regions are expected to have access to the internal EPA TRI data base, which will contain the identical data, as well as the information claimed as trade secret available only on the internal system. Therefore, the Government sector is not expected to be a substantial user of the publicly available data base.

Possible competition for usage of the TRI data base will result from the "other means" of public access to be provided by the EPA. Other means under consideration include hard copies of forms and data reports, computer disks and tapes, and optical disks. Regardless of the packaging and distribution of toxic release inventory data, a computer tape will be available to the public through the National Technical Information Services (NTIS) for all entities which have capabilities for computer processing. Access to the Section 313 data may also be available through Freedom of Information Act (FOIA) requests. Therefore, the estimated user group may be affected by the needs of the user community and by the "other means" of distribution that EPA implements.

#### **1.4            *BIBLIOGRAPHY***

Much research and analysis have been undertaken towards the definition and evaluation of data base options. Previous studies and publications are referenced in this report, where appropriate, without repetition of the details of those studies. Conclusions will be restated where necessary. References, indicated with a number in parentheses, are listed in the bibliography in Appendix A.

#### **1.5            *GLOSSARY OF TERMS***

A glossary of terms and acronyms used in this report is provided in Appendix B.

## **SECTION 2**

### **APPROACH**

The approach used to prepare this options paper included information gathering, documentation of the assumptions associated with the implementation of the data base, concerns for factors that have potential to affect the system development, and definition of the system requirements and criteria to be used for evaluating the alternatives.

#### **2.1            *INFORMATION GATHERING***

##### **2.1.1        *REQUIREMENTS ANALYSIS***

OTS has performed an analysis of user requirements for EPA Headquarters, EPA Regions, other Federal Agencies, and the public to identify the types of searches, reports, information subsets and aggregates which are expected to serve the needs of a large portion of the potential user community.<sup>(3)</sup>

In January 1987, a draft options paper <sup>(4)</sup> was prepared which presented four options for computer-based TRI access, including an EPA controlled clearinghouse, commercial vendor support, an interagency agreement (e.g., with the National Library of Medicine (NLM)), and a cooperative agreement (e.g., following the model of the National Pesticide Information Retrieval System (NPIRS) set up by Purdue University). The options paper was the focus of a public meeting held April 20, 1987, generating additional comments and concerns for public availability of the TRI data.<sup>(5)</sup>

##### **2.1.2        *FIRMR***

The Federal Information Resources Management Regulation (FIRMR) <sup>(6)</sup> is concerned with the basic policies and regulations for managing information resources within the Federal Government. The FIRMR establishes policy for the analysis of alternatives (41 CFR 201-30.009) and resource sharing (41 CFR 201-31.002) with regard to resource acquisition and utilization.



The FIRMR requires that an alternatives analysis consider the following with respect to information management: the use of non-ADP resources, the use of existing facilities and resources on a shared basis (e.g., Federal Data Processing Centers (FDPC) and Interagency Agreements), the use of commercial ADP services, the redesign of existing applications (software), and the augmentation and upgrading of existing configurations and system components (hardware). Consideration must also be given to requirements for facilities, personnel, and other applicable factors, including such things as privacy or security implications. The alternatives presented in this document in Section 3 are in keeping with these requirements, as briefly summarized in Exhibit 2-1.

In keeping with the FIRMR requirements, several existing government systems were examined.<sup>(7)</sup> These systems included the Justice Retrieval and Inquiry System (JURIS), the Security and Exchange Commission Pilot Data base (EDGAR), Systems of the Nuclear Regulatory Commission (NRC), the Storage and Retrieval of Water Quality Information System (STORET), the Program Integration Project Queries Use in Interactive Command (PIPQUIC), the Utility Emission Inventory Data (UTIL), the National Institute of Occupational Safety and Health (NIOSH) Systems, and the Patent and Trademark Office (PTO) System. The possibility of providing public access to the TRI data base using a methodology similar to one of these has not proved to be feasible.

### **2.1.3 ETD DEMAND ANALYSIS**

The Economics and Technology Division (ETD) of OTS prepared a preliminary demand analysis to determine the potential market size and user community for the public TRI data base. Using a simple modeling approach, usage, in terms of number of online accesses, was assumed to be a function of price using a constant elasticity formulation, i.e., relative elasticities across user categories were estimated, based on assumptions about user category price sensitivity. The analysis was based on data on the overall online information market from the Information Industry Factbook '87 (Digital Information Group), estimates of data base usage by customer type and average session duration from an online data vendor of scientific information, customer usage frequency estimates from contractor interview reports, and online data base prices from price lists obtained from most of the major private online data base vendors of both

| REQUIREMENT | OPTION  | RATIONALE   |
|-------------|---|---|
| 1.          | Non-ADP Implementation  | The TRI data base cannot be limited to non-ADP resources because of legislative mandate for a computer data base and because the size and complexity of the data require the use of ADP resources. Data will also be distributed by other, non-ADP means, as mandated in the legislation. |
| 2.          | Shared Resources  | Option 1, EPA Clearinghouse (FDPC)<br>Option 3, Interagency Agreement<br>Option 4, Non-Profit Organization Information Provider   |
| 3.          | Commercial ADP Services   | Option 2, Commercial Vendors.   |
| 4.          | Redesign Existing Applications  | Redesign of the EPA internal TRI application to accommodate public usage was rejected based on the EPA obligation to preserve trade secret data, and the EPA determination of how best to fulfill that obligation.  |
| 5.          | Augmentation and Upgrading of Existing Configurations and System Components | Option 1, the EPA Clearinghouse, involves the augmentation and upgrading of the National Computer Center (NCC), the FDPC used for existing EPA ADP Systems.<br>Option 3, Interagency Agreement, may also require augmentation and upgrading of existing systems.                          |

Application of FIRMR to Selection of Options  
Exhibit 2-1

text and numerical data bases. Estimates of potential academic, non-profit and media users were obtained from Statistical Abstracts, World Environmental Directory and Environmental U.S.A.

The estimates of TRI demand were based on an estimated total potential user population of about 250,000. The total was distributed among user categories based on estimates of data base usage by customer type. Since only a fraction of the users of any online data base service would be expected to use the TRI data base even once a year, an estimate of usage of chemistry related files as a percentage of total data base usage was used as a proxy for the potential usage of TRI data. This percentage was 0.8%. Using an average online session of 20 minutes and an average number of sessions of three per month total usage was estimated to be 12 hours per user per year. Current prices for online public data bases produced by Federal agencies range from \$25 to \$84 per hour of usage. The price of an hour usage of scientific/technological and medical data bases offered through private vendors ranges from \$36 to \$153.

Using the above inputs, the number of potential TRI users was calculated to be approximately 2,000. Given an assumed price of \$100 per hour of usage, the estimated number of annual data base accesses is about 88,000, assuming that no fee waiver is provided. Since the users most likely to meet the "public interest" requirement for fee waivers are also likely to be price sensitive, it is estimated that total usage could be increased substantially through provision of fee waivers. However, available data do not allow this increase to be quantified reliably.

#### 2.1.4 *COMMERCE BUSINESS DAILY*

The EPA published an announcement in the Commerce Business Daily (CBD) (9) in June 1987, seeking information on online interactive systems for the Community Right-to-Know Data Base. Functional capabilities requested included powerful and easy to use indexing, searching, and retrieving of numerical and textual data on chemical substances; complementary files, with integrated search capability for health and environmental effects, toxicology and chemical industry data; support for statistical analysis; standard and *ad hoc* report generation with possible graphic output; and support for monitoring trends over and between reporting years. In addition to data base system functions, user support, training, marketing, documentation, user account

management, and telecommunications were also specified as necessary. System availability for a minimum of 16 hours per day, 7 days a week was also defined as a requirement.

Responses to the CBD announcement were received from commercial data base vendors and academic institutions (non-profit organizations) already engaged in distributing information through publicly available data bases. Others interested in developing the public data base system also responded. The information provided by responders was analyzed to determine the potential for the private sector vendors to meet EPA's specifications for the TRI data base. The information was used for preparing financial estimates for Option 2, i.e., the commercial vendors, and Option 4, i.e., agreement with non-profit organization information providers. The information, in conjunction with the information obtained from the survey of government data bases and systems, was also used as a basis for estimating the potential costs for operating an EPA clearinghouse.

## **2.1.5 ANALYSIS OF ADDITIONAL INFORMATION**

Additional information was collected and analyzed in the form of literature searches, interviews with operators of existing clearinghouses, and reviews of other existing EPA systems. Interviews were conducted with the following clearinghouses: the National Air Toxics Information Clearinghouse (NATICH), the local Public Document Rooms of the Nuclear Regulatory Commission (NRC), National Pesticide Telecommunications Network (NPTN), the Toxic Substances Control Act (TSCA) Information Service, the Chemical Emergency Preparedness Program (CEPP) Hotline, the Educational Resources Information Center (ERIC), the Environmental Information Exchange Network of the Environmental Defense Fund, and the National Chemical Response and Information Center (NCRIC) and the Chemical Transportation Emergency Center (CHEMTREC) of the Chemical Manufacturers Association (CMA). This information contributed to the definition of the EPA clearinghouse option and to compilation of additional information for the other options.

## **2.2 ASSUMPTIONS**

Assumptions which have been made regarding the implementation of the TRI data base for public access are the following:

- o The public data base will be maintained with semi-annual incremental updates or reloads from files extracted from the EPA internal system. Ongoing data base update, data validation and quality assurance, and data correction will take place on the EPA internal system.
- o Storage capacity is required for approximately 480Mb of uncompressed data in a flat file in the first year.<sup>(10)</sup> In the second year, as the threshold for reporting decreases for manufacturers and processors, it is expected that about 576 Mb of data will be added, i.e., about 20% more submitters will provide yearly reports. For each subsequent year, reporting is anticipated to increase by about 20% over the previous year. The data base with an accumulation of five years' reporting would therefore contain 3.3 Gb, i.e., 680% of the first year's data.
- o The publicly available system will require preprogrammed access to enable accurate search, retrieval and analysis. This is necessary to avoid risks involved with *ad hoc* searching and data analysis, where incomplete data may be compiled to give inconclusive and erroneous results.
- o Standard pre-programmed data analysis, e.g., means, averages, ratios, percentages, standard deviation, and linear regression, will be included in the basic system. More sophisticated statistical analysis is an optional characteristic for an embellished system.
- o Access via telecommunications will be available to individual users, without the need for requesting searches from professional information intermediaries.
- o Fee waivers will be managed according to policy established by the EPA based on economic and other factors.<sup>(11)</sup>
- o The user community for the purpose of this study and the estimated data base usage is based on the ETD demand analysis.<sup>(8)</sup> However, assuming that the analysis is based on full implementation of the TRI system through Tier C (see Section 2.4), this options paper has estimated Tier A usage at one-half and Tier B at three-fourths of the total usage provided in the demand analysis. Usage increases which may result as the result of an expanded user community have not been considered in these estimates.
- o The identity of chemicals for which a trade secret claim has been substantiated will not be included on the public data base. However, release data for the chemical will be contained in the data base, with a generic descriptor for the chemical.
- o Regardless of other implementation options, magnetic tapes containing all data for the public version will be provided to NTIS for distribution to the public on an ongoing basis.

## **2.3**

### **POTENTIAL SYSTEM IMPACTS**

Issues were identified in the performance of this study which may impact the system usage and the cost of the options to the EPA and to the public. The impact of the following issues cannot be determined until experience has been obtained from TRI reporting and TRI system usage:

- o The exact user community is not well-defined for the system. User estimates have been based on a non-representative sample of the bibliographic systems as opposed to numeric systems like the TRI data base content. Therefore, the user community could be much larger or smaller than is indicated in this document.
- o The actual amount of TRI data is unknown. The estimates are based on early studies, and actual quantity of data will not be known until after the submission of data. This will impact the estimated costs for data storage, and potentially the level of interest in the system by the public.
- o It is unclear how many of the submissions will be claimed as trade secret, which will impact the amount and utility of the data for interface with complementary data files and other EPA data files.
- o The data quality is unknown at this point in the study. Poor data will certainly affect the amount of usage of the system.
- o The "other means" by which TRI will be available will impact the potential user community for the public access data base. It is unclear this early in the study how many people will actually take advantage of the "other means" and how adequate the "other means" will be to satisfy the user's needs.
- o The public may prefer to address questions to the EPA Headquarters, EPA Regions, the States, or even local offices (e.g., local fire department). This method of access will dilute the potential user community of the online data base.
- o Vendors may acquire the TRI data through NTIS and develop systems independently. The estimated user community would be expected to make the TRI data base profitable for vendors who provide user friendly systems and a good complement of other benefits, e.g., analysis and display tools and health and environmental data files.
- o Information gathering from potential providers on capabilities and cost estimates does not provide guaranteed commitments to those projections until actual contractual negotiations are undertaken. The options described in this report are to be considered generic, rather than confirmed agreements.

## **2.4**

### ***EVALUATION CRITERIA***

The evaluation criteria for the proposed options are the overall access requirements and functional requirements as specified by the EPA in the CBD announcement<sup>(9)</sup> and in an OTS memorandum documenting the consensus at a meeting held on June 22, 1987 regarding the public data base.<sup>(12)</sup> The evaluation criteria will form the basis of the comparison analysis in Section 4. The functional characteristics have been grouped according to three levels (tiers) of implementation.<sup>(13)</sup> Tier A is considered by EPA to be the minimum requirements for data base implementation. Tiers B and C are considered to be optional characteristics which are desirable to enhance the use of the system, but may not be required in the initial system. Criteria for all three tiers are evaluated for the potential of any option to meet user and system needs.

#### **2.4.1**

##### ***MINIMUM REQUIREMENTS***

The minimum requirements for the public data base are those mandated by the legislation and those specified by the EPA in interpretation of the legislation or in response to comments from the user community. The minimum requirements (Tier A) are identified as follows:

- o **Hardware and software to accommodate data base load, data storage, indexed retrieval, online and offline report formats.**
- o **Accessible to a large user community.**
- o **Reasonable cost for access. An average cost of \$100 per hour for the use of the data base<sup>(8)</sup> was used as a "reasonable cost" in the ETD demand analysis. This amount includes connect time, telecommunications charges, and all other charges associated with the use of the data base to the user. Printing, training, and documentation charges are not included.**
- o **Aggregation capabilities for standard arithmetic functions of addition, subtraction, multiplication, and division, and statistical functions for average, percentage, maximum, minimum, mean, ratio, standard deviation, and linear regression.**
- o **Accumulation of at least five years of data.**
- o **Search and retrieval based on Boolean logic (i.e., AND, OR, NOT).**
- o **Downloading capabilities for subsets of data to floppy diskette and magnetic disk or tape.**

- o **Account management** to assign and manage user accounts and to adjust access costs for fee waiver.
- o **System availability**, as a minimum to include the daily core hours of 6:00 a.m. to 12:00 midnight EST<sup>(10)</sup> to meet the needs of users across the country.
- o **User support** through hotline inquiry, training, documentation, and search assistance.

Systems that could not meet Tier A requirements were not considered in this evaluation

## 2.4.2 *OPTIONAL CHARACTERISTICS*

Optional characteristics are desirable for the TRI public data base are included in either Tier B or Tier C. The capability of an option to meet the optional characteristics for both tiers is included in the evaluation of each option.

### 2.4.2.1 *Tier B*

The optional characteristics, which will enhance the availability and usage of the public data base when added to Tier A for the second level of implementation, include the following:

- o Access to SAS or other sophisticated statistical analysis software for ad hoc analysis and graphics display capabilities, e.g., to prepare graphs and charts related to release data.
- o Complementary files for health and environmental effects, toxicology, chemical industry data, marketing data (e.g., Dun and Bradstreet files), and geographical data (e.g., longitude and latitude based on ZIP code for off-site releases).

Tier B characteristics are not required for the initial implementation. However, these features are expected to enhance the value of the data base in terms of enhanced data utility and enhanced analysis and display characteristics. The system usage would be expected to expand with the addition of Tier B. Therefore, the evaluation of options in Section 3 includes estimated increases in usage and return of user fees as part of the cost analysis for Tier B.



#### **2.4.2.2 Tier C**

The optional characteristics which in addition to Tiers A and B will enhance the utility of the public data base at the third tier of implementation include the following:

- o Menu driven screens to improve user friendliness.
- o Mapping capabilities to create maps that show releases by geographic area.
- o Crosslinks to other EPA data bases, e.g., data bases relevant to TSCA Inventory updated production data, National Pollutant Discharge Elimination System (NPDES) permits, air releases, etc.

Tier C level of implementation is not a requirement for final TRI implementation. However, it would provide value-added analytical and display characteristics that will be especially valuable for the multi-year data base. The potential ability of each option to meet Tier C is addressed in Section 3.

## **SECTION 3**

### **DATA BASE OPTIONS**

#### **3.1 BACKGROUND**

A definition of the public data base system and identification of the key features of the data base system is necessary to adequately discuss and compare alternatives for public access. Common features expected to be met by all data base options are described in this section as are the criteria for costing an option. Following the general system description, the four data base support methods selected for consideration for the public access data base are discussed and ranked according to specific system characteristics, costs, and benefits. The following options are discussed:

- o An EPA Clearinghouse.
- o Commercial vendors/information providers.
- o An IAG with a government agency provider.
- o An agreement with a non-profit organization information provider.

Appendix C contains a table of the system features for each option, including system features which may not be in the current environment but have been proposed for TRI public data base support.

All discussion of system capabilities and estimated costs in this report is based on interviews and discussions held with technical experts representing each of the various options considered. In all cases, where the analysis is based on the information provided, it is assumed that the best professional estimate was provided, and no attempt to assess the reliability of the data was made on the part of EPA or its contractors.

#### **3.1.1 OVERALL SYSTEM DESCRIPTION**

The public data base will be installed in a hardware/software environment which can accommodate 480 Mb of uncompressed data storage for the first year and, to accommodate five years of multiyear data, estimated at a 120-160% increase for each subsequent year.

The system will be required to meet the Tier A system requirements as a minimum. System development and programming may be required to support the TRI data within a software environment which allows basic search and aggregation capabilities, report generation, and downloading capabilities.

In addition to the minimum requirements necessary to meet legislative mandates, users have identified additional capabilities that will enhance the utility of the TRI data base. The additional capabilities are evaluated for Tier B (i.e., access to SAS or other sophisticated statistical analysis software and interfaces to complementary data files) and for Tier C (i.e., menu-driven queries, mapping software and interfaces with other EPA data bases).

The system will be accessible at least 16 hours per day, 7 days a week and in the core hours of 6 a.m. to 12 midnight EST through major telecommunications networks and WATS line access. A specific hardware/software vendor will not be defined (e.g., IBM versus DEC). However, system access must be enabled through dumb terminals. Subsequent enhancements may require special access equipment, e.g., graphics terminals or personal computers (PCs) for utilization of mapping capabilities or graphic display of sophisticated statistical analyses.

User support will be provided in the form of training sessions, user documentation, inquiry hotline for system and data questions, and search assistance. The plans for the public access data base will include extensive marketing in the form of meetings, brochures, journal advertisement, and convention exhibits, in order to reach the widest audience possible for the system, and thereby to maximize its use.

### *3.1.2 SYSTEM COSTS*

In order to estimate the cost of an option to the EPA, the following costing criteria were evaluated for each option, where applicable:

- o Hardware lease and/or purchase.
- o Software lease and/or purchase.
- o Hardware/Software maintenance fees.
- o Data storage fees.

- o System development.
- o Telecommunications.
- o Marketing.
- o Systems operations and maintenance staff.
- o User support.
- o Account management.
- o Fee Waiver.
- o Miscellaneous costs, such as printer, paper or facility preparation costs.

All four options will require EPA staffing to support the fee waiver account management. Therefore, no additional cost is included for fee waiver account management for any of the options. However, costs for fee waiver usage have been calculated and included for each option to reflect an estimate of 200 fee waiver accounts, according to the ETD demand analysis. Fee waiver costs to the EPA are expected to increase by tier, because usage is expected to increase with improved capabilities.

### **3.1.3      *USER COSTS***

Analysis of the cost to the user is focused on the subscription fees, access fees, and telecommunications fees, where applicable, as provided by each option's technical expert. There are additional cost factors which contribute to user cost but were not evaluated because they are primarily user-driven costs which will vary among users (e.g., offline and online printing, save search, training, etc.).

## **3.2          *EPA CLEARINGHOUSE***

### **3.2.1      *SYSTEM DESCRIPTION***

The EPA Clearinghouse option represents the EPA's establishment of a dedicated computer system and services group to support the public access data base. Preliminary analysis (4) of the EPA Clearinghouse option explored the option of supporting the TRI public access data base within the existing EPA IBM 3090 timeshared environment. However, it was determined that the most effective manner to

support the EPA Clearinghouse environment is with a dedicated IBM 4381 hardware/software environment.

Exhibit 3-1 contains a diagram of the proposed system configuration. Computer operations will be established on an IBM 4381 mainframe and peripheral hardware, purchased or leased solely for the support of the TRI public access, and maintained in the EPA National Computer Center (NCC) facility at Research Triangle Park (RTP), North Carolina. The application system will be a copy of the internal EPA ADABAS TRI system minus the trade secret data. Additional software application development will be necessary to provide some aggregation capabilities and user friendly access to the software developed for the EPA internal TRI data base.

There is no convenient command language, such as is used by most commercial vendors, available within the EPA for usage with the EPA ADABAS system. Therefore, menu driven search and retrieval are expected to be implemented at the Tier A level for this option.

A gateway approach has been taken with this option to accommodate the Tier B characteristic of complementary file enhancement, without the additional expense of loading and maintaining additional files in the EPA Clearinghouse hardware/software environment. With this approach the user must establish his/her own account with a vendor, e.g., The National Library of Medicine (NLM) or Chemical Information System (CIS), if he/she wants to access complementary health and environmental data and other data files, such as MEDLINE, TOXLINE, RTECS, HSDB, AQUIRE, CCRIS, and CHRIS. A one-way communications path will be made available on the EPA Clearinghouse machine. The user will be prompted for a vendor access and password identification. A search argument file established with the TRI system, e.g., a list of CAS Registry numbers, will be transmitted to the vendor machine. In order to continue, the user will remain logged onto both his/her TRI account on the EPA machine and his/her vendor account. The gateway arrangement proposed will not allow transfer of vendor data back to the EPA machine, because two-way communications would entail a more costly networking arrangement and a formal agreement between the vendor and EPA.

# EPA CLEARINGHOUSE

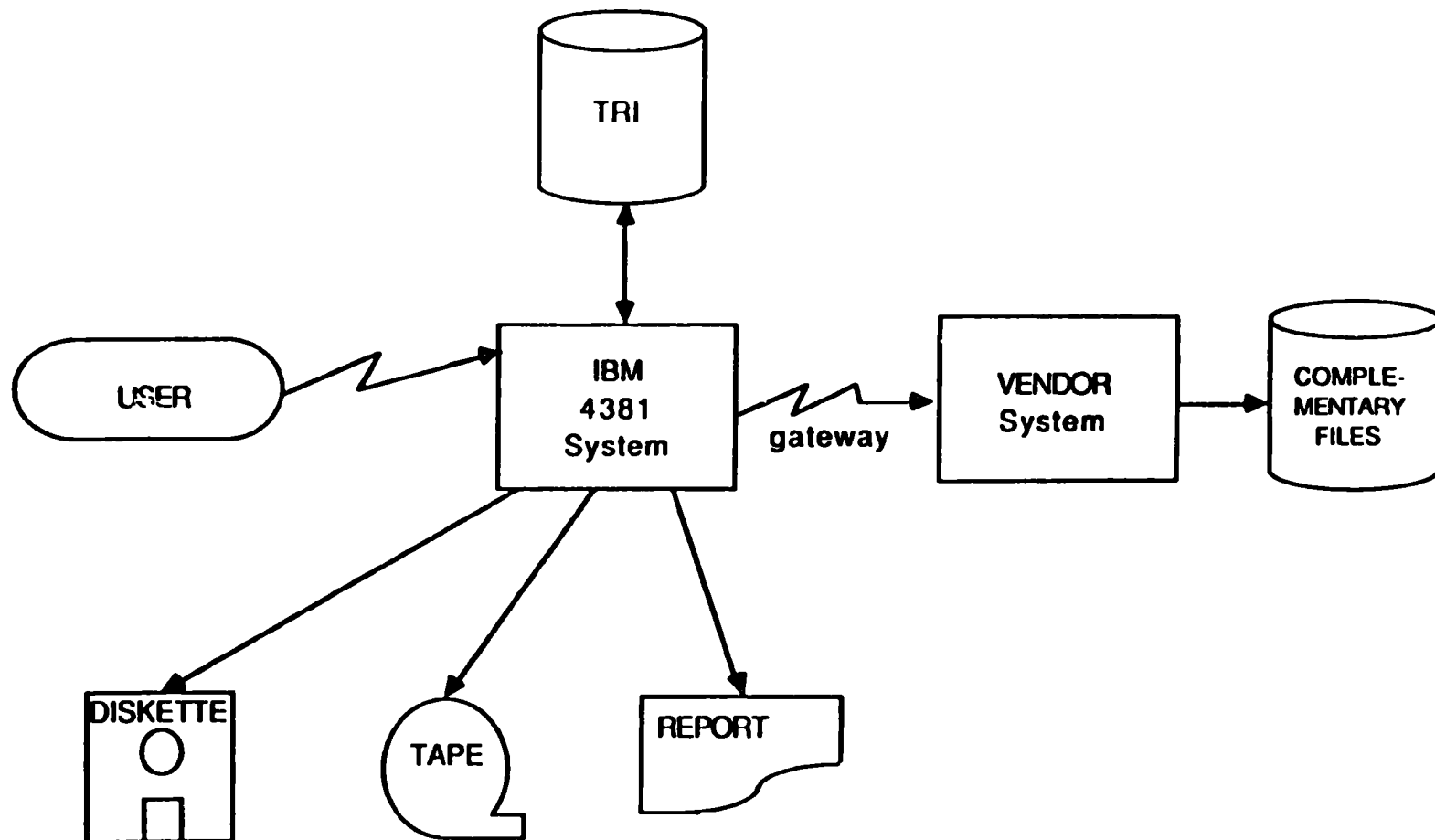


Exhibit 3-1

Tier B statistical analysis will be accommodated with SAS software for statistics, graphic reporting and display. UNIMAP and UNEDIT graphics software packages will be included for interactive contour mapping and modeling under Tier C.

The proposed IBM 4381 hardware configuration will not accommodate the addition of other EPA data bases for the Tier C crosslinked EPA database requirement. In order to accommodate this, it is expected that an IBM 3090 hardware configuration would be necessary. The cost for an IBM 3090 configuration was not calculated. However, it is estimated that it would be an order of magnitude higher than the costs for an IBM 4381 configuration provided in this document.

User support will be provided in the form of a hotline. Personnel to provide search services, operations and maintenance, documentation, training, and marketing will also be the EPA's responsibility.

### 3 2 2 *COSTS*

A summary table of costs is shown in Exhibit 3-2, including initial costs for first year implementation and five year annualized averages. For the EPA Clearinghouse, the upper limit of the range represents the cost to the EPA if no access fees are recovered from the user community, e.g., no usage of the system. The lower end cost of the EPA option is based on the return expected to the government if user fees are collected from the maximum estimated user community based on the demand analysis calculations. It should be noted that the net cost of the EPA Clearinghouse option is very dependent on the system usage and the amount of user fee to be charged. The costs which were used to compile Exhibit 3-2 include:

- o Hardware and software costs which include the purchase/leasing cost by year. Because it is assumed that this computer will be maintained in the NCC facility, under existing EPA contractual agreements, no additional fees are charged for ADABAS, NATURAL, and other software which is site-licensed to that facility. Also included are hardware/software maintenance fees, telecommunications setup and ongoing maintenance, and system development costs for Tier enhancements.

| <u>EPA<br/>Clearinghouse</u>    |                             | <u>Vendor</u>               |           | <u>IAG</u> | <u>Non-Profit<br/>Organization<br/>Information<br/>Provider</u> |
|---------------------------------|-----------------------------|-----------------------------|-----------|------------|---|
|                                 |                             | <u>A</u>                    | <u>B</u>  |            |   |
| First Year Implementation Cost: |                             |                             |           |            |   |
| TIER A                          | \$2,324,220-<br>\$2,684,220 | \$ 493,000-<br>\$1,058,000  | \$315,000 | \$437,686  | \$330,400-<br>\$355,000   |
| TIER B                          | \$2,399,640-<br>\$2,939,640 | \$ 604,000-<br>\$1,169,000  | \$360,000 | \$627,686  | \$370,400-<br>\$395,000   |
| TIER C                          | \$2,337,000-<br>\$3,057,000 | \$1,016,000-<br>\$1,581,000 | \$650,000 | \$742,686  | \$610,000-<br>\$634,600   |
| Five Year Annualized Average:   |                             |                             |           |            |   |
| TIER A                          | \$1,411,684-<br>\$1,771,684 | \$ 173,000-<br>\$2,327,000  | \$366,200 | \$485,401  | \$190,400-<br>\$251,900   |
| TIER B                          | \$1,339,104-<br>\$1,879,104 | \$ 204,000-<br>\$2,358,000  | \$411,200 | \$535,401  | \$230,400-<br>\$291,900   |
| TIER C                          | \$1,226,912-<br>\$1,946,912 | \$ 296,000-<br>\$2,450,000  | \$541,200 | \$570,401  | \$310,000-<br>\$371,500   |

Summary of Cost to Government  
Exhibit 3-2



- o Costs of support staff, including marketing costs which have been estimated for the establishment of a marketing staff and the preparation and presentation of monthly meetings, brochures, and quarterly convention exhibits. User support costs which include training, hotline for inquiries, search assistance, and documentation.
- o Fee waiver costs reflect the annual estimated usage costs incurred from 200 accounts, based upon estimated \$25/hour estimated access fee.(8) Fee waiver costs are expected to increase by tier, because access time is expected to increase as capabilities improve.

### **3.3            *COMMERCIAL VENDORS/INFORMATION PROVIDERS***

#### **3 3 1        *SYSTEM DESCRIPTION***

The commercial vendors/information providers option represents the use of an existing commercial environment to provide access to the TRI public access data base. With this option, the system will be supported in an existing hardware/software environment, with an established support infrastructure for marketing, training, and user support. The TRI data base will be enhanced by the vendor's complementary and crosslink files, as well as through the use of other vendor services.

The offered features and costs for this option are included for two separate vendors to illustrate the variance between the vendors that might be expected. The detailed system characteristics for each of the vendor systems are provided in Appendix C. Commercial vendor systems commonly are implemented with a user friendly command language, as required for Tier A implementation. Vendors interviewed for this analysis indicated that menu-driven query systems are also a viable option for future system enhancements.

The Vendor "A" approach is to develop customized, menu-driven software for analyzing TRI data at various levels of aggregation. In addition to listing user-selected data, the system would produce graphic presentations of facility, company, industry, and locational shares and rankings based upon user-specified quantities and weights. The proposed system would allow users to map combinations of variables to identify geographic patterns and relationships.

### 3.3 2

### *COSTS*

The Vendor "A" approach requires EPA to pay up-front development costs for implementation of a customized search, retrieval, and analysis system. Vendor "A" will supply the access mechanism, the user support, the training, documentation, account management, all without further cost to the EPA, given the expected user community. The Vendor "A" option also includes data storage fees, chargeable to the EPA in the event the user community does not materialize. The Vendor "A" option includes a \$280 initiation fees and \$1,200 of minimum annual usage, neither of which apply to fee-waiver accounts. Computational and other charges are included in the minimum usage fees, but are independent of connect time. Different rates apply to larger users. Vendor "A" stressed that the actual costs and pricing schedules for TRI may vary significantly from those presented here depending on the number and quality of reports submitted under Section 313, and depending on the actual size and character of user demand for online TRI data.

The Vendor "B" approach requires the EPA to provide the TRI data in a compatible format, and to pay some of the aggregation development costs, in addition to the monthly storage costs and the marketing costs. At the Tier C level, EPA would also be requested to pay for the implementation of any menu-driven search and retrieval software. Vendor "B" will provide the telecommunications access and user support, training, and account management, all within the user fees.

The costs for Vendor A support, which were used to compile Exhibit 3-2 reflect:

- o Data storage fees, which will be charged to the EPA only if there is not sufficient usage to offset the costs to the vendor. The storage fee ranges from no cost to EPA for >200 users to highest cost if there are no users. The data storage fees will also increase as the data base grows and crosslinked files are added.
- o Initial load costs, which reflect development costs to be charged to the EPA for system development and internal EPA costs for the EPA to provide the TRI data in an ASCII file format to the Vendor.
- o Additional development costs reflect the development of Tier A aggregation capabilities, customized Tier B health and environmental effects files, and customized software for analysis and mapping of the crosslinked Tier C files.

- o Fee waiver costs reflect vendor charges to the EPA for one large account subscription fee and for initiation and connect-time fees for 200 user codes. These costs are based on large-user prices. They do not include variable costs other than connect-time, such as computational and other charges. Vendor A also indicated that they would absorb account management for FOIA requests as a public service.

The costs to the EPA for Vendor B support reflect:

- o Data storage fees, which will increase as the data base grows, due to multiyear storage. Data storage fees will also increase with the addition of crosslinked EPA files. The Vendor indicated that the fixed storage cost is based upon 500 MB of uncompressed storage.
- o Initial load costs, which reflect the cost for the EPA to convert the TRI data to the file format used by the vendor. This will include systems analysis and programming support by the vendor and the EPA.
- o Additional system development costs, which reflect the Tier A development of aggregation capabilities, Tier C menu-driven access, and development and marketing of a PC-based mapping package for Tier C. The package would be charged to the users as a separate item, on request.
- o Marketing costs, which are included for monthly meetings, brochures, and exhibits.
- o Fee waiver charges, which reflect the estimated 200 user usage fees which will be billed to an EPA account. The vendor indicated that the subscription fee for the fee waiver account will be waived.

A chart showing comparative costs for first year and five year annualized net costs is provided as Exhibit 3-2. For the commercial vendor support, there is a wide range of cost between Vendors A and B. The Vendor A approach was to charge the EPA data storage fees for the system if the user community of 2400 users does not materialize. This accounts for the wide price difference in the Vendor A range. In the case of Vendor B, the premise of the proposed support is that the EPA will pay a fixed level of support, regardless of the user community.

Because Vendor A proposed to charge EPA for storage in the event of non-usage of the public access data base, "best" and "worst" cases have been shown for this vendor which differ by storage cost estimates. While Vendor A indicates that the storage charge will unlikely be necessary, it reflects vendor information. Therefore, costs are shown as gross and net for Vendor A option.

### **3.4 IAG WITH GOVERNMENT AGENCY PROVIDER**

#### **3 4 1 SYSTEM DESCRIPTION**

This option requires an interagency agreement (IAG) with another government agency to share computing resources to provide public access to the TRI data base. The government agency selected for discussion has had several years of experience providing user support, training, and marketing for a user community of primarily health care professionals. Specific features of this option are provided in Appendix C

#### **3 4 2 COSTS**

A table, comparing the costs for each option, was provided as Exhibit 3-2. This option reflects total costs regardless of usage, although the agency indicated the costs may decrease with usage.

Cost for the IAG option are based upon:

- o The additional systems and telecommunications hardware to support the TRI public access data base within the existing configuration. Prices increase by year proportional to the increase in data.
- o Additional staffing requirements to expand the existing user support and systems maintenance staff. Marketing costs were also estimated to expand the current marketing budget for a broader market base.
- o Fee waiver cos. which reflect the charge to the EPA for system usage for 200 estimated user accounts. Fee waiver costs are expected to increase through Tiers B and C because access time is expected to increase as capabilities improve.

The information provider for this option has indicated that the existing system environment is undergoing continuous enhancement. As new system enhancements are made to the entire system, all users and file owners share in these improvements. Hence, many features considered as enhancements for the TRI public access data base may be implemented by the agency and the EPA costs are then minimal with respect to these enhancements, especially menu-driven screens and enhancements to user friendliness.

### **3.5            *AGREEMENT WITH NON-PROFIT ORGANIZATION INFORMATION PROVIDER***

#### **3.5.1        *SYSTEM DESCRIPTION***

This option requires an agreement with a non-profit organization information provider. The basis used for analysis of this option is an existing university provider. A university will provide public access to the TRI from its contractor hardware/software environment. The university will provide the user support, access, training, marketing, and ongoing operations through its own resources. The specific university consulted for this option would plan to load the existing internal TRI data bases in ADABAS/NATURAL and the EPA internal system's applications software onto the university's system. Some additional development, beyond that planned for the EPA internal system and comparable to that for the EPA Clearinghouse Option, will be needed to enhance user friendliness and to provide aggregation capabilities. This option would be expected to use menu-driven software at the Tier A and B levels, because it will be using enhanced EPA software which does not include a convenient command language for search and retrieval.

#### **3.5.2        *COSTS***

A table, comparing the costs for each option, is provided as Exhibit 3-2. Because the organization proposed to charge EPA for storage in the event of non-usage of the public access data base, "best" and "worst" cases have been shown which differ by storage cost estimates.

Costs are discussed below:

- o Data storage fees will be charged to the EPA only when there is little or no system usage to defray these costs. System usage to the level of the maximum demand analysis estimates will defray the costs.
- o System development costs are included for the Tier A and B aggregation capabilities and the Tier C crosslinked file loading and modification of a PC-based mapping package which has already been developed. The mapping package would be charged to the users as a separate item, on request.
- o The existing marketing support for this option is not expected to be adequate for supporting the TRI user community. Therefore, marketing costs have been estimated by the organization to supplement the effort.
- o Fee waiver costs reflect 200 estimated user account subscription fees and usage fees which will be billed to the EPA. The organization indicated that subscription fees will be waived for these accounts. The organization also indicated that there is precedent for establishing different rate structures for managing the fee waiver matter, although this is not considered for this analysis.

### **3.6 EVALUATION OF REQUIREMENTS**

The previous subsections detailed the proposed approach of the information providers to meet the Tier A, B, and C requirements specified in Section 2.3 and the costs associated with each Tier. This section provides an evaluation of the proposed capability of each option to meet user requirements.

Ranking scores were assigned in an attempt to quantify the potential for each option to meet the requirements at each tier of implementation. The ability of each option to meet the requirements are compared on the table in Exhibit 3-3 on a scale of 1-5, with 5 representing the most advantageous option choice, 3 representing average acceptability, and 1 representing least advantageous.

| TIER A:                     | <u>EPA<br/>Clearinghouse</u> | <u>Vendors</u> |          | <u>IAG</u> | <u>Non Profit<br/>Organization<br/>Information<br/>Provider</u> |
|-----------------------------|------------------------------|----------------|----------|------------|---|
|                             |                              | <u>A</u>       | <u>B</u> |            |   |
| 1) Hardware and Software    | 3                            | 4              | 4        | 4          | 4   |
| 2) Accessible               | 4                            | 3              | 3        | 5          | 3   |
| 3) Reasonable Cost to User  | 5                            | 1              | 3        | 5          | 3   |
| 4) Aggregation Capabilities | 3                            | 5              | 5        | 3          | 3   |
| 5) Accumulation of Data     | 3                            | 3              | 3        | 3          | 3   |
| 6) Search and Retrieval     | 2                            | 3              | 4        | 4          | 2   |
| 7) Downloading Capabilities | 3                            | 3              | 3        | 3          | 3   |
| 8) Account Management       | 3                            | 2              | 2        | 2          | 2   |
| 9) System Availability      | 3                            | 5              | 5        | 5          | 5   |
| 10) User Support            | 1                            | 4              | 4        | 4          | 3   |

Ability to Meet Requirements  
Exhibit 3-3  
(Page 1 of 2)

|   | <u>EPA</u><br><u>Clearinghouse</u> | <u>Vendors</u> |           | <u>IAG</u> | <u>Non Profit</u><br><u>Organization</u><br><u>Information</u><br><u>Provider</u> |
|---|------------------------------------|----------------|-----------|------------|---|
| <b>TIER B:</b>  |                                    | <u>A</u>       | <u>B</u>  |            |   |
| 11) Sophisticated<br>Statistical Analysis<br>Software | 3                                  | 3              | 3         | 2          | 3   |
| 12) Complementary Files                               | 1                                  | 4              | 4         | 5          | 2   |
| <b>TIER C:</b>  |                                    |                |           |            |   |
| 13) Menu Driven<br>Access                             | 3                                  | 5              | 3         | 4          | 3   |
| 14) Mapping<br>Capabilities                           | 3                                  | 5              | 2         | 2          | 2   |
| 15) Crosslink to<br>EPA Data Bases                    | <u>1</u>                           | <u>4</u>       | <u>2</u>  | <u>2</u>   | <u>2</u>  |
| Subtotal Tier A:                                      | 30                                 | 33             | 36        | 38         | 32  |
| Subtotal through Tier B:                              | <u>34</u>                          | <u>40</u>      | <u>43</u> | <u>45</u>  | <u>36</u>   |
| Total through Tier C:                                 | 41                                 | 54             | 52        | 53         | 43  |

**Legend:**

5 - Most Advantageous

4 -

3 - Average

2 -

1 - Least Advantageous



### **3.6.1 TIER A**

#### **3.6.1.1 Hardware and Software**

All of the options adequately meet the requirement for hardware and software to accommodate data base load, data storage, indexed retrieval, and online and offline report formats. While the EPA Clearinghouse was ranked at the average acceptability level, the other options were assigned higher advantage due to the existing and proposed environment.

#### **3.6.1.2 Accessible to a Large User Community**

All of the options have the potential to meet the requirement that the TRI public access system be accessible to the user community discussed in Section 1.3 of this document. The EPA Clearinghouse is considered to be more advantageous because it will be specifically designed to meet the needs of the appropriate community. The IAG discussed reaches a large segment of the anticipated user community. The agreement with the non-profit organization information provider and the commercial vendors have a more limited existing user community.

#### **3.6.1.3 Reasonable User Cost**

Exhibit 3-4 contains a chart of the average annual user access cost estimates usage is based upon the ETD demand analysis for system usage.<sup>(8)</sup> The costs to the user presented in this chart are based upon subscription fees, access fees, and telecommunications fees, where applicable. It was assumed that the maximum usage would occur at Tier C implementation. Therefore, an estimate of 12 hours of access time/year/user was used for Tier C estimates with a correspondingly lower usage of 9 and 6 hours for Tiers B and A, respectively.

With the exception of the EPA Clearinghouse, the estimated user access fee decreases with usage, due to the wider distribution of the subscription fees by hour. In the case of the EPA Clearinghouse option, the user access fee increases for Tiers B and C, reflecting the increased cost to the user for gateway access to complementary files, where connect time will be charged by EPA and by the vendor.

|                     | <u>EPA<br/>Clearinghouse</u> | <u>Vendor</u><br>A _____ B       |                                | <u>IAG</u>                   | <u>Non-Profit<br/>Organization<br/>Information<br/>Provider</u> |
|---------------------|------------------------------|----------------------------------|--------------------------------|------------------------------|---|
| TIER A <sup>1</sup> | \$25/hour<br>\$150/year/user | \$243/hour**<br>\$1480/year/user | \$118/hour*<br>\$705/year/user | \$25/hour<br>\$150/year/user | \$84/hour*<br>\$502/year/user                                   |
| TIER B <sup>2</sup> | \$40/hour<br>\$360/year/user | \$164/hour**<br>\$1480/year/user | \$104/hour*<br>\$930/year/user | \$25/hour<br>\$225/year/user | \$78/hour*<br>\$703/year/user                                   |
| TIER C <sup>3</sup> | \$40/hour<br>\$480/year/user | \$123/hour**<br>\$1480/year/user | \$95/hour*<br>\$1140/year/user | \$25/hour<br>\$300/year/user | \$75/hour*<br>\$904/year/user                                   |

<sup>1</sup>Tier A estimate is based upon six hours of access time/year/user.

<sup>2</sup>Tier B estimate is based upon nine hours of access time/year/user.

<sup>3</sup>Tier C estimate is based upon 12 hours of access time/year/user.

\*Includes subscription fee.

\*\* Includes \$280 initiation fees and \$1200 minimum annual use, which may allow for even more than 12 hours of use. Different rates apply to larger users.

Average Annual User Access Cost  
Exhibit 3-4

Using \$100/hour as a maximum acceptable access cost, the EPA Clearinghouse option at \$25-40/hour, and the IAG option at \$25/hour represent good value to the user community. Vendor B costs range from \$118-95, by Tier and the non-profit organization information provider ranges from \$84-75 by Tier, therefore, were assigned acceptable cost scores. However, the usage cost of Vendor A is unacceptable based upon the minimum annual usage and initiation fee requirements.

### *3.1.4 Aggregation Capabilities*

Aggregation capabilities represent standard arithmetic functions of addition, subtraction, multiplication and division, and statistical functions for average, percentage, maximum, minimum, ratio, standard deviation, and linear regression.

The EPA Clearinghouse, IAG, and agreement with non-profit organization information provider options will require system development in order to provide the minimally adequate aggregation capabilities for the data base. The Vendor Systems already provide these capabilities within their existing systems.

### *3.6.1.5 Accumulation of Data*

All of the options have been designed to adequately support the accumulation of at least five years of data at the rate of growth expected, and, therefore, have been assigned average scores.

### *3.6.1.6 Search and Retrieval*

The IAG and the Vendor B currently have command level software which enables TRI searching. Vendor A proposes extensive programming to develop a customized search and retrieval system. Currently no generic, command language software is available for either the non-profit organization information provider or for the EPA Clearinghouse. Therefore, additional software development would be required.

### **3.6.1.7      *Downloading Capabilities***

All options will support download to floppy disk by an online PC user and have been assigned average acceptability. The data base will also be available to the public on magnetic tape through NTIS.

### **3.6.1.8      *Account Management***

Account management to assign and manage user accounts is available for all options. However, for all options, it is assumed for this analysis that access adjustments for fee waiver will be handled by the EPA, and not by individual providers, therefore, all options have been assigned less acceptability than the EPA Clearinghouse option.

### **3.6.1.9      *System Availability***

The EPA Clearinghouse option is designed to meet the minimum core hours of 6:00 a.m. to 12:00 midnight and has been assigned average acceptability. The other options represent 24 hour per day 7 days per week availability and have been assigned most advantageous scores.

### **3.6.1.10     *User Support***

The commercial vendors and the IAG option already have an established infrastructure to support the user community with search assistance, hotlines, documentation, and training; therefore, they have been assigned high scores. The university considered for the non-profit organization information provider option has limited experience in this capacity, however, and is at an average acceptable level.

The EPA Clearinghouse is at a major disadvantage in regard to user support services, because support infrastructures for public data base access do not currently exist within the Agency and will have to be established for this option.

### **3.6.2 TIER B**

#### **3.6.2.1 Sophisticated Statistical Analysis Software**

With the exception of the IAG option, all of the options will have SAS available for statistical analysis as well as graphics report generation enabled by SAS and have been assigned average scores.

Based on the data files, usage and system configurations currently in place for the IAG, it is unlikely that SAS or equivalent statistical software will be available for potential users under that option. Therefore, a PC-based statistical package was proposed for accommodation of the Tier B requirement, at additional development cost to the EPA. This alternative may be available for use, regardless of the option selected.

#### **3.6.2.2 Complementary Files**

Complementary file availability is most advantageous with the IAG option, which currently makes available to users a multitude of files containing toxicity data which may enhance the TRI data utility. The vendors also have an impressive list of complementary files to enhance TRI data utility. These will vary with choice of vendor and are an important consideration when selecting a vendor.

Of least benefit are the EPA Clearinghouse and the agreement with the non-profit organization information provider. The EPA Clearinghouse will make these files available at increased cost to the user through gateway access to a vendor system. Complementary files can be added for the non-profit organization information provider option as specified in the agreement. However, this may increase the cost to the EPA for storage and system development/conversion costs.

### **3.6.3 TIER C**

#### **3.6.3.1 Menu-Driven Access**

Menu-driven access screens to improve user friendliness will be available for the EPA Clearinghouse and for the non-profit organization information provider option, resulting from the EPA TRI internal data base development. These options been assigned average acceptability. Both vendors proposed menu-driven access development

at a cost to EPA. However, Vendor A proposes a very sophisticated capability which is quite user friendly and so has been assigned highest marks. The IAG currently offers menu-driven capabilities and has, therefore, been assigned above average acceptability

#### *3 6 3 2 Mapping*

Mapping capabilities specified for Tier C will be enabled on the EPA Clearinghouse through the UNIMAP and UNIEDIT software on the system. Vendor A has proposed a very sophisticated geographical software development approach as a basis for the TRI public access data base in recognition of the maximal benefit of this approach to the user for TRI data. This vendor has been assigned maximum values for the requirement.

In order to satisfy the mapping enhancement for the other options, a PC-based mapping package will have to be developed and distributed. This would make the enhancement available only to a user with access to a PC and, therefore, decreases the acceptability

#### *3 6 3 3 Crosslink Files*

Lack of crosslink file availability is a disadvantage for all of the options, with the exception of Vendor A who currently has some EPA files available to the user in the vendor's data processing environment and has proposed extensive crosslinked file availability. The other options with the exception of the EPA Clearinghouse Option, have a few files currently available and additional files could be negotiated to be provided for the other commercial vendor, for the IAG, and for the non-profit organization information provider. However, this will increase the storage and system development costs to the EPA. Access to crosslinked files is not accommodated by the EPA Clearinghouse option.

### **3.7                    *COMPARISON OF POTENTIAL RISK***

In addition to the evaluation of costs and the ability to meet the requirements, it is necessary to include a discussion of potential risks involved in the selection of an option, such as:

- o     Timely negotiation.
- o     Timely system implementation.
- o     Long-term availability of the system.

Exhibit 3-5 shows a comparison of the relative risk potential for each option and each is discussed below. The options were compared to each other and ranked on a scale of 1-5 for each risk, where low score indicates high risk potential.

#### **3.7.1                *TIMELY NEGOTIATION***

The timeliness risk reflects the length of time to obtain a formal agreement to support the public access system. Acquisition of resources to manage information requires that agreements be negotiated between the EPA and potential providers. Contract vehicles exist for most of the hardware/software proposed for the EPA Clearinghouse. However, acquisition of equipment, preparation of the facility, and development of a support group will be time consuming and has the potential to delay the negotiation. The wide range of available services and potential costs associated with the private sector will require more time for evaluation and negotiation for agreement with commercial vendors and with the non-profit organization information provider than for negotiating an agreement with a Government Agency. Therefore, the IAG has been determined to have the most acceptable (lowest level) of risk for timely negotiation of an agreement.

#### **3.7.2                *TIMELY SYSTEM IMPLEMENTATION***

The risk associated with system implementation is whether the proposed system can be implemented in a timely manner. Timeliness for implementation is expected to be most favorable for the non-profit organization information provider at the Tier A level, based on the one month implementation schedule expected for the loading and implementation of the ADABAS files after an agreement is signed.

| <u>Risks</u>                    | <u>EPA<br/>Clearinghouse</u> | <u>VENDORS</u> |          | <u>IAG</u> | <u>Non-Profit<br/>Organization<br/>Information<br/>Provider</u> |
|---------------------------------|------------------------------|----------------|----------|------------|---|
|                                 |                              | <u>A</u>       | <u>B</u> |            |   |
| 1) Timely Agreement             | 1                            | 1              | 1        | 4          | 1   |
| 2) Timely System Implementation | 2                            | 3              | 4        | 3          | 4   |
| 3) Long-term Availability       | <u>5</u>                     | <u>1</u>       | <u>1</u> | <u>4</u>   | <u>3</u>  |
| Total                           | 8                            | 5              | 6        | 11         | 8   |

Legend:    5    Lowest Level of Risk  
               4  
               3    Moderate Level of Risk  
               2  
               1    Highest Level of Risk

Comparison of Potential Risk  
 Exhibit 3-5



Favorable risk for system implementation is also anticipated for the IAG and adequate level of risk for commercial Vendor B. These two options represent taking a flat ASCII data file and implementing it in an existing application environment.

On the other hand, the EPA Clearinghouse and Vendor A represent high potential of risk for timely system implementation. While the EPA Clearinghouse, like the non-profit organization information provider is an ADABAS environment, hardware/software acquisition is expected to be timely, as is the establishment of the support environment. The Vendor A option will require the design and implementation of a sophisticated application environment with an estimated six month development and implementation schedule.

### **3.7.3      *LONG TERM AVAILABILITY***

Long term availability of the TRI public access database is considered in regard to the legal requirement for the EPA to keep the public access data base available. Long term system availability is most assured with a dedicated system, i.e., the EPA Clearinghouse. With the other options, ongoing support will be renegotiated at the end of each contract or agreement. The actual system usage will have an impact on this consideration. Non-profit institutions and government agencies are less likely to be affected by marketing conditions that are not favorable. The potential for vendor renewal of a contract is very dependent on system usage as it balances against resource utilization and profits, and so both were rated as high risk.

### **3.8          *COMPARISON OF SYSTEM COSTS***

System costs have been compared by Tier for first year implementation costs and for the estimated annual costs, based on a five-year average. In addition, costs have been compared according to the level of data base usage, where access by a minimal number of users has the potential for significantly increasing the cost to the Government. Exhibit 3-6 provides overall comparison of each option by Tier. The first year cost and ongoing cost to the Government is significantly greater with the EPA Clearinghouse and the Vendor A options, particularly if system usage is realized at the minimum estimated level. Other options, including Vendor B, and IAG, and an agreement with a non-profit

|        | <u>EPA<br/>Clearinghouse</u> | <u>VENDORS</u><br><u>A</u> <u>B</u> |          | <u>IAG</u> | <u>Non-Profit<br/>Organization<br/>Information<br/>Provider</u> |
|--------|------------------------------|-------------------------------------|----------|------------|---|
| Tier A | 1                            | 2                                   | 4        | 4          | 4   |
| Tier B | 1                            | 2                                   | 4        | 3          | 4   |
| Tier C | 1                            | <u>2</u>                            | <u>3</u> | <u>3</u>   | <u>3</u>  |
| Total  | 3                            | 6                                   | 11       | 10         | 11  |

Legend:

|   |               |
|---|---------------|
| 5 | No Cost       |
| 4 |               |
| 3 | Moderate Cost |
| 2 |               |
| 1 | Highest Cost  |

Comparison of System Costs  
Exhibit 3-6

organization information provider, do not require that magnitude of initial investment, and in fact, exhibit relatively low costs for the first year, increasing by year as the need for additional data storage increases.

## **SECTION 4 SUMMARY**

### **4.1           *COMPARISON ANALYSIS***

Section 3 of this document presented the discussion of four alternatives to meet the legislative mandate for a publicly available data base for the TRI data. The discussion included costs, level of ability to satisfy the requirements defined in Section 2, and a discussion of the potential risk in supporting a given option. Exhibit 4-1 compares the options by Tier for the net cost to the Government for the first year and as a five-year annualized average, average annual costs to the user, capability to meet the implementation requirements, and potential for risk. A discussion of each comparison factor follows, with additional consideration for the factors described in Section 2.3 which may affect the system usage and thereby alter the relative costs and benefits.

#### **4.1.1           *COMPARISON OF SYSTEM COSTS***

System costs have been compared by Tier for first year implementation costs and for the estimated annual cost, based on a five-year average. In addition, costs have been compared according to the level of data base usage, where access by a minimal number of users has the potential for significantly increasing the cost to the Government. As is indicated in Exhibit 4-1, the first year cost to the Government is significantly greater with the EPA Clearinghouse and the Vendor A options, particularly if system usage is realized at the minimum estimated level. Other options, including Vendor B, an IAG, and an agreement with a non-profit organization information provider, do not require that magnitude of initial investment, and in fact, exhibit relatively low costs for the first year, increasing as the need for additional data storage increases.

#### **4.1.2           *COMPARISON OF USER COSTS***

The user fees for the EPA Clearinghouse were defined as being equal to the fees of the IAG, to keep the fees in line with the other Government provider. However, the user costs associated with the EPA Clearinghouse are greater than for the IAG at the Tier B and Tier C levels, because the use of a gateway to other information

## COMPARISON OF OPTIONS

|                     | <u>EPA<br/>Clearinghouse</u> | <u>Vendor</u><br><u>A</u> <u>B</u> |          | <u>IAG</u> | <u>Non Profit<br/>Organization<br/>Information<br/>Provider</u> |
|---------------------|------------------------------|------------------------------------|----------|------------|---|
| <b>TIER A</b>       |                              |                                    |          |            |   |
| *Cost to User       | 5                            | 1                                  | 3        | 5          | 3   |
| *Cost to Government | 1                            | 2                                  | 4        | 4          | 4   |
| *Potential Benefits | 25                           | 32                                 | 33       | 33         | 28  |
| **Potential Risk    | <u>8</u>                     | <u>5</u>                           | <u>6</u> | <u>11</u>  | <u>8</u>  |
| Tier A Total:       | 39                           | 40                                 | 46       | 53         | 43  |
| <b>TIER B:</b>      |                              |                                    |          |            |   |
| *Cost to User       | 3                            | 1                                  | 3        | 5          | 3   |
| *Cost to Government | 1                            | 2                                  | 4        | 3          | 4   |
| *Potential Benefits | 29                           | 39                                 | 40       | 40         | 33  |
| **Potential Risk    | <u>8</u>                     | <u>5</u>                           | <u>6</u> | <u>11</u>  | <u>8</u>  |
| Tier B Total:       | 41                           | 47                                 | 53       | 59         | 48  |
| <b>TIER C:</b>      |                              |                                    |          |            |   |
| *Cost to User       | 3                            | 1                                  | 3        | 5          | 3   |
| *Cost to Government | 1                            | 2                                  | 3        | 3          | 3   |
| *Potential Benefits | 36                           | 53                                 | 49       | 48         | 40  |
| **Potential Risk    | <u>8</u>                     | <u>5</u>                           | <u>6</u> | <u>11</u>  | <u>8</u>  |
| Tier C Total:       | 48                           | 61                                 | 61       | 67         | 54  |

\* High score indicates high benefit

\*\* High score indicates less risk

Comparison of Options  
Exhibit 4-1

providers for complementary files requires that the user incur charges simultaneously from the EPA Clearinghouse and from the other information provider while accessing the complementary files

User costs are therefore most favorable to the user with the IAG, followed by the EPA Clearinghouse and the non-profit organization information provider. User costs would be expected to be least favorable with commercial vendors, although reasonable user costs can be realized by this option.

#### *4.1.3 COMPARISON OF CAPABILITY TO MEET REQUIREMENTS*

All options have the potential to meet the basic Tier A requirements. However, the specific user-friendliness of the system, including the support services, would be expected to vary among the options. The EPA Clearinghouse, being dedicated to TRI system support has the potential for meeting Tier A requirements at a high level of satisfaction. However, the long term support through the EPA Clearinghouse option would be expected to be less favorable, due to the lack of convenient access to complementary files at the Tier B level and the lack of availability of other EPA data bases at Tier C.

The agreement with the non-profit organization information provider also has the potential for high satisfaction at Tier A. Continued satisfaction at Tier A, B, and Tier C levels is dependent upon the development of additional systems with appropriate complementary files and other EPA data bases.

The IAG and commercial vendors have the highest potential for meeting requirements at the Tier A, Tier B and C levels.

#### *4.1.4 COMPARISON OF POTENTIAL RISKS*

The potential risks considered for the TRI implementation include timeliness factors, both for negotiating an implementation agreement and for accomplishing the developed system, and long term availability of the system. Acquisition of resources to manage information requires that agreements be negotiated between the EPA and potential providers. Contract vehicles exist for most of the hardware/software proposed for the EPA Clearinghouse. However, acquisition of

equipment, preparation of the facility, and development of a support group will be time consuming and has the potential to delay the implementation. The wide range of available services and potential costs associated with the private sector will require more time for evaluation and negotiation for agreement with commercial vendors than for negotiating an agreement with a Government Agency. With respect to the non-profit organization information provider, uncertainties exist with respect to a mechanism for reaching an agreement.

Timeliness for implementation is expected to be most favorable for the non-profit organization information provider agreement at the Tier A level, because the EPA internal system, with some enhancements, can be loaded on their existing computer. Timely implementation for the IAG will be affected by the need to acquire additional equipment and by software enhancements to meet aggregation requirements.

Long term system availability is more assured with a dedicated system, i.e., the EPA Clearinghouse. Vendors are less likely to provide ongoing support if business appears to be not favorable. Non-profit institutions and government agencies are less likely to be affected by marketing conditions that are not optimum. Therefore minimal risk is associated with the IAG and the non-profit organization information provider options with respect to long-term availability.

#### 4.1.5 SUMMARY OF COMPARISONS

A bar chart is provided as Exhibit 4-2 to illustrate the potential cumulative benefits of the options when equal consideration is given to system costs, costs to the user, conformance to requirements, and least risk for implementation. This chart can best be viewed as providing a relative comparison of the various options. The chart is based on Tier B requirements, because all but EPA have the existing capability for providing complementary files. The costs and risks are based on first year factors, with consideration for the limiting factors described in Section 2.3. The bar at the left illustrates a system where maximum benefits would be achieved. The IAG appears to provide benefits which are closest to optimum.

### CUMULATIVE BENEFITS FOR TIER B AT FIRST YEAR COSTS

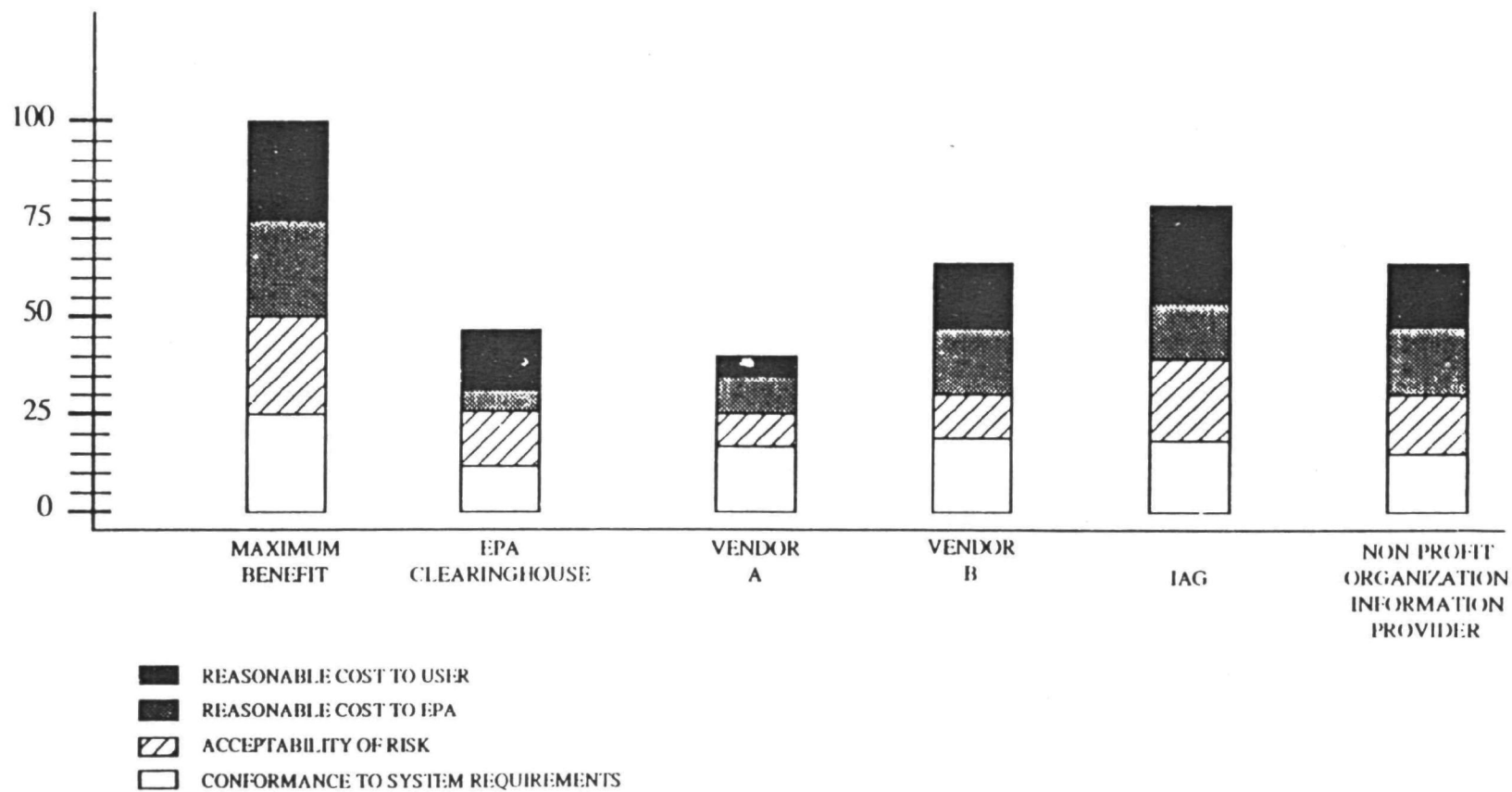


Exhibit 4-2



Prior to selection of an implementation strategy, two essential points must be considered

First, the foregoing analysis needs to be considered in the context of the identified uncertainties concerning the exact user community, the actual amount and quantity of the TRI data, the impact of the "other means" on database use, and possible public preference for addressing questions to the states and EPA regions or to EPA headquarters. While an attempt has been made to look quantitatively at the various options to aid the decision-making process, it must be understood that the degree of certainty implied by the numbers has not been achieved. The numbers do, however, aid in understanding the relative rankings of the various options.

Second, there are strengths and varying degrees of weakness for all the options. For example, the EPA Clearinghouse provides reasonable cost to the user and the potential for long term system availability. On the other hand, weaknesses relate to the fact that this option represents high cost to the government and lacks an existing user support structure for a large user community. It also will require costly and extensive software development to provide access to complementary health and environmental effects files and to crosslink to other EPA files. There is also a significant risk that the EPA Clearinghouse could not be made available within the necessary timeframe. Similar pros and cons are applicable to each of the options.

The task for the decision-maker is to decide which mix of strengths and weaknesses would, on the whole, achieve the public policy objectives for the database. Two possible strategic approaches seem appropriate for consideration, taking into account the factors discussed above, including system capabilities, risks (timeliness, certainty of development), costs to government, and costs to users. These approaches are presented, followed by a recommendation for the preferred option relating to each approach.

The unembellished approach focuses only upon Tier A system requirements, meeting the basic statutory mandate but does not offer significant system enhancements, without significant additional funding. Two options that conceptually fit within this approach are the EPA Clearinghouse and the agreement with the non-profit organization information provider. Of these, the non-profit organization information provider is recommended, assuming that an appropriate mechanism can be found to enter into an agreement. While the possible system would lack many of the "bells and whistles" of the more comprehensive proposals examined, this option represents a moderate risk approach, ranking positively in terms of low cost to government and moderate cost to users. Without enhancements, the database would not achieve maximum ease of use, and the current market coverage of the possible providers is limited. The EPA Clearinghouse option is not recommended both because of its high initial cost and the significant risk that it could not be made available in a timely fashion.

The three options, which fall within this approach, i.e., Vendors A and B and the IAG, call for up-front development and/or planning for the type of basic data base described above, plus a variety of system enhancements that go significantly beyond the statutory requirements and which in various ways add value to the utility of the basic data. These options differ in terms of degree of risk, cost, and particular enhancements offered. Of these options, the IAG option is recommended because it represents the lowest level of overall risk as well as high satisfaction of requirements at a low to moderate cost to EPA. This option also represents a very low cost alternative for the user community.

If the IAG option is chosen, it is recommended that the IAG be initiated at the Tier A level of effort due to the uncertainties mentioned and the funding constraints. With this option, partial "Tier B" capabilities of complementary files will be available to the public at this initial level of support. If the market materializes as expected, then the development of the front end statistics and mapping packages could proceed to the full Tier C level of effort.

### **4.3**

#### ***EPA PLANNED APPROACH***

The EPA planned approach has been to select the enhanced approach, using the IAG, which has been determined to be the option which most closely provides the best balance of all evaluation criteria established for this analysis. The EPA believes that complementary health and environmental effects data files will provide enhanced utility to the data collected under the Community-Right-To-Know legislation. The IAG system implementation, which has immediate access to complementary files of toxicological and other health and safety data, provides the most acceptable solution, given the high level of acceptability for other analysis factors.

The EPA will periodically evaluate and reassess the need to continue support or to enhance the system capabilities, as the uncertainties of data volume and user requirements are resolved and as system usage and utility become established.

In recognition of the importance of ongoing public/private cooperation, the EPA plans to make the private sector aware of the possible opportunities for maximizing the full capabilities of TRI data use by the public. In addition to the data base established for public access to the TRI by the EPA, the public (including commercial vendors) will be encouraged to make use of the magnetic tape of TRI data, available through the NTIS.

**APPENDIX A**  
**BIBLIOGRAPHY**

## BIBLIOGRAPHY

- (1) Proposed Toxic Chemical Release Reporting, Community Right-to-Know Rule, Federal Register, Volume 52, Number 107, June 4, 1987.
- (2) Title III Trade Secret Claims Clarification, Federal Register, Volume 52, Number 199, October 1987
- (3) "Overview of Requirements for the Toxic Chemical Release Inventory," CRC Systems, Inc., September 28, 1987.
- (4) Options for Making the Toxic Chemical Release Inventory (TRI) Database Available to the Public, Draft, January 1987
- (5) "Summaries of Comments from the April 20, 1987 Public Meeting on Toxic Release Inventory (TRI) Database Access," ICF Incorporated, International Square, Washington, D C , April 22, 1987
- (6) Code of Federal Regulations (41 CFR 201), July 1, 1987
- (7) U S Government Systems Review, Draft, August 1987
- (8) Revised Online TRI Data Demand Estimates for Key User Segments, Putnam, Hayes & Bartlett, Inc , December 22, 1987
- (9) "Interactive Systems for Community Right-to-Know Database," Commerce Business Daily, PS4-ps4-9390, July 28, 1987
- (10) "Revised ADP Requirements for the TRI Database," Memorandum from Michele Zenon to Linda Smith
- (11) "Discussion of Fee Policy Issues for the TRI Data Base", Draft, Putnam, Hayes & Bartlett, Inc , 1615 L Street, N W , Washington, D C.
- (12) Cost Estimates for an EPA Clearinghouse, OTS, December 1987.
- (13) "Title III Section 313 Public Database Characteristics," Memorandum from Linda A. Travers to Charles L. Elkins, June 26, 1987.

## **APPENDIX B**

### **GLOSSARY**

## **GLOSSARY**

|                 |  |
|-----------------|--|
| <b>ADP</b>      | <b>Automated Data Processing</b>   |
| <b>AQUIRE</b>   | <b>Aquatic Information Retrieval</b>   |
| <b>ASCII</b>    | <b>American Standard Code for Information Interchange</b>  |
| <b>CBD</b>      | <b>Commerce Business Daily</b>   |
| <b>CCRIS</b>    | <b>Chemical Carcinogenesis Research Information System</b>                                       |
| <b>CEPP</b>     | <b>Chemical Emergency Preparedness Program</b>   |
| <b>CERCLIS</b>  | <b>Comprehensive Environmental Response, Compensation, and Liability Act Information System.</b> |
| <b>CFR</b>      | <b>Code of Federal Regulations</b>   |
| <b>CHEMTREC</b> | <b>Chemical Transportation Emergency Center</b>  |
| <b>CIS</b>      | <b>NIH/EPA Chemicals Information System</b>  |
| <b>CMA</b>      | <b>Chemical Manufacturers Association</b>  |
| <b>CRC</b>      | <b>CRC Systems, Inc.</b>   |
| <b>EDGAR</b>    | <b>Security and Exchange Commission Pilot Data Base</b>  |
| <b>EPA</b>      | <b>Environmental Protection Agency</b>   |
| <b>ERIC</b>     | <b>Educational Resources Information Center</b>  |
| <b>ETD</b>      | <b>Economics and Technology Division</b>   |
| <b>FDPC</b>     | <b>Federal Data Processing Centers</b>   |
| <b>FINDS</b>    | <b>Facilities Index System</b>   |
| <b>FIRMR</b>    | <b>Federal Information Resources Management Regulation</b>                                       |
| <b>FOIA</b>     | <b>Freedom of Information Act</b>  |
| <b>GB</b>       | <b>Gigabyte</b>  |
| <b>HSDB</b>     | <b>Hazardous Substances Data Bank</b>  |
| <b>HWDMs</b>    | <b>Hazardous Waste Data Management System</b>  |

|                |   |
|----------------|---|
| <b>IAG</b>     | <b>Interagency Agreement</b>  |
| <b>IMD</b>     | <b>Information Management Division</b>                                |
| <b>JURIS</b>   | <b>Justice Retrieval and Inquiry System</b>                           |
| <b>MB</b>      | <b>Megabyte</b>   |
| <b>NATICH</b>  | <b>National Air Toxics Information Clearinghouse Data Base</b>        |
| <b>NCC</b>     | <b>National Computer Center</b>                                       |
| <b>NCRIC</b>   | <b>National Chemical Response and Information Center</b>              |
| <b>NRC</b>     | <b>Nuclear Regulatory Commission</b>                                  |
| <b>NIH</b>     | <b>National Institutes of Health</b>                                  |
| <b>NIOSH</b>   | <b>National Institute of Occupational Safety and Health</b>           |
| <b>NLM</b>     | <b>National Library of Medicine</b>                                   |
| <b>NPDES</b>   | <b>National Pollutant Discharge Elimination System</b>                |
| <b>NPIRS</b>   | <b>National Pesticide Information Retrieval System</b>                |
| <b>NPTN</b>    | <b>National Pesticides Telecommunications Network</b>                 |
| <b>NTIS</b>    | <b>National Technical Information Service</b>                         |
| <b>OARM</b>    | <b>Office of Administrative Resource Management</b>                   |
| <b>OTS</b>     | <b>Office of Toxic Substances</b>                                     |
| <b>PC</b>      | <b>Personal Computer</b>  |
| <b>PIPQUIC</b> | <b>Program Integration Project Queries Use in Interactive Command</b> |
| <b>PTO</b>     | <b>Patent and Trademark Office</b>                                    |
| <b>RCRIS</b>   | <b>Resource Conservation and Recovery Information System</b>          |
| <b>RTECS</b>   | <b>Registry of Toxic Effects of Chemical Substances</b>               |
| <b>RTP</b>     | <b>Research Triangle Park</b>   |
| <b>SARA</b>    | <b>Superfund Amendments and Reauthorization Act</b>                   |
| <b>STORET</b>  | <b>Storage and Retrieval of Water Quality Information</b>             |



|             |  |
|-------------|--|
| <b>TSCA</b> | <b>Toxic Substances Control Act</b>    |
| <b>TRI</b>  | <b>Toxic Release Inventory</b>         |
| <b>UTIL</b> | <b>Utility Emission Inventory Data</b> |

**APPENDIX C**  
**OPTION FEATURES**

## OPTION FEATURES

| <u>OPTION</u>              | <u>EPA<br/>Clearinghouse</u>                            | <u>VENDORS</u>                                       |  | <u>IAG</u><br>—   | <u>Non-Profit<br/>Organization<br/>Information<br/>Provider</u> |
|----------------------------|---|--|--|---|---|
|                            |   | <u>A</u>   | <u>B</u>   |   |   |
| <u>Feature</u>             |   |  |  |   |   |
| System Availability        | *16 hrs/day<br>7 days/week                              | 24 hrs/day<br>7 days/week                            | 24 hrs/day<br>7 days/week                              | 24 hrs/day<br>7 days/week                               | 24 hrs/day<br>7 days/week                                       |
| Basic Search Capabilities  | *menu driven  | command language<br>*menu driven                     | command language                                       | command language<br>*menu driven                        | command language<br>*menu driven                                |
| Downloading                | *floppy disk<br>*mag. tape                              | floppy disk<br>mag. tape<br>CD-ROM                   | floppy disk<br>mag. tape                               | floppy disk<br>mag. tape                                | floppy disk<br>mag. tape<br>COM. fiche                          |
| Printer                    | *Online<br>*Offline                                     | Online<br>Offline                                    | Online<br>Offline                                      | Online<br>Offline                                       | Online<br>Offline   |
| Fee Structure with Waivers | *Yes  | *Yes   | No   | No  | *Yes  |
| Sophisticated Statistics   | *Mainframe<br>SAS,                                      | SAS, SPSS  | Mainframe<br>SAS, MLAB                                 | *potential<br>as PC<br>package or<br>on system          | Mainframe<br>SAS  |
| Complementary Files        | *Depends on System<br>being accessed<br>through gateway | *chemical<br>*locational<br>*population<br>*facility | 5 health and<br>environmental<br>effects data<br>bases | 10 health and<br>environmental<br>effects data<br>bases | few health<br>and environmental<br>effects data bases           |

\* Proposed to be developed for TRI public access data base

## OPTION FEATURES

| OPTION                               | EPA<br>Clearinghouse   | VENDORS  |   | IAG   | Non-Profit<br>Organization<br>Information<br>Provider         |
|--------------------------------------|--|--|---|---|---|
|                                      |  | A  | B   |   |   |
| <u>Feature</u>                       |  |  |   |   |   |
| Mapping                              | *Mainframe<br>UNIMAP<br>UNIEDIT                                | *Yes   | *PC System<br>to be developed                       | *PC System<br>to be<br>integrated<br>with TRI                                     | PC System   |
| Crosslink to other<br>EPA Data Bases | None   | Negotiable   | Negotiable  | Negotiable  | Negotiable  |
| User Support                         | *documentation<br>*training sessions<br>*hotline<br>*searching | 6000-6500 users<br>training (weekly)<br>hotline<br>consulting services             | 2500-3000 users<br>documentation<br>online training | 13500 users<br>documentation<br>training(monthly)<br>hotline                      | 500 users<br>online training<br>hotline<br>PC-based tutorials |
| Established Market                   | None   | Industry<br>Government<br>Finance<br>International<br>Academic<br>Legal/Accounting | Industry<br>Government<br>Academic<br>International | Industry<br>Government<br>International<br>Academic<br>Health Care<br>Legal Firms | Industry<br>Government<br>Academic                            |

\* Proposed to be developed for TRI public access data base