See 7/9/86
E. Wick
Transmittal

NBSIR 86-3363 (NRC) (R)

An Analysis of the Requirements for a Computer Assisted Database for Reviews and Evaluations on High Level Waste Data

G. Interrante
C. Messina
S. Harrison
R. Shull
M. Kaufman
U. Bertocci
E. Escalante

U.S. DEPARTMENT OF COMMERCE
National Bureau of Standards
Institute for Materials Science & Engineering
Metallurgy Division
Gaithersburg, MD 20899

June 1986

Sponsored by:
U.S. Nuclear Regulatory Commission (NRC)
Division of Waste Management
Washington, DC 20555

NBSIR 86-3363 (NRC) (R)

AN ANALYSIS OF THE REQUIREMENTS FOR A COMPUTER ASSISTED DATABASE FOR REVIEWS AND EVALUATIONS ON HIGH LEVEL WASTE DATA

C. Interrante
C. Messina
S. Harrison
R. Shull
M. Kaufman
U. Bertocci
E. Escalante

U.S. DEPARTMENT OF COMMERCE
National Bureau of Standards
Institute for Materials Science & Engineering
Metallurgy Division
Gaithersburg, MD 20899

June 1986

Sponsored by:
U.S. Nuclear Regulatory Commission (NRC)
Division of Waste Management
Washington, DC 20555

U.S. DEPARTMENT OF COMMERCE, Malcolm Baldrige, *Secretary*
NATIONAL BUREAU OF STANDARDS, Ernest Ambler, *Director*

An Analysis of the Requirements for a Computer Assisted Database for Reviews and Evaluations on High Level Waste Data

C. Interrante, C. Messina, S. Harrison R. Shull,
M. Kaufman, U. Bertocci, and E. Escalante
National Bureau of Standards
Gaithersburg, MD 20899

DISCLAIMER

The motivation for this report was a request from the Nuclear Regulatory Commission (NRC) for NBS to analyze the hardware, software, and file structure that would be required to automate selected types of data for a specific application. This problem was analyzed at the NBS, in the context of the constraint that the recommended hardware and software would have to be compatible with that in use at the NRC.

Using primarily vendor literature and the experience of NBS workers, several commercially available database management systems (DBMSs) were examined. The DBMS that was found to be most suitable in relation to the requirements was tested to verify its adequacy.

This report identifies various hardware and DBMSs by trade names, as necessary to provide a descriptive characterization of their features and to answer a specific request on costs of the recommended hardware and software. Neither the recommendations nor the inclusion of any item of hardware or software in this report implies a recommendation or endorsement by the National Bureau of Standards for applications other that for which this study was undertaken. Further, the vendors were not asked to verify information for accuracy and clarity, and the recommendations are based on the technical judgement of the authors.

Due to the changing nature of the systems features and the user application environment, the information presented is current only to January 1986.

DISTRIBUTION

This document has been prepared for the use of the Nuclear Regulatory Commission. Responsibility for its further use rests with that agency. NBS requests that if release to the public is contemplated, such action be taken only after consultation with the Technical Information and Publications Division at the National Bureau of Standards.

SUMMARY

The establishment of a computer-assisted database has been initiated, and is suggested for further development and full implementation for storage and retrieval of reviews and evaluations of high level waste (HLW) data, composed principally of pertinent waste-package reports generated by DOE. Initial suggestions by materials scientists at the NBS on the type and form of information to be stored in the database were reviewed by NBS computer scientists, who have concluded that of the available software of packages, Revelation has been determined to be an adequate database management system (DBMS) for this application. Hardware that is compatible with existing NRC equipment is recommended. The format for storage and retrieval of information using Revelation will require further development for full implementation of this DBMS for use with the reviews of reports on the HLW package.

BACKGROUND

In support of the technical responsibilities of the Department of Energy (DOE) and the Nuclear Regulatory Commission (NRC), the NBS has undertaken for the NRC a task of conducting reviews and evaluations of DOE's technical reports on high level waste (HLW). It is expected that this effort will be an ongoing effort that will involve the evaluation of data on materials performance and properties, and that it will result in a substantial body of data that must be gathered and stored in a manner that will permit rapid access for efficient retrieval. As part of this task, this study has been conducted to assess the problems and costs associated with effectively storing this data as a computer retrievable database. An analysis of the requirements of a database for this application was initiated at the NBS after an initial exposure to selected HLW documents. The documents included two memoranda (references 1 and 2) on the "Data Base File Structure" and "HLW Data Base Requirements Analysis."

The above memoranda were reviewed and assessed by materials scientists at the NBS, with the intention of determining whether or not the structure for the database files and the requirements analysis for a HLW database proposed in the memoranda were adequate for the intended purposes. The conclusions of the requirements analysis were not endorsed because the software had not been adequately addressed. The proposed file structure was not endorsed because the NBS workers needed more familiarity with the problem. Further, the NBS workers view the software and the structure of the files as a set of problems that must be addressed together in the development of an effective computer-assisted database. One way to proceed would have been to buy the hardware and then write software to meet the needs dictated by the hardware and the structure of files created for containing the information. This approach is implied in the memoranda and is regarded by NBS workers as likely to be both inefficient and costly, both in the initial implementation and in the later maintenance of the system. Therefore, it would have been recommended only if existing software could not be found.

Constraints established by the sponsor at the outset of this work are described in references 3 through 5. Thus the recommendations given here for hardware and software are based on these constraints, and only a few of them are discussed in the text to follow.

PLANNING A DATABASE

The establishment of a computer database for the HLW application is viewed as a continuous process that involves familiarization with the overall problem and the existing information, and includes determinations of the types and quantity of information that will be stored and the types of queries that will be conducted on the stored information. Together with the form of reports to be prepared these establish the basic requirements of the database system. An additional consideration is what modes of operation are anticipated, as this partially governs hardware requirements. Next, the available software and hardware are studied and selected to meet the requirements. First, the software is chosen, and then the hardware can be specified. It is understood that the hardware must be compatible with existing NRC microcomputers. Further, both flexibility and speed of the database system must be adequate for various modes of operation. The available software packages must be evaluated for their capabilities in relation to the above requirements.

It is the NBS judgment that development of software (from scratch) is far too costly both for initial investment costs and for maintenance through the first two or more years of usage. Therefore, a commercially marketed database management system (DBMS) would be our first choice, provided that one is found that meets the requirements for this application. While a commercial DBMS generally can be expected to require computer professionals to write some software to adapt it initially to the unique needs of the database system, the level of effort is minimized by the judicious selection of a DBMS for a given application. Further, changes in the software required to maintain the database system are much less demanding when a commercial DBMS is used, and this maintenance software frequently can be written by personnel with lower levels of computer expertise.

The hardware requirements include the following: (1) The hardware must be compatible with existing NRC equipment and (2) Candidate software must be capable of being run on the IBM PC (or compatible) hardware.

MODES OF OPERATION

After the hardware is selected and the software is structured to meet the specific needs of the information, the data are entered, stored, and retrieved to complete the first cycle in the process of building a database. In this first cycle, the data will be taken from NBS reviews of selected HLW documents. The initial architecture of this database will be that required by these first reviews. This initial implementation will lead to further understanding of the requirements, and this will result in changes in the structure. Fields will be added, deleted, and changed in name, size, or type. Decisions for major restructuring of the database may even be in order, during the first few cycles. Thus, the software will continue to be modified until its form has been tailored to meet all of the user needs, which themselves probably will not be well understood until the system has been in operation for sometime.

Each NBS review of a document will be actually stored in one or more files, but for simplicity in this discussion, a file will be regarded as containing all stored information for one document. Two types of work on these files are development and production. In the development work, the structure of the database is altered. In the production work, the files of the DBMS are altered only by the data clerk responsible for updating the DBMS.

Reviewers will create the reviews that are modified by others and entered by a data clerk. The reviewer uses either of two media. One is a handwritten form that is entered into the DBMS by the data clerk after it is in final form. The other is a computer on which the reviewer can generate a file that is later merged into the DBMS by the data clerk. The computer may be one of the two PCs recommended for use at the NBS in a HLW Data Center, or it may be a portable unit that may be used by a reviewer in an office, at home, or in another library as needed. While the make and model of the computer best suited for this activity has not been established, it is intended that it should be a PC compatible unit that is highly portable. In the process of creating a review, the reviewer may require numerous searches of the database to answer questions pertinent to the review in question. Thus, another feature of a portable PC compatible computer is that it be capable of storing the files that contain the answers to the reviewers' queries. The 360Kb floppy diskette is the media of choice for this.

The files would be structured and restructured by the manager of the database or by a programmer skilled in information system management. The files are structured to facilitate information retrieval, as reports. Reports are prepared in an agreed upon format, with a report generally being a subset of the data contained in the DBMS, and with the format being one suitable for use by the NRC and its contractors. The stored information must be queried to satisfy a user's request for information. Normally, this query will yield answers for use in decisions made by the users: NRC workers, NBS reviewers, and other NRC contractors.

Index information will be stored in various fields that are accessed by user options, in the normal query. The text fields will be available for "manual" queries, which may be conducted by any user of the database system. The DBMS must support large text fields in a way that permits an optional search of the text. If minicomputers were used, their speed and capacity would likely permit relatively rapid searches of the text contained in the database.

After the appropriate documents have been identified by a query, the query furnishes a list of applicable documents. The user may opt to read the full review of each document or to use the review along with the detailed data contained in the original full document.

DATA STRUCTURE

The preliminary conclusion is that the HLW database should contain five logical entities of information, on each review of a document. One will be an index used for making quick queries of the contents of the document; the second will contain bibliographic information; the third will contain the reviewer's assessment of the contents of the document; the fourth will contain the reviewer's evaluation; and the fifth will contain textual information from other sources, such as an abstract or an executive summary. Searches conducted on any except the first of these entities can be expected to be slow, in general, especially on a microcomputer.

One important conclusion is that within these entities a search will have to be conducted on any of four types of fields: One-string fields, repeating fields, numeric fields, and text-like fields. While some index data can be stored in fields containing only one string, numerous other fields within the five entities will contain a number of words, perhaps more than 200.

Examples of one-string fields include the Identification Number of the document (e.g., CAT001.DOE, or CAT034.MCC), Portion of the Waste Package (e.g., canister, or waste form), and Temperature (e.g., 25C, or 100C). This last item typifies a category that presents a special problem. There may have been measurements taken at numerous temperatures and this would require a different storage and searching system than that which is used for one word, or one string like "waste form."

Examples of repeating fields, those that have multiple entries, follow:

> Material (e.g., alloy steel, 1020 C-Mn steel, or any of a number of ferrous and non-ferrous materials that were tested);
>
> Authors (e.g., Brown, A. W., Avery, R. R., Spike, G. T.)
>
> Environment (e.g., brine, groundwater, and NACE standard H2 solution)

For fields like those for Material, Authors, and Environment, the examples indicate that numerous multiple entries, may be required. These are examples of fields with repeating groups, i.e., the field may be repeated an indefinite number of times.

For a field like the Title of the Document, the information is similar in structure to that which would be given in a text (or comment) field. To search these types of fields requires "substring" search capabilities. A substring can be nearly anything, including one of the items listed in a repeating group.

Thus, for the types of data expected in an HLW database, fields may be one-string, repeating groups, numeric, or text. The DBMS must be able to conduct the query for each of these types of fields, so as to facilitate data entry, retrieval, and reporting. The expectation is that reports generated by the DBMS will require a format that differs significantly from that of the data entries.

One concern of NBS computer scientists who reviewed the requirements as set forth by the materials scientists is that DBMSs designed for microcomputer applications usually are designed to handle only the "one-string" type of data entry, and this would have proven to be very inefficient for the HLW database application.

Detailed numerical tables of data are believed to be unwarranted in the database itself. Other numerical data, such as temperature of test, temperature range over which the data may be applicable, pressure, various mechanical and physical properties of the material tested, etc., would be included in the index. Almost all of the index fields would not be numeric entries. These would include: Title, keywords, materials, environmental and other test conditions.

LIMITATIONS OF MICROCOMPUTERS

The NRC requires that the proposed HLW database should be compatible with their existing microcomputers. This requirement excludes from consideration existing software developed for minicomputers, as a portability problem would be associated with the mixing of mini- and microcomputers, and this problem would be costly to overcome. While this portability problem is not insurmountable, we regard the use of minicomputers for this application to be ill advised, provided that suitable software can be obtained for microcomputers. More importantly, in this application, a principal benefit of suitable commerically available software for microcomputers is related to portability of the entire database, the structure of the files and the data. On the other hand, with the minicomputer, portability problems could be very costly to solve.

In general, the software developed for microcomputers is preferred for this application, at least in the sense that it is often more user friendly than that developed for either the minicomputer or the mainframe computer. User friendliness translates into lower costs for initially establishing and for later maintaining the database system, with the software costs being kept to a minimum if the microcomputer can be used to do all or most of the storing and processing of the data.

DBMS SOFTWARE for the PC

The database management systems (DBMSs) developed for microcomputers have been predominately modeled after the relational database. Traditionally, the DBMS developed for the microcomputer has been designed to handle the one-string and numeric types of data entries, e.g., dBase III. However, some of the more advanced systems, such as Revelation, can also be used for the two other types of fields described above (repeating, variable length).

Software for use on minicomputers and mainframe computers have traditionally been developed to process information stored in "hierarchical" and "network" structures. The structure of either the hierarchical or the network database has the following characteristics: (1) It permits very fast searches, (2) it is relatively difficult to update while maintaining the integrity of the file, and (3) it requires a high level of skill to generate reports. Thus, the design of the database is determined by knowledge of how the database is to be used; and software, for large computers have better features, especially speed and the capabilities to handle various types of fields. For this application the disadvantages of mini- and mainframe computers are the costs of the software and its maintenance. In addition, both the initial and the maintenance costs for the hardware are generally higher than for microcomputers.

With the relational approach, the data are stored as tables that include a unique identifier for each row of the table. For example, each reviewed document will have a unique designation that is carried along in each of the files on that document. These identifiers facilitate the search process. The architecture of the files within the DBMS must be designed to access those data in a manner consistent with the intended usage. The best known relational software packages were written for microcomputers.

The Fixed-Field DBMS

In this type of DBMS, the amount of space required for storage equals the amount of space that is searched. Data are stored in fields that have their length fixed at a value corresponding to the maximum size that may be required for any single entry made in that particular field. For some fields, this can be a large field size, and it can contrast sharply with the mean length of an entry for the particular field, i.e. it may be much larger than the actual amount of information that is usually stored in that field. Because this anticipated length must be searched every time the field is searched, both the search time and the storage capacities are affected adversely when the fixed-field DBMS is used.

In most current DBMSs, including dBASE III which is the only DBMS listed (Ref.1) in the NRC Standard PC Software, the data are divided into fields (columns) of predetermined size. The largest field that dBASE III can hold is 4000 bytes, which is about one full page of text. A change in the size or position of a fixed-sized field requires an extensive knowledge of the DBMS and is not done lightly. It requires the supervision of the database manager, and the problems that might be encountered include (1) previously stored information may be lost, (2) available storage space may be exceeded, (3) forms for output must be altered to describe the new form of the database, etc. There are other problems with the use of fixed fields, as described below.

The imposition of limits on the size of the fields for either the keyword or comment fields would place unwanted constraints on the editors and it would complicate data entry.

When a predetermined space has been reserved for fields that normally would contain various lengths of information, such as for keywords, authors, or comments, a good part of the field will normally contain blanks, which are the spaces needed for the largest known entry. The presence of these blanks uses both RAM and disk space. Further, in the fixed-field system, when a data entry is larger than the reserved space for that field, one is very tempted to truncate, abbreviate, or omit information instead of restructuring the DBMS. Thus, the choices made by the builder are compromises that consider the available memory, the database requirements, and restructuring.

Thus, a fixed field size in a DBMS lends itself to very fast mathematical search schemes, but it complicates both the structure of the search language and the formats for generating of reports.

The Variable-Length Field DBMS

Some DBMSs have a free form, that is the lengths of the fields are variable and dependent only on the amount of information stored in the field. Thus, information is packed into this type of DBMS and there are no empty spaces in fixed-field DBMSs. Four commonly used free-form DBMSs are Dayflo, Oracle, Revelation, and Sci-mate. Revelation will be used as our prime example in the discussion to follow.

When compared with the above fixed-field types of DBMS, systems developed for fields of data containing strings of information of indeterminate length follow a different logic of storage. Examples of the latter, which is sometimes called free form, are the commercially available Revelation and ABCUP, developed at NBS on an HP-1000 mini computer.

Instead of a table of fixed size, the information is stored as books would be on library shelves. An entry is stored in its entirety and an empty entry takes up no space. Therefore, in these systems, no space is wasted on null information and the entire RAM and disk are utilized for meaningful data. Data restructuring is normally accomplished simply by adding new fields when desired. These fields are not limited in size or by any characteristics, except by the intended limitation dictated by the type of data to be stored.

When compared with numbers, strings of characters are more difficult to input and proofread. Therefore, for string data, computer assistance is needed to keep strict control on the contents of a field. Revelation can be instructed to (1) parse (check at the time of data entry) for specific "allowed" values within specified fields; (2) insert default values, (3) match the input against a predetermined set of acceptable strings. Further it allows development of BASIC code to mathematically manipulate data. For example, a code could be written to check validity of input or to take a unit [degrees F], convert it to another unit [degrees K], and store the conversion in a field used for searches. These examples of the advantages of some of the nonfixed-field DBMSs, provide for operations that are very difficult or impossible using a fixed-field DBMS, such as dBASE III.

Although the internal structure of Revelation makes it inherently slower for searches when compared with dBASE III, the difference in search speed could be minimized by thoughtful design of the fields and files making up the HLW database.

An indepth comparison of dBASE III vs Revelation appears in reference 6.

SIZE OF THE PROPOSED DATABASE

Each review of a report will be stored in a separate record. It is estimated that each record would contain no more than about 5 packed pages or about 25,000 characters. It is anticipated that about 1000 records will be stored over a period of about 5 years. This could generate as much as 25 Megabytes of data. This database size assumes that the variable-length-field concept will be used to store the data. If fixed-field DBMSs were to be used for this same task, a much greater memory storage capacity would be required.

ESTIMATES ON NBS USAGE

The database would involve activities of about a dozen workers: A database manager, a data clerk, and about 10 reviewers. The time that each worker will devote to interaction with the database is actually indeterminate and it is a strong function of time over the next five year period. Nevertheless, estimates for the immediate future are given below for NBS users only:

	time in hours per week
Development work by System Analyst--	0 to 30, ave. 15
Production work by Data Clerk--	20 to 40, ave. 25
Reviewers--	10 to 50, ave. 25

In addition to these NBS users, there are others who would use the data. Only NBS workers will alter the files. These will be the systems analyst who restructures the database and the database manager and clerk who enter data. Those who would only access the files include NBS users (reviewers, database manager and program manager), NRC Workers, NRC Contractors, and other users.

File alterations (development and production) would be conducted on a PC designated as a "server". This unit would have an internal hard disk for primary storage and an external storage media would back up these data files. A second PC, designated as a user unit, would serve the needs of reviewers primarily. Although this second unit is independant of the server, it could be and would be used, as needed, for the development and production operations. For example, the system analyst could work at it even while production work is being done by the data clerk.

The user unit will normally be for frequent and rapid access to the information in the files. This is done mainly by reviewers. It is anticipated that this unit will be set up in a multiuser mode, so that nearby and remote PC's, one of which already exist in the reviewer group, can be used to access these files. The printer for the second unit will serve two purposes and it need not have the quality of the primary printer used for formal reports. This unit will be used to obtain quickly paper copies for users and to assure compatibility with the NRC recommended hardware.

This compatibility issue is particularly important. It must be assured that printers, on- and off-site, will print the identical special characters (and super- and subscripts). The two units will be connected in a mode best suited for rapid communications. This is to facilitate transfer of data, especially that from the server to the user. The reviewers' file must be constantly kept current. After the system has been in operation and procedures have been developed for porting to a portable unit (discussed earlier), the user unit would be used by reviewers to create subsets of the database for use in the portable unit at remote locations.

SOFTWARE SELECTION

A list of 16 requirements were reviewed in the process of selection and these are given in Appendix I titled Software Selection. Software selection was based in part on the above considerations. Candidate DBMSs for this application were selected by screening a large list of commercial DBMSs that have been reviewed in "Project Database" which is published in PC Magazine. The issues examined are dated June 12 and 24, August 7 and 21, and September 4 and 18. All six are part of Volume 3 (Ref. 6). Features and restrictions (such as limits on record sizes and numbers of fields) for each listed DBMS were reviewed in relation to the requirements of this application. Two criterion used in reviewing DBMSs are: 1. The quantity and quality of searchable data is much more important than the speed of search, and 2. the database should have a free-form structure to accomadate a high expected variance in field length.

Thus, four free-form DBMSs were selected for further review to determine if any of them satisfied requirements established for this application. The result of this review of four DBMSs that use variable-length fields is the recommendation that Revelation is the most suitable DBMS for this application.

Other DBMS candidates considered include Dayflo, Oracle, and Sci-mate. All four were evaluated in terms of the requirements set forth in the Appendix. Revelation is currently being used at the NBS, on a Compaq, in the Occupational and Health Safety Division.

Although the requirements of this database have been evaluated by NBS, as set forth in this document, it is expected that they will be firmly established only after the suggested system has been in operation for some time. Some of these requirements are based on the expected modes of operation.

HARDWARE SELECTION

One of the principal guidelines set forth for this HLW database is that the system must be compatible with the IBM PC, IBM PC-XT, and Compaq microcomputers, which are currently in use and being purchased by the NRC. It is noted that while minicomputers are more powerful and while they may have significant advantages for this application in terms of available software, capacity, and speed, the problem of interfacing them to the existing NRC microcomputers could prove to be costly. Further, the initial investment in minicomputers would be higher as well. With this in mind the discussion that follows on microcomputers was prepared.

Comparison of the IBM PC-AT and PC-XT

The IBM PC-AT and IBM PC-XT are prime candidates for this application because of their compatibility with the current NRC hardware. The AT has the potential for much greater speed of processing and therefore shorter times may be required to make complex queries of the database. Revelation has been written to take advantage of the higher-speed chip that is resident in the AT. Further, code written for the XT can be used on the AT at a speed at least equal to that obtained with the XT, as the AT can be operated using the disk operating system (DOS) of the XT. Speed will be a most important consideration as the database grows and the number of records is increased to levels that require significant times for obtaining an answer to a query.

Basically, the AT has the more modern architecture developed for the PC line of microcomputers; an investment in it could also be warranted on this basis, as new developments of enhanced capabilities are much more likely to be serving the AT.

Recommendations

Clearly, either an AT, or an XT, or a PC with an attached disk drive (see minimum configuration in Appendix I) can be used for accessing the database. In each case, the information would be developed on the same operating system and there is no question of compatibility or transferability. Speed is the principal advantage of the AT, but other considerations mentioned above are very important to the practical usefulness of the database depending on the frequency of usage and the types of queries to be conducted. The initial implementation of this work has been completed using an IBM PC-AT.

Based on this experience and a cost analysis that is presented in the Appendix and discussed briefly below, it is recommended that two PC-AT units that are virtually identical should be purchased for this work. The usage estimates given above show that these units will not be idle and we estimate that the time saved by using the AT, which has a faster processor, will be justified in productivity gains of the workers, and other justifications are cited elsewhere in this report. The cost difference is only about $1,480 over that of the comparable XT system. There is no advantage to making one of these units different from the other; in fact, there are a number of disadvantages that naturally arise from the fact that a number of time-intensive users will want to use either machine without regard to questions of speed, mode of operation, or keyboard ergonomics.

The printers for the two units are different for reasons discussed above, one is for formal reports and the other is for users. The portable computer discussed above has not been selected, and it is recommended that this selection be made after the system has been in operation and time is available to determine which models will best serve the intended functions with a minimum in software development costs, if any.

The figures quoted below for the high usage sites include costs of all anticipated hardware required to make the system fully operational and they include the cost of a Bernoulli 20+ external storage unit. Two of these units are included in the recommendations given below, with the intention of using them as the primary transport media that will be used for keeping the files at the NRC current. The intention is to use one at the NBS for both backup and transporting, and the other at the NRC, where it could be used to access the data even when only the minimum recommended system is available. Transport of the entire database will be made using Bernoulli disks. Transport of occasional (low-volume) communications and networking operations can be conducted using a modem.

It is understood that transportability of the entire database within the NRC is desired. This requirement leads to the recommendation of the 20+ Bernoulli. If this were not a requirement, less costly alternatives to the 20+ Bernoulli could be recommended. The 20+ Bernoulli could be used (at NRC) either to upload the entire database to an internal hard disk or to support a PC having the minimum configuration.

A document reader will be used to reduce costs of data entry from non-PC sources. By judicious selection of a multiplicity of fonts the reader will permit rapid entry of abstracts, etc. directly from technical articles, and from any review forms that are not prepared on a PC.

COST ESTIMATES

Under Hardware Recommendations in Appendix I, costs are given for the recommended PC hardware that has been determined to be suitable for use with Revelation in this application. Three systems are priced. One is the cost of the AT systems recommended for high usage sites, with the Quietwriter printer ($10,699) and with an Epson printer ($10,102). The second is an XT system recommended for low-usage sites ($8622), and the third is a minimum recommendation configuration for low-usage sites ($3676).

If the first two are compared using the same printer, the difference in cost between an AT and an XT configuration can be obtained, and this value is $1480. It is given here as a reference for use in the recommendations given above.

A summary of the estimated costs for the hardware for the proposed computer assisted databsase for reviews and evaluations on high-level waste data is given in Table 1. The estimated costs of the recommended vender supplied software are given in Table 2. The tables indicate that hardware will cost $39,929 and the software $2005, for a total of approximately $42,000. It is expected that the additional cost required to fully develop and operate the database will be about $50,000 per year greater than the cost required to present the reviews in hard copy form only, i.e. with no database storage or retrieval of the reviews. This cost will diminish by about a factor of two after the first two or three years of operation, as they would include mainly the solution to software and hardware problems.

IMPLEMENTATION OF REVELATION FOR A HLW DATABASE

The proposed database has been implemented for the bibliographic parts of the information, using the DBMS Revelation. A list of data elements for the bibliographic file and their descriptions are given in Appendix IIC. In addition, the file structure for the entire database has been designed. Guide forms have been created to assist reviewers, to minimize ambiguity in the instructions, and to facilitate the establishment of a system for development of keyword lists (see Appendix IIA and IIB). The keyword lists will be useful in conducting a limited but rapid search of extensive data files in the proposed database. The various forms developed for use with this database, include the completed review forms (Appendix IID), and an output of the bibliographic file (Appendix IIE). This output is in raw form intended only to demonstrate the status of the work and is not intended as the final form for reporting.

Using Revelation, a user may search for any alphanumeric character string, in any field regardless of the size of the field. Thus, Abstract, General Comment, Type of Data, fields, etc., can be searched for information not in the keyword lists.

ACKNOWLEDGMENT

We gratefully acknowledge the support and assistance rendered by Mr. Everett Wick and Mr. John Voglewede of the Materials Engineering Section, Engineering Branch, Division of Waste Management of the Nuclear Regulatory Commission. Special thanks are due for Charles L. Sheppard of the Institute for Computer Sciences and Technology for his guidance and counsel in this work.

REFERENCES

1. T. Sullivan (BNL) memorandum to E.S. Wick (NRC) on "Database File Structure" dated September 20, 1984.

2. T. Sullivan (BNL) memorandum to E.S. Wick (NRC) on "HLW Data Base Requirements Analysis" dated Nov. 14, 1984.

3. H.S. Bassett (NRC) memorandum for Branch Chiefs and above on "Use of Microcomputers by NRC Contractors" dated December 13, 1984.

4. J.A. Shields (NRC) memorandum for NRC computer users on "PC Policy and Procedures" dated December 20, 1984.

5. M.R. Knapp (NRC) memorandum for R.E. Browning (NRC) on "Quality Assurance Plan for Software-OPS Plan Commitment 5212339" dated September 26, 1984.

6. "Project Data Base", PC Magazine, vol.3, No. 17, September 4, 1984.

Table 1

Summary of Costs of Recommended Hardware

PC-AT Table 1, the Appendix configuration

1 each with IBM Quietwriter printer	$10,699.
1 each with Epsom XF-85 printer	10,102.
Document reader (estimate)	13,500.
Portable computer (estimate)	4,500.
Six each extra Bernoulli cartridges	398.
100 each floppy disks - 360 Kb	200.
1.2 Mb	530.
Total	$39,929.

Table 2

Summary of Costs for Vendor Supplied Items

Software from Revelation - Estimated Software Costs	
Version G.2 of Revelation (for server)	$ 950.
Runtime (for use at NRC)	200.
Runtime (for use at NBS)	200.
Network Revelation (optional)	
Bump disk for 4 users at NBS	495.
Subtotal	$ 1,845.
Operating System from IBM	
DOS 3.1 (2 each)	$ 160.
Total	$ 2,005.

APPENDIX I

REQUIREMENTS ANALYSIS: The Second Phase of the Software Evaluation and Recommendations for Software and Hardware

REQUIREMENT ANALYSIS

The selected database management system for the NRC database must run on an IBM PC or compatible, and it must support:

- o variable length fields,
- o field lengths of 7000 characters,
- o record lengths of 22,047 characters,
- o a minimum of 44 fields,
- o repeating fields,
- o multiple large text fields,
- o indexing,
- o searching large text fields,
- o report generating,
- o data dumping,
- o bulk loading,
- o integrity checking,
- o structure modification,
- o a minimum of 1000 records,
- o menu building, and
- o programming.

The numbers in the above features are based on an initial set of data elements that are proposed for calculation purposes only. See the appendix for a listing.

Variable Length Fields. Assignment of variable length fields is essential for the conservation of storage space, which is at a premium on microcomputers. Variable length storage is required both for fields with specified upper limits and for those without. That is, string values that are assigned to fixed length field types should not be padded with spaces up to the maximum allowed character length.

Field Lengths of 7000 Characters. Each record in the NRC database will contain the "review and evaluation" information for a particular document. For this reason, an assigned field value may contain as much as one-half to one and one-half pages of text. Examples are the assignment of abstract or comment fields.

Record Lengths of 22,047 Characters. The sum of the maximum possible field lengths in the initial record structure for the NRC database total 22,047 characters. This number of characters is the minimum requirement of any selected data management software.

A Minimum of 44 Fields. The proposed structure for the NRC database identifies 44 fields. Thus, any selected data management software must allow defining a minimum of 44 fields.

Repeating Fields. Some fields for a given record occurrence (in the NRC database) will have more than one assignable value. These are repeating group fields which will require a special definition language. The language must be able to handle a variable number of assignable values to a given field.

Multiple Large Text Fields. The initial structure for the proposed NRC database specifies that several large text fields may be included in the record for any given document. Both the abstract and several different comments about a document will be assigned as separate fields.

Indexing. To speed up searching, a data management package will be required to have an indexing capability. This capability will prove to be increasingly beneficial as the database grows, because search time increases greatly with the number of records searched.

Searching Large Text Fields. There are also those cases when a search constraint may specify a field that is not indexed or that is not a key. This will be true during the early developments of the NRC database because comments and abstracts may include candidate keywords that were not evident initially. Searching record occurrences in either the abstract or comment fields will require the selected database management system to possess a powerful substring searching capability.

Report Generating. Reproduction of reviews and evaluations will be necessary, and in some cases, special reports will be generated. The report generator may require an interface with the database through a high-level programming language. Thus, the selected database management system should allow interfacing through a high-level programming language.

Data Dumping. The operations of a database at user sites will require periodic updates from the central office and perhaps comments to the central office. This transfer will require the selected database management system to possess both an import and export facility that can handle the necessary data transfers between locations.

Bulk Loading. The initial population of the NRC database will require extensive typing by clerical personnel unless the selected database management system can adequately handle bulk loading of any existing data. The database management system should possess a loading facility that can accept a plain ASCII file with clearly specified field separators.

Integrity Checking. Upon entering data, the data management software must support the option to: (1) test the format of the entered data, and (2) test the entered data against a specified set of values. Such options will ensure the integrity of critical data.

Structure Modification. The structure of the fields for the proposed NRC database is expected to change periodically in the initial development stages. Thus, the selected data management package should allow structural modifications with a minimum amount of difficulty.

A Minimum of 1000 Records. It is projected that the proposed NRC database will contain record information on at least 1000 documents over a period of about 5 or so years. Storing 1000 records is not a limitation for most data management software; however, limitation comes into play when those records may contain as many as 22,047 characters.

Menu Building. As mentioned previously, clerical help may be required to enter the data for record occurrences. This will require a system that is both simple and easy to use. However, the selected database management system should lend itself both to the development of menus and to the flexibility or ease of alterating existing menus. Such a menu building capability will allow tailoring the capabilities of the system's data manager to a specific application.

Programming. In some cases, it will be necessary to augment the capabilities of the system's data manager with those of a programming language. A programming language will allow the development of controls for the user/interface procedures that cannot be readily achieved otherwise.

EVALUATION PROCESS

The evaluation process to select a database management system for the NRC database involves two phases. In phase one, we reviewed the manuals of each candidate DBMS. The reviews entailed analyzing the features of a DBMS package using the criteria discussed in the requirement analysis section. If analysis revealed the absence of any features that are critical to the development of the proposed NRC database, the DBMS was dropped as a candidate and no further consideration was given to it.

In the second phase of our evaluation process, we will implement a version of the NRC database file structure and perform testing on a subset of the records that will populate the database. The structure of the database will vary from one DBMS package to the other. This fact will be noted in any performance testing. The vàriation is required because the database must be structured to the form required by the structure of the database management system.

EVALUATION RESULTS

Each of four software packages considered is capable of running on hardware that is compatible with the IBM PC family. The software under consideration include DayFlo, ORACLE, Revelation, and Sci-mate. Evaluating the features of each software package resulted in the table on the following page.

DAYFLO

Variable Length Fields. Specification of field lengths is not required. Fields are limited only by the maximum record length.

Maximum Field Length. A field may contain a maximum of

DBMS EVALUATION RESULTS

Feature \ Software	Dayflo	Oracle	Revelation	Sci-mate
Variable Length Field	yes	yes	yes	yes
Field Length of 7000 Characters	yes	yes	yes	no
Record Lengths of 22,047 Characters	yes	yes	yes	no
A Minimum of 44 Fields	yes	yes	yes	no
Repeating Fields	yes	no	yes	no
Multiple Large Text Fields	yes	no	yes	no
Indexing	yes	yes	yes	no
Searching Large Text Fields	yes	no	yes	no
Report Generating	yes	yes	yes	yes
Data Dumping	yes	yes	yes	no
Bulk Loading	yes	yes	yes	yes
Integrity Checking	yes	yes	yes	no
Structure Modification	yes	yes	yes	yes
A minimum of 1000 Records	yes	yes	yes	yes
Menu Building	no	yes	yes	no
Programming	no	yes	yes	no

32,000 characters. This is well over the required 7000 characters for the NRC database.

Maximum Record Length. Records may contain up to 32,000 characters. This length satisfies the 22,046 characters requirement for a record in the NRC database.

Maximum Number of Fields. The maximum number of fields is limited only by the 32,000 characters that are allowed per record.

Repeating Fields. Multivalued fields are allowed. However, there is no way to group repeating values across different repeating fields.

Multiple Large Text Fields. The number of large text fields is limited only by the maximum number of characters per record.

Indexing. Indexing is maintained through field names. A maximum of 32 indexes is allowed per record.

Searching Large Text Fields. Searching may be performed on text fields of any size (e.g. 7000 characters) through substring searching capabilities. However, large text fields can not fully utilize the indexing capabilities; therefore, searches will be performed sequentially and much more slowly than on a fully indexed field.

Report Generating. Document formats can be specified easily. These specifications identify which fields are to be printed and their page locations. However, reports are limited to the 2,500 records resident in the work space.

Data Dumping. Records may be dumped as an ASCII file. Field values may or may not be tagged by their respective field names.

Bulk Loading. Records can be loaded from an ASCII file. However, field values are not assigned to respective fields. This assignment has to be performed by the operator.

Integrity Checking. Integrity checking is performed through pattern matching. A specified pattern controls the string values that can be assigned to fields. If a string value does not match the specified pattern, it is not store as a field value.

Structure Modification. Structure modifications can be made easily. Modifications are made directly to the dictionary information (i.e., field size, field name, field addition and field deletion). The modification process is completely menu driven.

Maximum Number of Records. A database may hold up to 65,000 records or pages of information. Of this number, only 2,500 are available at any time for a designed worked area.

Menu Building. Not supported.

Programming. Not supported.

ORACLE

Variable Length Fields. Specification of field lengths is required except in the case of LONG data types. However, in all cases, space is allotted only as required by string values.

Maximum Field Length. A field may contain a maximum of 65,536 characters. This is well over the required 7000 characters for the NRC database.

Maximum Record Length. Records may contain a maximum of 125,256 characters. This length satisfies the 22,046 characters requirement for a record in the NRC database.

Maximum Number of Fields. The maximum number of possible fields is 254.

Repeating Fields. Not supported.

Multiple Large Text Fields. Not supported for text fields that contain over 240 characters.

Indexing. An index file can be generated for any number of field names.

Searching Large Text Fields. Not supported for text fields that contain over 240 characters.

Report Generating. Programming skills are required to generate the reports for the NRC database application.

Data Dumping. Data dumping is handled through an export facility which allows extracting portions of a database. Data dumping can also be performed by redirecting the listing of a database file to a file external to the database.

Bulk Loading. Bulk loading can be performed from either a plain ASCII file or an exported data file. In the case of a plain ASCII file, field values have to be fixed in length; that is, the beginning and ending character positions for field values have to be the same across input records.

Integrity Checking. Integrity is handled through a screen building capability. This capability permits data values to be checked against a set of allowed values. However, pattern matching can be handled only through its programming language interface.

Structure modification. Fields can be appended to the end of a definition table or can be increased in character length. However, fields can not be dropped from a table definition.

Maximum Number of Records. The number of records is limited only by the amount of available storage space.

Menu Building. Menu Building is available through the high level language interface. Currently, C is the only high level language available for interfacing on the PC family.

Programming. Programming is available through the external C language and an internal PASCAL-like language.

REVELATION

Variable Length Fields. Specification of field lengths is not required and is limited only by the maximum number of characters allowed per record.

Maximum Record Length. A record can contain up to 65,000 characters.

Maximum Field Length. A field can be assigned up to 65,000 characters.

Maximum Number of Fields. The maximum number of fields is limited by the screen size. This problem can be resolved by linking screens. This concept is called paging. Using paging the maximum number of fields is 65,000.

Repeating Fields. Fields can be either single or multivalued. If a field is a multivalued type, a set of values can be assigned to that field for a single record occurrence.

Multiple Large Text Fields. The number of large text fields is limited only by the number of characters allowed per record. Thus, the number of large text fields is determined by the selected partitioning of 65,000 characters.

Indexing. The index list can be maintained in any Revelation file. The record key is the indexed value. The field value is a multi-valued list of keys for the records in the file being indexed. Normally, one would create a seperate cross-reference file for each Revelation file that requires cross referencing. The cross-reference file is specified by name and field number.

Searching Large Text Fields. The substring capability allows searching any field size. Searches will be performed much faster on large text fields that have been indexed. Large text fields can be indexed using the Revelation command "invert.all" .

Report Generating. Reports in the form required by the NRC database will require programming capabilities that are more detailed than the general reporting capabilities.

Data Dumping. Data can be dumped through the use of its BASIC programming language or its export facility called PORTER. In either case, a plain ASCII file can be generated. The copy command will copy any group of Revelation records to an ASCII file. Pdisk from Revelation G.2 will redirect printer output to any DOS file.

Bulk Loading. Data can be bulk loaded through the use of its BASIC programming language or its import facility called PORTER. In either case, field values can be either fixed or variable length. However, if values are variable in length, they have to be delimited and separated by commas or some other user specified character.

Integrity Checking. Integrity checking is handled through the use of a pattern matching capability. A pattern is set up when a field is defined. It allows the user to establish editing criteria for that field. The data entered into that field must match the pattern before it can be accepted by the system as valid. A BASIC program can also be called to perform the function of pattern matching.

Structure Modification. New fields can be easily added to an existing structure. However, fields cannot be easily removed from an existing structure. In Revelation, the file's dictionary references each field in a record by its field position. You can remove a field with three lines of BASIC code, and then renumber the dictionary's fields, as appropriate. A better alternative would be to null all the values for that field. This would only take one extra character per record in the data base while eliminating the need to modify the dictionary.

Maximum Number of Records. The number of records is limited only by the amount of available storage space. Records may be joined from multiple disks.

Menu Building. Menu building is available in its purest form. The menu building facility allows a developer to tailor interfacing to a database to a specific application.

Programming. A version of the BASIC programming language is available and can be accessed from within the database management system. This version of BASIC is specifically designed for database management and it is an integral part of Revelation.

SCI-MATE

Variable Length Fields. Specification of the lengths of fields is not required when defining a data structure. Storage is allotted character-by-character.

Maximum Field Length. A field length is limited only by the maximum characters allowed per record.

Maximum Record Length. The maximum record length is 1,894 characters per record. However, this limit can be exceeded through a record overflow method. If more than 1,894 characters is assigned to a record, it will be linked to an overflow record which can hold up to 1,894 characters. This process is repeated until entry is completed or the storage space is depleted.

Maximum Number of Fields. Only 20 fields can be specified per record.

Repeating Fields. Not supported.

Multiple Large Text Fields. Fields can contain only the 1,894-characters limit.

Indexing. Not supported.

Searching Large Text Fields. Limited by the 1,894 characters allowed per record.

Report Generating. Reports can be generated only in columnar form. This is inappropriate for the Proposed NRC Database Structure, which has a minimum of 44 fields. For example, if each field was only 5 characters wide, a report page would be 220 characters wide. This page width exceeds the normal carriage widths of printers which are generally 80 and 132 characters.

Data Dumping. Not supported.

Bulk Loading. Records can be loaded from any text file that meets a required format. Records must be separated by "*****" and fields properly indicated by a coding scheme.

Integrity Checking. Not supported.

Structure Modification. Structure modifications can be easily made. Fields can be added or deleted at any time. The commands to do this are menu driven.

Maximum Number of Records. The maximum number of records varies with record size and available storage space. For example, if the available storage space is 10 megabytes and the record size is 102 characters, the maximum number of records would be 32,768; however, if the available storage space is 10 megabytes and the record size is 1,894 characters, the maximum number of records would be 5,161.

Menu Building. Not supported.

Programming. Not supported.

ANALYSIS OF SOFTWARE

The recommended software is Revelation, version G, for the development site. The compiled or run-time model is recommended for the distributed user sites. The respective costs are $950 and $200.

Revelation is recommended over the other packages because it satisfies all of the criteria specified in the feature analysis section while the other packages do not. For example, Dayflo does not support menu building and programming. Oracle does not support repeating fields, searching fields that contain more than 240 characters, and records that contain more than one field larger than 240 characters. Sci-mate is limited in so many ways that is not necessary to elaborate on them.

Future studies deal with development of the prototype database using Revelation and evaluating the performance of the software for this application.

HARDWARE RECOMMENDATIONS

The following recommendations are based on the assumption that Revelation is to be used as a DBMS to satisfy the requirements. For the development of the proposed NRC database and for high-usage sites, an IBM AT or equivalent with the following configuration is recommended:

Item	Price
30 Megabyte fixed disk (internal) 1.2 Megabyte floppy disk 512 K RAM memory	$4,196
Serial/parallel adapter	105
360 Kilobyte floppy disk	297
80287 Math Coprocessor	262
Monitor adapter	175
Monitor (monochrome)	192
Hayes 2400S modem (external)	675
IBM Quietwriter	947
9 Pin Serial adapter (with an RS232 conversion)	25
Printer cable	25
RS232 serial cable	44
20+ Megabyte Bernoulli	3,756
Total	$10,699
Cost with an Epson FX-85 printer	$10,102

The 30 megabyte fixed disk (internal), 1.2 megabyte floppy disk, and 512K RAM memory are sold as a unit price. The 80287 math coprocessor is recommended by the producers of Revelation for improving the performance of the software. The monitor adapter should be a monochrome card for better screen resolution. The Hayes 1200S modem will be use to communicate with other distributed sites. The IBM Quietwriter is recommended because of its report quality and dot matrix capabilities. The 9 pin serial adapter is necessary since the serial/parallel card, which is standard with the enhanced AT, has a 9 pin serial port which is not compatible with the standard 25 pin RS232 serial port. The 20+ Megabyte Bernoulli is recommended because of its storage volume as a data transfer medium. Additionally, it is recommended that the development site acquire copies of the DOS 3.1, DOS technical, and the AT technical reference manuals.

For low-usage sites, an IBM XT or equivalent (i.e., an

IBM PC upgrade) with the following configuration is recommend:

Item		Price
43 Megabyte fixed disk (internal)		
360 Kilobyte floppy disk		$2,942
256 K RAM memory		
Ast 6 Pak/with 256K RAM		318
8087 Math Coprocessor		145
Monitor adapter		175
Monitor (monochrome)		192
Epson FX-85		350
2400 baud modem (external)		675
Printer cable		25
RS232 Serial cable		44
20+ Megabyte Bernoulli		3,756

	Total	$8,622

The 43 megabyte fixed disk (internal), 360 kilobyte floppy disk, and 256K RAM memory are specified as a unit price. To get the additional 256K RAM of memory, the AST 6 Pak card is required. The 8087 math coprocessor is needed to improve the performance of Revelation. The IBM monochrome card has both monitor and parallel adapters. The Epson FX-85 is recommended for the same reason as previously stated. The 20 Megabyte Bernoulli is recommended for the same reason as previously stated.

Futhermore, as a minimum configuration, an IBM PC could be configured with the following hardware, with the understanding that it would be used in conjunction with a Bernoulli as discussed in the text of the report.

Item	Price
IBM PC....with 256K RAM 2 DS/DD floppy drive keyboard	$1606.00
Monitor Adapter Card	175.00
AST 6 Pak w/256K	318.00
8087 Math Coprocessor	161.00
IBM Mono Monitor	192.00
Epson FX85	350.00
Printer Cable	25.00
2400 Baud Modem (external)	675.00

A Bernoulli Controller Card		130.00
RS232 Serial Cable		44.00

	Total	$3,676.00

Backup storage media have been omitted from the above, because floppy diskettes and the 20 megabyte Bernoulli carteriages can be used initially. After the database has grown too large for floppys to be effective and practical, it is anticipated that very economical alternatives with very high storage capacities and quick access times will be on the market at relatively low cost, when compared with the backup hardware that is now marketed.

In addition to the frequency of usage per week, another criteria for recommending an IBM AT over an IBM XT is the response time when searching large volumes of text. The IBM XT should be regarded as a machine for use only when indexed items are to be searched. The reason for this is that search time on large volumes of text is likely to be in terms of minutes even when the AT is used; therefore, use of an XT for such searches is impractical because the AT runs twice as fast as the XT.

APPENDIX I

A Preliminary File Structure Used to Develop Software and Hardware Recommendations

PROPOSED NRC DATA FILE STRUCTURE

Data element 1: Citation No. (8 chars.)
Data element 2: Publication type (10 chars.)
Data element 3: Authors (10 to 110 chars.)
Data element 4: Editors of Publication (10 to 50 chars.)
Data element 5: Article No. (1 letter plus 5 numbers)
Data element 6: Article Title (100 chars.)
Data element 7: Publication Series (30 chars.)
Data element 8: CODEN (10 chars.)
Data element 9: Vol., Ed. or No. (2 numbers)
Data element 10: Issue (2 numbers)
Data element 11: Page (3 numbers)
Data element 12: Contract No. (6 chars. plus 5 numbers)
Data element 13: Publication Date (8 chars.)
Data element 14: Publisher, City, State (100 chars.)
Data element 15: Sponsor and Address (100 chars.)
Data element 16: Site Address (60 chars.)
Data element 17: Patent Nos. (9 numbers)
Data element 18: Application or Meeting Date (8 chars.)
Data element 19: Dissertation Degree (15 chars.)
Data element 20: Date of last Data record update (6 numbers)
Data element 21: Country of Patent issue (10 chars.)
Data element 22: No. of Authors (2 chars.)
Data element 23: Country (Authors) (12 chars.)
Data element 24: Language (8 chars.)
Data element 25: ISSN, ISBN (10 chars.)
Data element 26: Abstract Source (15 chars.)
Data element 27: Abstract Nos. (6 numbers)
Data element 28: Related Citation Nos.
(0 to 36 numbers; 6 numbers repeating up to 6 times)
Data element 29: CA Registry No. (6 numbers)
Data element 30: Availability (15 chars)
Data element 31: Document Key Words (60 to 200 chars.)
Data element 32: Property and Form of Data Comments (10 to 1200 chars.)
Data element 33: Property and Form of Data Keywords (50 chars.)
Data element 34: Materials and Specimen Geometry Comments
(50 to 1200 chars.)
Data element 35: Materials and Specimen Geometry Key Words (50 chars.)
Data element 36: Test Conditions Comments (50 to 1800 chars.)
Data element 37: Test Conditions in temperature
(5 numbers, repeating up to 12 times)
Data element 38: Test Conditions in pressure
(5 numbers, repeating up to 12 times)
Data element 39: Test Conditions in time
(5 numbers, repeating up to 12 times)

Data element 40:	Experimental Methods (100 chars.)
Data element 41:	Comments on Data Validity and experimental methods (500 chars.)
Data element 42:	Document Abstract 400 to 4000 (chars.)
Data element 43:	Comments on Computational analysis (70 to 5000 chars.)
Data element 44:	Executive Summary (2000 to 7000 chars.)

APPENDIX II

IMPLEMENTATION OF REVELATION FOR A HLW DATABASE:

A. SCHEMA FOR A HIGH LEVEL WASTE DATABASE

B. REVIEWER'S INSTRUCTIONS FOR THE WASTE PACKAGE DATA REVIEW FORM

C. LIST OF DATA ELEMENTS FOR THE BIBLIOGRAPHIC FILE AND THEIR DESCRIPTIONS

D. SAMPLE OF A COMPLETED REVIEW

E. SAMPLE OF BIBLIOGRAPHIC CITATIONS

APPENDIX IIA

SCHEMA FOR A HIGH LEVEL WASTE DATABASE

General Data Element Categories Containing Content Description

C - Data Element is not Limited to a Keyword List
K - Data Element is Limited to a Preset Keyword List
M - Multiple Entries Permitted in a Data Element
N - Numerical Data Element
S - Single Entry Permitted in a Data Element

Titles for the Checklists of Keywords

1.	Scope of Work	(M,K)
2.	Model/Methodology	(M,K)
3.	General Environment	(M,K)
4.	Water Present	(M,K)
5.	Other Materials Present in the Environment	(M,C)
6.	Material Studied (General Type)	(M,K)
7.	Material Studied (Standard Designation)	(M,C)
8.	Material Condition Prior to Tests	(M,K)
9.	Electrolytes	(M,K)
10.	Radionuclides and Material Containing Them	(M,K)
11.	Measurements	(M,K)
12.	Mechanical and Thermophysical Properties	(M,K)
13.	Failure Modes or Phenomena Studied	(M,K)

General Comment Data Elements from the
WASTE PACKAGE DATA REVIEW FORM
(their relationship to Checklists 1 to 13)

1.	Type of Data (1., 13.)	(S,C)
2.	Materials/Components (6., 7.)	(S,C)
3.	Test Conditions (12., 3., 4., 5., 8., 9., 10.)	(S,C)
4.	Method of Data Collection/Analysis (11., 2.)	(S,C)
5.	Amount of Data (Data and Graph Summary Tables)	(S,C)
6.	Uncertainties in Data	(S,C)
7.	Deficiencies/Limitations in Data Base	(S,C)
8.	General Comments	(S,C)
9.	Abstract (if any) from HLW entry	(S,C)

Summary of Graphs z = f(X,Y)

z = property studied
X = variable on x axis
Y = variable on y axis

1.	Property Measured	(S,K)
2.	Property of X Variable	(S,K)
3.	Units of X Variable	(S,K)
4.	Minimum of X Variable	(S,N)
5.	Maximum of X Variable	(S,N)
6.	Property of Y Variable	(S,K)
7.	Units of Y Variable	(S,K)
8.	Minimum of Y Variable	(S,N)
9.	Maximum of Y Variable	(S,N)
10.	Figure Title	(S,C)

Data Summary Table z = f(X,Y,t) where

z = property studied
t = time
X = X variable (such as Temperature)
Y = Y variable (such as Pressure)

1.	Property Studied (name of z)	(S,K)
2.	Numerical Value of z	(S,N)
3.	Units of z	(S,K)
4.	Uncertainty of z assigned within HLW entry	(S,N)
5.	Name of X variable	(S,K)
6.	Numerical Value of X	(S,N)
7.	Units of X	(S,K)
8.	Name of Y variable	(S,K)
9.	Numerical Values of Y	(S,N)
10.	Units of Y	(S,K)
11.	Numerical Time Value (t)	(S,N)
12.	Time Unit Name	(S,K)
13.	Table Title	(S,C)

1. Scope of Work (M,K)

1. |__ data analysis
1. |__ experimental data
1. |__ literature review
1. |__ planned work
1. |__ theory
1. |__ other ______________________________.

2. Model/Methodology (M,K)

2. |__ Latin Hypercube
2. |__ Monte Carlo
2. |__ PDF (probability distribution functions)
2. |__ sampling
2. |__ scoping test
2. |__ other ______________________________.

3. General Environment (M,K)

3. |__ J-13 water
3. |__ basalt
3. |__ field site
3. |__ granite
3. |__ laboratory
3. |__ radiation field (alpha)
3. |__ radiation field (gamma)
3. |__ salt
3. |__ simulated field site
3. |__ tuff
3. |__ other ______________________________.

4. Water Present (M,K)

4. |__ J-13 water
4. |__ PH
4. |__ basalt composition
4. |__ brine
4. |__ deionized
4. |__ flow rate
4. |__ granite composition
4. |__ redox condition
4. |__ salt concentration
4. |__ significant dissolved species concentration
4. |__ tuff composition
4. |__ other ______________________________.

5. Other Materials Present in the Environment (M,C)

5. |__ Cl
5. |__ Cu
5. |__ Fe
5. |__ J-13 water
5. |__ Keller's reagent
5. |__ Ni
5. |__ sulfur ions
5. |___other ____.

6. Material Studied (General Type) (M,K)

6. |__ brass
6. |__ bronze
6. |__ cast iron
6. |__ cladding
6. |__ copper base
6. |__ electrolyte
6. |__ general environment
6. |__ nickel base
6. |__ packing
6. |__ radionuclide
6. |__ stainless steel
6. |__ steel
6. |__ titanium base
6. |__ water
6. |__ weld
6. |__ zircaloy
6. |__ zirconium base
6. |___other ____.

7. Material Studied (Standard Designation) (M,C)

7. |__ 304 stainless steel
7. |__ 304L stainless steel
7. |__ 308L weld filler wire
7. |__ 316L stainless steel
7. |__ AISI 317L
7. |__ AISI 321
7. |__ AISI 347
7. |__ AISI 1020
7. |__ AISI 1025
7. |__ J-13 steam
7. |__ J-13 water
7. |__ bentonite
7. |__ deaerated distilled water
7. |__ distilled water
7. |__ grey cast iron
7. |__ high-nickel alloy 825
7. |__ nodular cast iron
7. |__ zircaloy-4
7. |___other ____.

8. Material Condition Prior to Tests (M,K)

8. |__ case hardened
8. |__ cast
8. |__ cold worked
8. |__ irradiated
8. |__ magnetized
8. |__ mill annealed
8. |__ prestressed
8. |__ sensitized
8. |__ sintered
8. |__ solution treated
8. |__ stress relieved
8. |__ textured
8. |__ welded
8. |__ wrought
8. |___other ____.

9. Electrolytes (M,K)

9. |__ J-13 water
9. |__ acetic
9. |__ alkaline
9. |__ aerated
9. |__ chloride
9. |__ deaerated distilled water
9. |__ irradiated
9. |__ neutral
9. |___other ____.

10. Radionuclides and Materials Containing Them (M,K)

10. |__ Co60
10. |__ Np237
10. |__ Pu239
10. |__ commercial high level waste (CHLW)
10. |__ defense high level waste (DHLW)
10. |__ spent fuel (power reactors)
10. |__ spent fuel (water reactors)
10. |___other ____.

11. Measurements (M,K)

11. |__ adsorption
11. |__ electrochemical
11. |__ microscopy
11. |__ neutron diffraction
11. |__ slow strain rate
11. |__ sorption
11. |__ spectroscopy
11. |__ surface film
11. |__ tensile test
11. |__ thermal history
11. |__ visual examination
11. |__ weight change
11. |__ x ray diffraction
11. |___other ____.

12. Mechanical and Thermophysical Properties (M,K)

12. |__ bent beam tests
12. |__ creep strength
12. |__ density
12. |__ elongation
12. |__ heat (conduction)
12. |__ heat (convection)
12. |__ heat (radiative)
12. |__ heat capacity
12. |__ hydrostatic head
12. |__ lithostatic pressure
12. |__ modulus of elasticity
12. |__ stress-strain
12. |__ tensile strength
12. |__ thermal conductivity
12. |__ thermal expansion
12. |__ yield strength
12. |___other ____.

13. Failure Modes or Phenomena Studied (M,K)

13. |__ buckling
13. |__ corrosion (crevice)
13. |__ corrosion (general)
13. |__ corrosion (intergranular)
13. |__ corrosion (local)
13. |__ corrosion (microbial)
13. |__ corrosion (pitting)
13. |__ corrosion (stray current)
13. |__ corrosion (stress cracking - SCC)
13. |__ creep
13. |__ creep buckling
13. |__ dealloying
13. |__ debonding
13. |__ deformation (elastic)
13. |__ deformation (plastic)
13. |__ degradation (spent fuel)
13. |__ devitrification (glass)
13. |__ diagenetic-like changes
13. |__ fatigue (corrosion)
13. |__ fatigue (high cycle)
13. |__ fatigue (low cycle)
13. |__ fatigue (thermal)
13. |__ fracture (brittle)
13. |__ fretting
13. |__ galvanic
13. |__ hydration (glass)
13. |__ hydrogen attack (CH3 formation)
13. |__ hydrogen embrittlement
13. |__ leaching (radiation enhancement)
13. |__ leaching (spent fuel)
13. |__ matrix dissolution (glass)
13. |__ passivity
13. |__ poisoning (chemical)
13. |__ radiation effects
13. |__ relaxation (thermal)
13. |__ rupture (ductile)
13. |__ rupture (stress)
13. |__ sensitization
13. |__ spalling
13. |__ thermal instability
13. |___other ____.

REVIEWER'S INSTRUCTIONS FOR THE WASTE PACKAGE DATA REVIEW FORM

DRAFT:RShull:1-30-86

TYPE OF DATA:

(1) Scope of the Report: e.g., Experimental, Theoretical, Literature Review, Data Analysis

(2) Failure Mode or Phenomenon Studied: e.g., Corrosion, Creep, Fatigue, Leaching, Pitting, Hydrogen Embrittlement, Debonding, Dealloying

MATERIALS/COMPONENTS:

Description of the material studied (or component part, if specifically addressed as such; e.g., the screw-on type cap on a waste cylinder): e.g., 304L Stainless Steel, Brass, Zircalloy Cladding, Welds in 316 Stainless Steel, Packing Material, Basalt

TEST CONDITIONS:

Includes (1) the State of the material being tested, and (2) the Environment of the material being tested, e.g.:

(1) Cold worked or annealed 304L Stainless Steel, Thermo-mechanical history of the material (or component) being studied

(2) Aqueous environment, Radioactive surrounding, Electrolytes or corrosive agents present, Temperature and pressure (externally applied or not) during the test

METHODS OF DATA COLLECTION/ANALYSIS:

Includes Data Measurement Methods and Types of data measured, as well as Data Analysis Techniques, e.g.:

Electron microscopy, weight loss vs. time, slow strain rate tensile test, x-ray diffraction, differential thermal analysis, A.C. electrical resistivity using a Wheatstone bridge, mass spectroscopic chemical analysis of the corrosive environment, Latin Hypercube method, Monte Carlo techniques

AMOUNT OF DATA:

Includes the number of tables and graphs of data together with their titles and axes (indicating the range in values), e.g.:

5 tables of temperature and time data for five molten glass pouring operations, each table including the data from ten sensor locations. The temperatures ranged from 1100 °C to 0 °C over a time period of 24 hours.

UNCERTAINTIES IN DATA:

Included here are error bars and uncertainties in the data as stated by the author. This also includes qualitative statements by the author on the reliability of the data, e.g.:

Temperatures carry an accuracy of ±5 °C while the times are reported to within ±15 sec. It was felt that under real glass pouring operations (without well controlled crucible cooling) the temperature-time curves will be shifted to somewhat higher temperatures than shown here.

DEFICIENCIES/LIMITATIONS IN DATABASE:

Includes statements by the author on the applicability of the data, e.g.:

Extrapolation of the temperature-time (time < 24 hrs) data presented here to times in excess of 100 years should not be performed. The data presented here is useful only for indicating trends and qualitative parameter relationships, not for the purpose of presenting absolute values.

APPLICABILITY OF DATA TO LICENSING:

Includes information pertaining to specific Listed Licensing Issues in an NRC site characterization plan (ISTP). If an issue is specifically addressed in the paper, then the "key data" box should be marked. Otherwise, the paper is "supporting data." Comments on the listing of the document under subcategories (a) and (b) may be made in those subsections. The comments section in this category are the reviewer's comments pertaining to licensing.

GENERAL COMMENTS:

The reviewer's general comments on the document. This category is wide open as far as content. It contains information the reviewer did not enter in any of the above categories, but which is considered important for the reader to know, e.g.:

This is a very comprehensive review of the literature on the temperature sensitization of stainless steels. Even though it neglects the definitive work of Bertocci, Shull, Kaufman, and Escalante [Phys. Rev. J13, (1879), pp. 15-358] in this area (presumably because of the difficulty in locating this document), this review still considers a sufficiently large number of other investigations to provide a good understanding of the present status of the field. The one discordant note here, however, is that it would have been a much more useful review if stainless steel types 301, 303, 304, 316, and 440C had also been addressed.

It would be in this section that the reviewer's own comments on the deficiencies and uncertainties in the data and analysis would appear.

DATA SOURCE:

Full document reference. This section will be completed for the reviewer before he/she receives the document.

KEY WORDS:

These are already entered, as they are included in the entries of the above categories.

DATE REVIEWED:

The date the document review was completed.

Directory for Key Word List Completion

The following is a listing of the major headings of the Waste Package Data Review Form indicating (both by title and list numbers) which key word lists may be filled out by the reviewer from the information that appears in that review form heading.

TYPE OF DATA: Scope (1), Failure Mode (13)

MATERIALS/COMPONENTS: Material Studied (6, 7)

TEST CONDITIONS: (a) State of the Material (5, 8, 12)
(b) Tests (3, 4, 9, 10)...

METHOD OF DATA COLLECTION: Measurement Type (11)
Model (2)

AMOUNT OF DATA: Number of tables, graphs, and titling numerical results and their axes and ranges

UNCERTAINTIES IN DATA:

DEFICIENCIES/LIMITATIONS IN DATABASE:

APPLICABILITY OF DATA TO LICENSING:

(a) Relationship to Waste Package Performance Issues Already Identified

(b) New Licensing Issue

(c) Comments

GENERAL COMMENTS:

DATA SOURCE:

KEY WORDS: From List of Elements

APPENDIX IIC

LIST OF DATA ELEMENTS FOR THE BIBLIOGRAPHIC FILE AND THEIR DESCRIPTIONS

DATA ELEMENTS LIST FOR THE BIBLIOGRAPHIC FILE

Entry Program for HLW Reference File (I-1)

HLW.ENTRY

01 CIT.NO
02 PUB.TYPE
03 AUTHOR.S
04 EDITOR.S
05 CHAP.NO
06 SUB.TITLE
07 PUB.TITLE
08 PUB.ORG
09 PUB.SPONSOR
10 SITE.REL
11 PUB.AVAIL
12 PUB.NO
13 CONTRACT.NO
14 PAT.NO
15 PUB.VOL
16 PUB.ISSUE
17 PUB.PAGE
18 PUB.CODEN
19 PUB.ISSN
20 PUB.DATE
21 OTHER.DATE
21 START.DATE
23 CHANGE.DATE
24 CIT.REL

Descriptions of the Bibliographic Data Elements

01 CIT.NO

A unique NBS-assigned reference number assigned to every entry in the HWL collection.

02 PUB.TYPE

The form of the publication: 1 general; 2 journal; 3 book; 4 conference; 5 patent; 6 magnetic media; 7 dissertation; 8 draft or in press; 9 communication.

03 AUTHOR.S

The persons (not organizations) who created the entry. The patronymic name comes first (i.e., Smith, W. L. A., Jr., or O'Ryan, C.).

04 EDITOR.S

The editors and/or translators of the entry (not the NBS or other reviewers).

05 CHAP.NO

A number pertinent to the entry such as article number, chapter number, conference paper number, or report number within a larger document.

06 SUB.TITLE

The subtitle of the pertinent section within a HLW entry such as: 1 an article title; 2 chapter title in a book; 3 journal paper or report title within a conference report.

07 PUB.TITLE

HLW entry title: 1 book title; 2 journal title; 3 report series name; 4 conference or proceedings name; 5 dissertation series name.

08 PUB.ORG

The name and possibly a short address of the organization that created the HLW entry.

09 PUB.SPONSOR

The name and address of the sponsoring organization: 1 contract report sponsor; 2 conference sponsor; 3 patent sponsor.

10 SITE.REL

Additional site locations if needed such as: 1 meeting place of conference (if not included in the title); 2 university of dissertation.

11 PUB.AVAIL

Availability of the document to the general public, and ordering information.

12 PUB.NO

The number or numbers assigned to the HLW entry by the creating organizations for cataloging purposes (i.e., an ORNL number, a NUREG number, etc.).

13 CONTRACT.NO

The contract number under which the report was generated. There may be more than one report under a contract number.

14 PAT.NO

Patent number, patent pending number.

15 PUB.VOL

1 volume number of journal; 2 edition of book; 3 number of conference or proceedings.

16 PUB.ISSUE

The issue number of a journal or publication.

17 PUB.PAGE

The page number of document (start-end), or total pages

18 PUB.CODEN

The ASTM CODEN, with check character, as assigned by the Chemical Abstracts Services. A CODEN is a six character code describing a serial publication such as a journal.

19 PUB.ISSN

The International Standard Serial Number (ISSN) or the International Standard Book Number (ISBN).

20 PUB.DATE

The date of publication, issue date of patent, or date of conference proceedings publication (not meeting date, see OTHER.DATE).

21 OTHER.DATE

Any additional date needed to describe the publication, as date of patent application, or date of conference meeting.

22 START.DATE

This is the date that this data entry was created.

23 CHANGE.DATE

This is the date that this data entry was last changed.

24 CIT.REL

The internal (NBS) numbers of other entries in the HLW data collection that are directly related to this entry.

APPENDIX IID

SAMPLE OF A COMPLETED REVIEW

WASTE PACKAGE DATA REVIEW FORM

TYPE OF DATA

Experimental study of the possible effects of a 3-year exposure of a 304L canister in Climax stock quartz mangonite and the implications for waste disposal in a tuff environment. The canister was examined by metallurgical examination and chemical analysis, and the water taken from the liner rock annulus was also analyzed for smilarities with anticipated tuff water chemistry.

MATERIALS/COMPONENTS

304L canister with 308L weld filler wire; plain C steel liner

TEST CONDITIONS

Three-year exposure; underground granite; canister containing spent fuel; water containing a significant concentration of dissolved species (including Cl) in contact with base of the welded 304L canister; maximum temperature ~ 140 °C; ambient pressure; total γ dose: 3.2×10^8 rads

METHODS OF DATA COLLECTION/ANALYSIS

Metallographic observation after exposure; thermocouple measurements during storage; chemical analysis of canister materials and well water before and after exposure

AMOUNT OF DATA

One table comparing water analysis of Climax facility with J-13 well water; one table listing results of chemical analysis of base metal, weld metal, and weld wire; one graph of temperature (20 - 120 °C) vs. time (2.4 - 3.6 yrs) for the canister and liner

UNCERTAINTIES IN DATA

Not addressed.

DEFICIENCIES/LIMITATIONS IN DATABASE

The authors discuss the validity of using the data from this environment for predicting the behavior in a tuff repository and conclude that the environment studied is more aggressive than the tuff environment indicating that the performance of 304L in tuff should be adequate for repository purposes.

APPLICABILITY OF DATA TO LICENSING

[Ranking: key data ( ), supporting data (X)]

(a) Relationship to Waste Package Performance Issues Already Identified

The data are considered supporting for issues 2.2 and 2.2.1 in the ISTP for NNWSI.

(b) New Licensing Issues

(c) General Comments

DATA SOURCE

(a) Organization Producing Data

Lawrence Livermore National Laboratory, Livermore, CA 94550

(b) Author(s), Reference, Reference Availability

Weiss, H., Van Konynenburg, R. A., and McCright, R. D., Metallurgical Analysis of a 304L Stainless Steel Canister from the Spent Fuel Test--Climax, UCID-20436, April 1985. Available from NTIS.

KEY WORDS *(These will be generated from the Keyword Checklist.)*

GENERAL COMMENTS

The authors suggest in their discussion that more severe chloride cracking occurs at the lower temperatures encountered in the spent-fuel test than at the higher temperatures expected for the repository. This is questionable based on previous work done in these alloys over a range of temperatures. For example, Kowaka and Kudo (Trans. JIM, 16 (1925) 385) show time to failure (t_f) curves for 304 as a function of temperature where there is a minimum in t_f at approximately 140 °C. Thus, it would be useful to perform similar experiments at slightly higher average temperatures (i.e., 130 - 145 °C) where SCC is most pronounced.

DATE REVIEWED: 10-28-85/Revised 12-2-85/1-13-86

APPENDIX IIE

SAMPLE OF BIBLIOGRAPHIC CITATIONS

001 Cit: 001 |Authors: Claiborne, H.C.;Croff, A.G.;Griess, J.C.;Smith, F.J. |Pub.title: Repository Environmental Parameters Relevant to Assessing the Performance of High-Level Waste Packages in Basalt, Tuff, and Salt |Pub.org: OAK RIDGE NATIONAL LABORATORY, Oak Ridge, Tennessee 37831 operated by MARTIN MARIETTA ENERGY SYSTEMS, INC. |Pub.sponsor: Division of Waste Management, Office of Nuclear Material Safety and Safeguards, NRC, Wash., DC 20555 |Pub.no: NUREG/CR-4134;ORNL/TM-9522/R1 |Contract.no: NRC FIN NO. B0288 |Pub.page: 338 |Pub.date: 09/00/85 |P: 1 |Start.date: 12-06-85 |Change.date: 12-06-85

002 Cit: 002 |Authors: Stephens, K.;Boesch, L.;Crane, B.;Johnson, R.;Moler, R.;Smith, S.;Zaremba, L. |Pub.title: Methodologies for Assessing Long-Term Performance of High-Level Radioactive Waste Packages |Pub.org: Eastern Technical Division, THE AEROSPACE CORPORATION, Wash., DC, |Pub.sponsor: Office of Nuclear Material Safety and Safeguards, NRC, Wash., D.C. |Pub.no: ATR-85(5810-01)-1ND |Contract.no: F04701-83-C-0084 |Pub.page: 188 |Pub.date: 05/00/85 |P: 1 |Start.date: 12-06-85 |Change.date: 12-06-85

003 Cit: 003 |Authors: Ballou, Lynden B.;McCright, R. Daniel |Pub.title: Overview Information on NNSWI's Corrosion Program |Pub.org: Lawrence Livermore National Laboratory |Pub.no: MRB-0418 |Pub.date: 03/08/85 |P: 9 |Cit.rel: 004;005 |Start.date: 10-25-85 |Change.date: 10-25-85

004 Cit: 004 |Authors: Ballou, Lynden B.;McCright, R. Daniel |Sub.title: NNWSI Test Plan for Copper and Copper Base Alloys |Pub.title: Overview Information on NNWSI's Corrosion Program |Pub.org: Lawrence Livermore National Laboratory |Pub.no: MRB-0418 |Pub.date: 12/10/84 |P: 9 |Cit.rel: 003;005 |Start.date: 12-06-85 |Change.date: 12-06-85

005 Cit: 005 |Authors: Ballou, Lynden B.;McCright, R. Daniel |Sub.title: Metal Barrier Testing--Objective and Issues Addressed |Pub.title: Overview Information on NNWSI's Corrosion Program |Pub.org: Lawrence Livermore National Laboratory |Pub.no: MRB-0418 |P: 9 |Cit.rel: 003;004 |Start.date: 12-06-85 |Change.date: 12-06-85

006 Cit: 006 |Authors: Juhas, M. C.;McCright, R. D.;Ga[illegible], R. L. |Sub.title: Behavior of Stressed and Unstressed 304L Specimens in Tuff Repository Environmental Conditions |Pub.title: Corrosion 85 NACE Annual Meeting |Pub.org: Lawrence Livermore National Laboratory |Pub.sponsor: U.S. Department of Energy |Site.rel: Boston, MA |Pub.no: UCRL-91804 |Contract.no: W-7405-ENG-48 |Other.date: 03-25-85 |P: 8 |Start.date: 10-25-85 |Change.date: 10-25-85

007 Cit: 007 |Authors: McCright, R. Daniel;Weiss, Haskell |Sub.title: Corrosion Behavior of Carbon Steels Under Tuff Repository Environmental Conditions |Pub.title: Materials Research Society 1984 Annual Meeting |Pub.org: Lawrence Livermore National Laboratory |Pub.sponsor: U.S. Department of Energy |Site.rel: Boston, MA |Pub.no: UCRL-90875 |Contract.no: W-7405-ENG-48 |Other.date: 11-26-84 |P: 8 |Start.date: 10-25-85 |Change.date: 10-25-85

008 Cit: 008 |Authors: Bates, John K. ;Oversby, Virginia M. |Sub.title: The Behavior of Actinide Containing Glasses During Gamma Irradiation In a Saturated Tuff Environment |Pub.title: Materials Research Society 1984 Annual Meeting |Pub.org: Lawrence Livermore National Laboratory |Pub.sponsor: U.S. Department of Energy |Site.rel: Boston, MA |Pub.no: UCRL-90818 |Contract.no: W-7405-ENG-48 |Other.date: 11-26-84 |P: 8 |Start.date: 10-25-85 |Change.date: 10-25-85

NW.REF DATA..

009 Cit: 009 |Authors: Fox, Michael J.;McCright, R. Daniel |Pub.title: An Overview of Low Temperature Sensitization |Pub.org: Lawrence Livermore National Laboratory |Pub.sponsor: U.S. Department of Energy |Pub.no: UCRL-15619 |Contract.no: W-7405-ENG-48 |Pub.page: 31 |Pub.date: 12/00/83 |P: 1 |Start.date: 12-06-85 |Change.date: 12-06-85

010 Cit: 010 |Authors: Glass, Robert S.;Overturf, George E.;Garrison, Robert E.;McCright, R. Daniel |Pub.title: Electrochemical Determination of the Corrosion behavior of Candidate Alloys Proposed for Containment of High Level Nuclear Waste in Tuff |Pub.org: Lawrence Livermore National Laboratory |Pub.sponsor: U.S. Department of Energy |Pub.no: UCID-20174 |Contract.no: W-7405-ENG-48 |Pub.page: 38 |Pub.date: 06/00/84 |P: 1 |Start.date: 12-06-85 |Change.date: 12-06-85

011 Cit: 011 |Authors: Rothman, A. J. |Pub.title: Potential Corrosion and Deqradation Mechanisms of Zircaloy Cladding on Spent Nuclear Fuel in a Tuff Repository |Pub.org: Lawrence Livermore National Laboratory |Pub.sponsor: U.S. Department of Energy |Pub.no: UCID-20172;MRB-0418 |Contract.no: W-7405-ENG-48 |Pub.page: 49 |Pub.date: 09/00/84 |P: 1 |Start.date: 12-06-85 |Change.date: 12-06-85

012 Cit: 012 |Authors: Oversby, V. M.;McCright, R. D.;Westerman, R. E.;Yow, J. |Sub.title: Lists of Planned Field and Laboratory Tests |Pub.title: LLNL NNWSI Program |Pub.page: 10 |P: 9 |Start.date: 12-06-85 |Change.date: 12-06-85

013 Cit: 013 |Authors: Weiss, H.;Van Konynenburg, R. A.;McCright, R. D. |Pub.title: Metallurgical Analysis of a 304L Stainless Steel Canister from the Spent Fuel Test--Climax |Pub.org: Lawrence Livermore National Laboratory Nuclear Waste Management Projects |Pub.sponsor: U.S. Department of Energy |Pub.no: UCID-20436 |Contract.no: W-7405-ENG-48 |Pub.page: 20 |Pub.date: 04/23/85 |P: 1 |Start.date: 12-06-85 |Change.date: 12-06-85

014 Cit: 014 |Authors: Smith, H. D.;Oversby, V. M. |Pub.title: Spent Fuel Cladding Corrosion under Tuff Repository Conditions-Initial Observations |Pub.org: Lawrence Livermore National Laboratory Nuclear Waste Management Projects |Pub.sponsor: U.S. Department of Energy |Pub.no: UCID-20499 |Contract.no: W-7405-ENG-48 |Pub.page: 4 |Pub.date: 06/00/85 |P: 1 |Start.date: 12-06-85 |Change.date: 12-06-85

14 Records Processed

NBS-114A (REV. 2-8C)

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET (See instructions)	1. PUBLICATION OR REPORT NO. NBSIR 86-3363	2. Performing Organ. Report No.	3. Publication Date JUNE 1986

4. TITLE AND SUBTITLE

"An Analysis of the Requirements for a Computer Assisted Database for Reviews and Evaluations on High Level Waste Data"

5. AUTHOR(S) C. Interrante, C. Messina, S. Harrison, R. Shull, M. Kaufman, U. Bertocci, and E. Escalante.

6. PERFORMING ORGANIZATION *(If joint or other than NBS, see instructions)*

NATIONAL BUREAU OF STANDARDS
DEPARTMENT OF COMMERCE
WASHINGTON, D.C. 20234

7. Contract/Grant No. NRC No. FIN A-4171-6

8. Type of Report & Period Covered

9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS *(Street, City, State, ZIP)*

U.S. Nuclear Regulatory Commission (NRC)
Division of Waste Management
Mail Stop SS 965
Washington, D. C. 20555

10. SUPPLEMENTARY NOTES

☐ Document describes a computer program; SF-185, FIPS Software Summary, is attached.

11. ABSTRACT *(A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here)*

The establishment of a computer-assisted database has been initiated, and is suggested for further development and full implementation for storage and retrieval of reviews and evaluations of high level waste (HLW) data, composed principally of pertinent waste-package reports generated by DOE. Initial suggestions by materials scientists at the NBS on the type and form of information to be stored in the database were reviewed by NBS computer scientists, who have concluded that from the available software packages, Revelation is the most suitable database management system (DBMS) for this application. Hardware that is compatible with existing NRC equipment is recommended. The format for storage and retrieval of information using Revelation will require further development for full implementation of this DBMS for use with the reviews of reports on the HLW package.

12. KEY WORDS *(Six to twelve entries; alphabetical order; capitalize only proper names; and separate key words by semicolons)*

Database; database management system (DBMS); fields; file; hardware; high level waste; microcomputer; query; records; software.

13. AVAILABILITY

☐ Unlimited
[XX] For Official Distribution. Do Not Release to NTIS
☐ Order From Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.
☐ Order From National Technical Information Service (NTIS), Springfield, VA. 22161

14. NO. OF PRINTED PAGES 61

15. Price $11.95

USCOMM-DC 6043-P80