Lawrence Berkeley National Laboratory
Recent Work

Title
Investigation of Network-based Information System Model

Permalink
https://escholarship.org/uc/item/56r8g263

Author
Konrad, A.

Publication Date
1996-09-09
Investigation of Network-Based Information System Model

A.M. Konrad, M. Perez, J. Rivera, Y. Rodriguez, M.J. Durst, D.W. Merrill, and H.H. Holmes

Information and Computing Sciences Division

September 1996
Investigation of Network-Based Information System Model

A.M. Konrad, M. Perez, J. Rivera, Y. Rodriguez, M.J. Durst, D.W. Merrill, and H.H. Holmes

Information and Computing Sciences Division
Ernest Orlando Lawrence Berkeley National Laboratory
University of California
Berkeley, California 94720

September 1996

This work was supported by the Assistant Secretary for Environment, Safety and Health, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.
DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.
Table of Contents

Acknowledgments
Abstract

1. Statement of problem ........................... 1
2. Approach ........................................ 2
3. Architectural Issues .............................. 3
4. System evaluation ................................ 10
5. Prototype Implementation ....................... 14
6. Results .......................................... 16
7. Significance of this Work ....................... 17
8. Possible Future Work ............................ 18
9. Summary of Deliverables ....................... 19

APPENDIX A  DBMS Evaluation Criteria, detail
APPENDIX B  Products and Systems Evaluated
APPENDIX C  Table of Evaluations (interim)
APPENDIX D  Selected Screens from WWW Access to CEDR Structured Documentation
APPENDIX E  Selected Screens from WWW Access to Epidemiology Guides (site archive records)
Acknowledgments

This project was supported by the LBNL Center for Science and Engineering Education (CSEE). We acknowledge the support of Laurel Egenberger and Nancy Salee of CSEE. We acknowledge the support of Barbara Brooks, Heather Stockwell, and Patrick Heinig with the office of the Assistant Secretary for Environment, Safety and Health in the U.S. Department of Energy, Environment and Health and Safety, who provide support for the CEDR Information System at LBNL. Guidance from colleagues at SLAC and LBNL was essential in concluding a successful Summer Research Program. We thank Louise Addis, George Crane, Dick Guertin (Stanford/RLIN), Pat Kreitz, and John Halperin, for their support from SLAC and Stanford, and John McCarthy and Mark Dedlow from Berkeley.
Abstract

The objective of the DOE-LBNL summer student research program in computer and information sciences focused on investigating database-based http-based information architectures, and implementation of a prototype using DOE's Comprehensive Epidemiologic Data Resource (CEDR) metadata or Epidemiology Guide content. We were successful in identifying the components of such an information system, an appropriate configuration given the requirements, and in implementing a prototype. This work comprised investigation of various information system architectures or variants, evaluation and selection of various tools, products, and packages, preparation of databases, database content, output formats, and graphical (World Wide Web-compatible) interfaces. We successfully prepared and demonstrated network access to content from both the CEDR structured documentation and from the DOE Epidemiology Guides (site archive records).
1. Statement of problem

The Information Problem, how to connect people with information, is complex, in part because there are many types of obstacles and intermediaries to mitigate those obstacles. The widespread use of network-based database technologies, including the internet, the world wide web (WWW), browsers, and server-to-database interfaces (Common Gateway Interfaces, CGI) have solved a number of access problems which confront the information seeker, while requiring information system designers to become familiar with a continuous stream of new architectures, tools, and methods.

During the course of this summer, Madel Perez, Yamilet Rodriguez and Jose Rivera worked under the supervision of Allan Konrad and Mark Durst with the Comprehensive Epidemiologic Data Resource (CEDR) program to investigate new information system architectures for CEDR. CEDR is a Department Of Energy (DOE) program to disseminate by various means, including electronic, access to research results and data from DOE-related epidemiologic studies during the last 30 years.

The tasks for this Summer Research Program were to develop a set of criteria for WWW-based access to the CEDR structured documentation (metadata) database, then an architecture, then determine what components were required, and of those, which could be obtained externally and which had to be developed ab initio. First it was necessary to evaluate a group of WWW-capable database management systems (DBMS) and select one. The time to finish all tasks for this Summer Research Program was 10 weeks.
2. Approach

The following was the work plan provided to the student research assistants at the beginning of the program.

Week 1  **Review of internetworking. Browsers.** Review HTML command language via Yahoo.  
          Search Engines. Review homepage. Scan course pages.
          
          Tasks: Individual homepage edited and mounted.

Week 2  **The Research Problem:** Develop model of components and functional requirements for Webware. Understand data flows. Review of information retrieval (IR) functionality. CGI tools. DBMS. email servers.
          
          Tasks: Survey candidate packages. 
          Presentation of plan and approach. Identify bibliographic materials as needed.

Week 3  **Analysis and selection.** Analysis and selection of packages and products for functionality and interoperability. Selection criteria. 
          Possible Applications: 
          Webware I (web-based groupware- email in, http in, email out, http out).  
          Webware II (CEDR metadata)  
          Webware III (web-based listserv archive)  
          Webware IV (CEDR catalog)  
          Webware V (instructional modules)  
          Webware VI (Epidemiologic Guides)
          
          Tasks: Status report. Selection of component specialization by each student.

Week 4  **Begin Implementation:** Procurement and installation

Week 5  **System Installation.**

Week 6  **Application implementation.**

Week 7  **Load demonstration data (Webware II).**

Week 8  **Demonstration.**

Week 9  **Write project reports.**

Week 10  **Wrap up, Poster Sessions.**
3. Architectural Issues

At the beginning of the project, we envisaged an architecture consisting of a web server (httpd), a database management system (DBMS) with several CEDR databases, and one or more Common Gateway Interface (CGI) scripts to convey data between the httpd and DBMS and further transform the data as needed in each direction. We anticipated as a major subtask of the summer project, writing these CGI scripts ourselves, using the PERL language.

Early into the project, we learned from LBNL colleagues and other sources, that this “home brew” approach to CGI scripts was only the first step in an evolution. The CGI Evolution, Phase I below indicates the originally anticipated configuration.

3.1 CGI Evolution, Phase I

By the “home brew” approach, we mean:

- the hand-crafting of a CGI script specifically for the target database, i.e., with specific business logic regarding data base name, database elements, and searchable indices;

and where the systemic operation is:

- request by client of a graphical interface form from server for specification of inquiry specifications;
- generation of GET/POST requests to be sent to httpd from client;
- requests received from client by httpd translated into CGI variables or stdin;
- requests received by CGI script as variables or stdin, reformatted by the CGI script as legal DBMS commands;
- requests received by the DBMS, processed against the database, results generated by the database, and then returned to CGI script;
- results received by CGI script, translated, with substantial business logic encoding, into HTML, and forwarded to the httpd;
- httpd conveys html using http to client browser.
A block diagram, figure 1, illustrates this architecture, comprised of the following:

**Client:**
- C1 Platform (hardware, OS, network connectivity)
- C2 Browser (http-capable if not multi-protocol)

**Network:**
- N1 Network

**Server:**
- S1 Platform (hardware, OS, network connectivity)
- S2 HTTP Daemon
- S3 Page containing a Link (optional, but likely) to search form
- S4 Form Page
- S5 CGI script
- S6 DBMS
- S7 HLI/means of send/receive data to other processes (If on same machine*)

*I.e., if S6 is on a different network address (i.e., different machine) than S5, then S7 becomes a network server, to mitigate the network boundary, and listens at a (well known) port, and binds an incoming socket to the port to support a specific client process.*
The components provide the following services:

C1  human interface service and network connectivity
C2  http communications capability, HTML rendering capability, 'helper' applications
N1  TCP/IP

S1  All of: HTTPD service, DBMS service, and network connectivity

S2  Send and Receive http commands (as/from HTML documents) to/from Client, translate commands, and send to/from CGI Script

S3  Store HTML text and URL pointing to a forms page

S4  Provide forms capabilities (fields, radio buttons, etc.) to specify a search or browse, and return values to HTTPD

S5  Both:
Receives forms input, translates into DBMS-acceptable languages, conveys to DBMS;
Receives DBMS response, performs HTML markup, conveys to Client Browser (C2) via S2.

Note: S5 might run on a machine with a different network address than httpd (S2) and thus be supported as a CGI program encapsulated in a network server, listening for connection requests, queuing, etc.

S6  DBMS capability (organized repository, indexing, input, output/report generator, query processing)

S7  Interoperability with DBMS by process other than a terminal, e.g., a CGI script
    DBMS same machine as CGI: Host Language Interface
    DBMS different machine as CGI: server listening at a port
How Capability is Provided (Interoperability):

Pertinent interoperations are:

C2-S2  http, ftp, gopher, etc.
S2-S3  HTML, etc.
S2-S4  GET (constructs a URL)
       POST (constructs MIME object)
S2-S5 (S4-S5) CGI recognition defined for httpd functionality
       To CGI: GET/URL stores variables in memory accessible to CGI
       POST/MIME sends to CGI via stdin
       From CGI: stdout
S5-S6/S7 To DBMS: Same Machine: HLI.
         Different Machine: network server
       From DBMS: Same Machine: HLI.
         Different Machine: network server

Does Application Developer Provide the Component?
C1  No, User
C2  No, User
N1  No, System
S1  No, System
S2  No, System
S3  Yes, we develop
S4  Yes, we develop
S5  Yes, we develop
S6  No, Vendor
S7  No, Vendor
3.2 CGI Evolution, Phase II

The essence of this phase is that laborious hand-crafting of CGI middleware to provide the services for S5 *supra* is virtually eliminated by redistributing the functionality of S5 (CGI script) as follows:

S5 still receives requests from httpd and passes to DBMS, and receives requests from DBMS and returns to CGI

However, most translation tasks and other database-specific tasks, e.g., translating forms/CGI requests into DBMS requests, and translating DBMS responses into HTML objects, are no longer performed by the CGI, but are partitioned to either the httpd, in the case of translating forms requests into dbms-legal commands), or assigned to the DBMS itself (generating HTML-ready output).

This implies that the httpd "knows" the syntax of DBMS-legal requests. This is accomplished by 'built-in' recognition of translations between forms specifications and legal DBMS command language, and facilitated by using a configuration file which is consulted by the httpd where application-specific attributes are specified.

Indeed, major database vendors are bringing to market their own httpd servers which provide these translation services without need for individual coding of CGI scripts. The CGI functionality of conveying requests in each direction is reduced to a generic capability that can be coded and compiled by the httpd/DBMS vendor as simply another module.

A Phase II block diagram (figure 2) illustrates this modified architecture. This model comprises additional components:

S8 config files containing application-specific information such as name of database, user, passwords, preference for error processing, logging, format names, other specifications as required.

S9 Stored procedure call processor accepts requests from httpd as configured according to individual config file specifications and transmitted by the CGI, and submits the query to the specified database and returns the result. This processor might be recognized to the DBMS as simply another (logged on) user, or process (if using S7 host language interface).

The benefits are:

1. The labor required to implement a forms-based www-based dbms application is greatly reduced.

2. The functionality remaining in the CGI requires no user-specific or application-specific information and can be treated as simply a compiled module that runs regardless of the database requested.
Figure 2
3.3 CGI Evolution, Phase III

This phase builds upon the benefit of Phase II by establishing connectivity between S2 (httpd) and S6 (DBMS) such that each successive DBMS request does not require reinitialization of the connection between S2 and S6, and accompanying generation of an image of database access modules. Where architecture of both S2 and S6 is under the control of one party, e.g., a database vendor, this connection can be established through efficient, if proprietary, means.

The primary benefit of Phase III is enhanced performance where the overhead of establishing a connection to a database or an image of a database access module in memory is substantial or where there is a high volume of requests as a multiplier of the overhead.

3.4 Relationship to Summer Research Project

In Section 4 below, is described our evaluation of various tools and systems considered for our information system architecture. During the course of our investigation, we determined:

- A significant training investment is required in developing PERL expertise;
- A significant coding investment is required in developing CGI scripts by hand
- We have access to "Phase II" and "Phase III" DBMS obviating the need to take a "Phase I" approach.

As the system evaluation below indicates, we did identify a set of Phase III tools that were within our scope for (the relatively short 10-week) summer research project and purchasing constraints (time and dollars).
4. System and product evaluation

The objective of the product and system evaluation effort was to identify tools with sufficient 'built-in' WWW accessibility that the overall work of implementing WWW access to CEDR databases would be minimized. Initially, we evaluated:

S5 (CGI script) tools

S6 Database Management Systems

Evaluation of S5 (CGI script) tools.

As explained supra, we learned during the course of our investigation that a Phase I approach might not be the most desirable, and that at least a Phase II approach was possible within the context of the Summer Research Program. Consequently, our initial effort devoted to developing both PERL and CGI expertise, while useful from a breadth-of-discipline standpoint, was abandoned in favor of a Phase II/III approach.

Evaluation of S6 Database Management Systems:

We understand database management system to comprise:

1. An organized repository of goal records stored "in" the DBMS according to a definition or schema also "in" the DBMS

2. An indexing engine, which builds indices of various kinds (word, phrase, date, personal name, etc.) from goal records newly added to (1) the organized repository; modifies index entries appropriately when goal records are modified; and deletes index entries when goal records are deleted from the organized repository.

3. A search engine which receives queries, processes them against index records to obtain search results (a set of pointers to goal records selected), and then uses those pointers to provide the count of the records satisfying the search result.

4. A report generator which uses search result sets or direct (e.g., display or sequential processing) commands to fetch specified goal records and then present them in a specified format either to a specified file, to a terminal, or to another program, such as a CGI program, for further processing or transmission to another program means.

5. Support for text/character manipulation capabilities (please see Appendix).


7. Support for optionally-occurring elements.

8. Support for dates

9. Support for pre-defined procedures (on records upon input to organized repository, on records upon output from dbms, on goal records as passed to index records, on query values as processed against index records)

10. Support for personal name processing.

11. Support for variable length elements.

12. Support for web/http (CGI) interoperation (Phase II or Phase III)
Systems were identified as candidates, and then evaluated against the above twelve criteria. Candidate systems were identified from a variety of sources, including:

a) WWW
   1) Free Databases list from Yahoo/Database page
   2) Non Free Databases - company homepages
b) FTP sites
c) Information from LBNL colleagues
d) Information from external colleagues

From these sources, a preliminary list of candidate systems was developed (please see Appendix B).

Evaluation was performed on each candidate system initially only to determine whether the first four criteria were met by the candidate system. This resulted in a narrower list of finalists for which further evaluation was performed using criteria 5 - 12.

Evaluation on any candidate system was done by whatever means were available to us within the constraints of our Summer Research Program. We did not invite any vendor for a presentation. We began evaluation of nearly every candidate with information from the vendor or distributor's web pages, ftp sites, or other information we were able to obtain. We utilized:

1) Telephone
2) E-mail
3) Fax
4) Regular Mail
5) Web sites
4) FTP sites
Results of Evaluation.

Twenty two (22) systems received preliminary evaluation. Of these:

Seven (7) were "web search engines only". That is, they did not support goal records in an organized repository of their own, but used agents to scan the web, fetch documents, and build word indices. These were considered unsuitable because, without an organized repository and a record structure under CEDR's control, we could not control the format of output, or even generate reformatted output at all. These were:

1) Glimpse 3.0
2) lq-text
3) FFW- Free text for WWW
4) MG Information Retrieval System
5) qt (Query Text)
6) Topic Internet Server
7) WAIS

Four (4) were non-responsive. In this category, we include those for which no telephone number or network address was found for obtaining further information, or for which the distributor indicated there was 'no support available' for the product, or where repeated telephone inquiries were not returned, or where our inquiries were answered with "the only documentation we have comes with our system - get our system and install it, and then you can read the documentation". These included:

1) Postgres 95
2) MORE- Multimedia Oriented Repository Environment
3) System II
4) SMART

More detailed evaluation was performed upon the eleven (11) remaining systems, with the following results:

There were several products which did not meet functional requirements related to bibliographic-type applications (criteria 5-11):

1) RDB
2) ORACLE (base product)
3) Sybase (base product)

Sybase and ORACLE both recommended use of separate ancillary products to meet the bibliographic/textual requirements. We read these recommendation to mean that the core DBMS offerings themselves were not adequate to meet our requirements in the judgement of the vendors. One of these ancillaries was a third-party product, one was a product offering of the primary DBMS manufacturer. Although the combination of these DBMS products and their application add-ons might have met our functional requirements, we discontinued further evaluation because the investments required, both in licensing and in staff training time, were beyond the scope of our Summer Research Program. If an evaluation process less constrained by time and other resources becomes possible, these two, Oracle with Context, and Sybase with Fulcrum, might be given further consideration.

There was a class of products which can be described as WWW-DBMS middleware. Although these did not meet our DBMS evaluation criteria, they could provide an interesting alternative approach in a less constraining evaluation program. We encountered one such product (although we encountered mention of a small number of others):

1) ESQL/Seekersoft
One system was not evaluated because we determined that it was still in beta stage on the basis of postings to a listserv established for users and beta testers. However, a future evaluation should include this system:

1) ISITE/ISEARCH from Center for Networked Information Discovery and Retrieval (CNIDR)

The remaining six (6) appeared to be finalist candidates:

1) SPIRES
2) AIRS II
3) BRS/Search
4) Microsoft Access
5) Basis Plus
6) Fulcrum Search Server

Each of these systems or products deserve full evaluation in the context of a conventional evaluation process.

By this point in our Summer Research Program, having identified these six from a field of twenty-two, we were confronted with the pressure of selecting and installing a product, loading CEDR data, and developing WWW interface forms, all within about six (6) weeks and with negligible funding or time resources to conduct any sort of convention procurement.

At this juncture, we received indications from a sister DOE laboratory, the Stanford Linear Accelerator Center (SLAC) near Palo Alto, of interest in inviting our Summer Research Project to use one of the six finalists candidates, SPIRES, using their computing resources, at no cost and with no procurement burden. Further, this alternative offered a Phase III approach, thus mitigating resource-intensive CGI development, a comparable computing platform to CEDR's (SunOS 4.1.3), and favorable licensing terms if migration to an LBNL platform were desired. Interestingly, SPIRES provides this Phase III capability, the performance enhancement of optimized reinitialization, without a proprietary httpd, but by instantiating the initialization functionality in S9 rather than in S2.

SPIRES, developed with DOE funding and the foundation for a large (DOE ER/SLAC/CERN/DESY/Japan) high energy physics bibliographic database, is recognized by portions of DOE (ER) as an appropriate tool for this type of application.

Finally, this alternative had the advantage that a project staff member had sufficient prior working knowledge of the system that there was a high probability that we could define CEDR databases, load CEDR data, and develop WWW interfaces within the tight time constraints. Additionally, we had established a good working relationship with SLAC staff over the past two decades which augured for good technical support when needed by the Summer Research Project staff.

SLAC's motivation for this offer was consistent with other cooperative work between SLAC and LBNL in the past: mutual benefit. SLAC was interested in our experience with developing WWW-based applications in SPIRES which they might use to improve the transition of their mainframe-based SPIRES high energy physics databases to a unix-based www-based environment.

Consequently, SPIRES was chosen from the six finalist candidate systems with which to implement our prototype system for the Summer Research Project.
5. Prototype Implementation

The prototype information system implemented in the Summer Research Project was a Phase III type information system as described above.

Three variations to the Phase III model are used in the WWW-based SPIRES system:

First, noted supra, the performance enhancement of optimized reinitialization is achieved without a proprietary httpd by instantiating the initialization functionality in S9 (stored procedure processor) rather than in S2 (httpd) or S5 (CGI).

Second, use of a proprietary httpd is further avoided by provision of a proprietary module which provides forms-DBMS request translation unique to SPIRES. This appears in figure 3 as component S10, DBMS-forms Interface.

Third, the network address of the httpd server (S2) and the network address of the machine running the CGI script (S5) are different, i.e., they are different machines. Consequently, the CGI function is encapsulated in a server (S5) which listens in the standard way for requests from S2, otherwise behaving as an ordinary Phase II CGI component (passing requests only), although some additional security features have been implemented in this CGI server.

The following table indicates the names used in the CEDR WWW-based SPIRES information system for each of the generic components:

<table>
<thead>
<tr>
<th>Component</th>
<th>Generic Name</th>
<th>CEDR-WWW-SPIRES</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>Platform</td>
<td>SunOS Release 4.1.3_U1 (SERVER-4M) (all versions of unix believed to be supportable)</td>
</tr>
<tr>
<td>S2</td>
<td>HTTP Daemon</td>
<td>Version 3.0_pre5 of the CERN server code is running on <a href="http://www.slac.stanford.edu">www.slac.stanford.edu</a> (the Unix production server).</td>
</tr>
<tr>
<td>S3</td>
<td>Page containing link to search form</td>
<td>not public at present</td>
</tr>
<tr>
<td>S5</td>
<td>CGI function (Phase II type)</td>
<td>encapsulated in John Halperin CGI-SPIRES server ~jxh/pgms/spires/jrunsrvr.csh (netcall.pl) (interpreted rather than compiled at present)</td>
</tr>
<tr>
<td>S6</td>
<td>DBMS</td>
<td>SPIRES, uSPIRES (unix)</td>
</tr>
<tr>
<td>S8</td>
<td>config files</td>
<td>stored in: /afs/slac.stanford.edu/www/spires/find/ cedrdfs cedrfile cedrvars cedrcode</td>
</tr>
</tbody>
</table>

When spiface receives the request from a form, then it consults the options files for that database, reforms the command to be a 'qspires' command and sends it on to Halperin's server.
S9  Stored procedure call processor
quSPIRES (a version of qSPIRES for unix)

S10  DBMS-Forms interface
SPIFACE v. 2.0

---

dashed line indicates S3 and S4 stored on spires.slac.stanford.edu but AFS-accessible from www.slac.stanford.edu.

The same configuration can be used for S8.

---

Figure 3
6. Results

1. We defined a file definition (schema) for the Epidemiology Guides comprising two 'databases':
   - Epiguides
   - EpiDE (the 'data elements' defined by HAI - a lookup table)

2. We converted three of the eight Epi Guides for loading, and loaded.

3. We developed prototype output formats, including for web presentation.

4. We implemented www-accessible forms for access to the Epidemiology Guides, for browsing, and for searching.

5. We defined a file definition (schema) for the CEDR metadata comprising five 'databases':
   - CEDRDFS (Data file set level metadata)
   - CEDRFILE (File level metadata)
   - CEDRVARS (Variable level metadata)
   - CEDRCODE (Codes metadata)
   - CEDRICD (Death Summaries)

6. We developed conversion routines to convert each type of metadata from CEDR dump format into database load format. This was problematic owing to the inconsistency of CEDR data and some of the CEDR dumping routines, and other eccentricities inherent in the current CEDR.

7. We converted all CEDR metadata resident in cedr.lbl.gov/data/cedrprod as of mid-July.

8. We loaded all the CEDR metadata and ICD summaries into their respective databases.

9. We established output formats for each metadata database, including for www access.

10. We established configuration files and www-accessible forms for access to all five databases.

11. We itemized a number of aspects of CEDR data deserving of QA/QC attention.

12. We characterized potential future work (Section 8).
7. Significance of this Work

All network access to CEDR material previously was to stored "flat files" each of which had to be separately maintained.

With this prototype, interaction with CEDR structured documentation is directly to the authoritative data in the database, giving these benefits:

1. The laborious load-dump-filter/overflow file procedure can eventually be abandoned.

2. Updated CEDR information is now accessible instantly - when the dbms record is updated, it is available on the web at that moment.

3. Multiple output formats all operate on single source - when CEDR metadata is updated, it can be instantly accessible in www format, CEDRtext format, and in phototypesetter format.

4. The HyperCEDRtext concept is now obsolete.
8. Possible Future Work

Transforming the Summer Research Project prototype to a production CEDR information system might require at least:

1. Additional output formats to generate CEDRtext files.
2. Development of a protocol to enable publish-on-demand CEDR catalogs.
   An important aspect of this work is that a CEDR staff member can update a metadata record, and the update is immediately available, and could be used to print a catalog rapidly. Such on-demand catalog printing might be triggered by receipt of email with a valid postal address.
3. Development of a CEDRtext applet:
   Currently, the only functional benefit of CEDRtext over www access is that CEDRtext only sends to the user/terminal that portion of the numerical data (i.e., data, not metadata) that fits in a small (15x80) window. Were the CEDR data files to be simply accessed as HTML, the larger CEDR files would exceed the capacity of the net and the client's browser.
   However, a Java-like applet might be developed with this portion of CEDRtext functionality, and would work as follows:
   - Information Seeker opens CEDR homepage
   - (registered) Information Seeker clicks on link to CEDR data
   - A CEDR applet is conveyed from server to client's browser
   - The applet provides a window with vertical and horizontal scroll bars. Manipulation of the scroll bars causes command language on the server to fetch and send to the client only those records that fit, thus avoiding the problem of sending whole CEDR files to the client.
4. Complete installation of the remaining Epi Guides.
5. Aesthetics. The on-screen presentation of both CEDR structured documentation and epidemiology guides can be improved.
6. Include HTML links in CEDR metadata to other metadata (Epiguide goal records 'data element' element now linked to the 'Epi data element' database).
7. Complete loading of all CEDR structured documentation.
9. Webware III (web-based listserv archive)
10. Webware V (instructional modules)
9. Summary of Deliverables

Online Deliverables:

2. CEDR Metadata: http://www.slac.stanford.edu/~lbl5spi/formcedr.html

Paper Deliverables:

1. LBNL-39163 Investigation of Network-Based Information System Model, Konrad, Perez, Rodriguez, Rivera, Durst, Merrill, Holmes. September, 1996 (this report).
Appendix A

DBMS System Evaluation Criteria

Our first four general requirements are:
   a) Organized Repository of Goal Records
   b) Generation and maintenance of Index Records derived from Goal Records
   c) Query of Index Records
   d) Presentation of Goal Records both by search result or directly (Report Generator)

Additionally, the following requirements relate to bibliographic database applications:

a) Text/Character manipulation capabilities.
   1) Breaking string into words for passing to index on query string.
   2) Force to upper case for passing to index and on query string.
   3) Change a string to different string.
   4) Insert text in a string.
   5) Return size of data element value.
   6) Squeeze leading and trailing blanks.
b) Concurrent update control.
c) Optionally-occurring elements
   1) The data elements (variables) use storage or cells having no value?
d) Dates.
   1) What's the DATE FORMAT?
   2) Is it possible to change the DATE in one format to different output format?
e) Pre-defined procedures
   1) What Pre-defined procedures process values as input to DB?
   2) What Pre-defined procedures process output or displayed values from DB?
   3) What Pre-defined procedures process index values?
   4) What Pre-defined procedures process search values?
f) Personal name processing.
   1) How does DB store personal names?
   2) How does DB retrieve personal names?
   3) How does DB process and present personal names sub-components?
g) Variable length elements.
   1) Can all data elements in a DB be variable length?
   2) How is storage of variable length values supported?
h) Data repository.
   1) Is the data stored in the DB or externally?

i) Search engine.
   1) Complexity of query language.
   2) Report Generator
   3) Web/http (CGI) interface.
   1) Description of DBMS for interoperability with CGI scripts.
Appendix B

Evaluation of products and systems

**Microsoft Access**

Name of the manufacturer: Microsoft Corporation  
Name of the vendor: Microsoft Corporation  
Cost: $399.00  
Type: Relational Database

Summary: Microsoft Access 2.0 is a good DBMS for Windows, Windows 95 and Windows NT. There is Web access for Windows NT server. No Web access was found for Windows or Windows 95. Examples of features were provided but no implementation details were available.

**Topic Internet Server**

Name of the manufacturer: Verity Inc.  
Name of the vendor: Verity Inc.  
Cost:  
Type: Search Engine

Summary: Topic Internet Server is a product of topic SEARCH which is part of Verity's topic family of products. No detailed paper documentation was provided and there was no answer to telephone calls and mail messages.

**Postgres 95**

Name of the manufacturer: Andrew and Jolly Chen  
Name of the vendor: University of California  
Cost: Free  
Type: Relational Database

Summary: Terminated evaluation at criterion number 5 (Optionally-occurring elements) on lack of merits:

1) Absence of product support  
2) Absence of text capability  
Resources do not permit further review.

**WAIS**

Name of the manufacturer: Thinking Machines, Inc.  
Name of the vendor: CNIDR  
Cost: Free  
Type: Information retrieval system
Summary: Free WAIS is an Information retrieval system that implements free text search using a client/server architecture. Since free WAIS is not a DBMS there is no further consideration of this product.

**Basis plus**

Name of the manufacturer: Information Dimension Inc.
Name of the vendor: Information Dimension Inc.
Cost: 
Type: Text DBMS

Summary: BASIS plus is a great text DBMS with almost all the characteristics CEDR is looking for. The only characteristic it don't has is the personal name processing. There was no answer to the telephone call to check on the cost of the product. Further consideration of the of this product is justify. No paper documentation was provided.

**System II**

Name of the manufacturer: Sybase Inc.
Name of the vendor: Sybase Inc.
Cost: 
Type: RDBMS

Summary: System II is a family of Database products that runs on a variety of platforms from PCs to multi-cpu super servers. Including UNIX and SunOS. Some information was received by fax, however no detailed documentation about the DB was available.

**SMART**

Name of the manufacturer: Gerard Salton
Name of the vendor: Cornell University
Cost: Free
Type: Not a DB, Search Engine

Summary: There is not enough documentation available, because it is necessary to install the system to get reasonable documentation. The system is only a search engine.

**MG Information Retrieval System**

Name of the manufacturer: I.H. Witten, A. Moffat & T.C. Bell
Name of the vendor: 
Cost: Free
Type: Not a DB, Information retrieval system

Summary: The information available is only to get the free software. To obtain documentation it is necessary to install the system or buy their new book. It is a research prototype, not a production-caliber product. Evaluation was terminated.
**FFW-Free text search for WWW**

Name of the manufacturer: Multitorg project  
Name of the vendor: Telenor R & D, Norway  
Cost: free for non-commercial use  
Type: Not a DB, is a Search Engine

Summary: The system it's only a search engine. It is a package made to provide easy-to-use searching facilities over HTML documents. Evaluation terminated.

**RDB**

Name of the manufacturer: Walter V. Hobbs  
Name of the vendor: RAND  
Cost: Free  
Type: Compliant with the relational model

Summary: Text/character manipulation capabilities are not supported by RDB. PERL could be used or UNIX utilities. This is a very simple relational DBMS. I think that using this DBMS with a good search engine would be a great help. [J. Rivera]

**Iq-text**

Name of the manufacturer: Liam R. E. Quin  
Name of the vendor: University of California at Berkeley  
Cost: Free  
Type: Not a DB. Is a search engine.

Summary: Iq-text is a text retrieval engine. The system is not what we need. Does not comprise an organized repository (DBMS requirement).

**gt (Query Text)**

Name of the manufacturer: John Conover  
Name of the vendor: (organization) Vixie Enterprises  
Cost: Free  
Type: Not a DBMS; search engine only.

Summary: The documentation sent by Mr. Conover was very good. The system is a text information retrieval system. It creates, maintains and queries a full DB. Evaluation terminated. I believe this search engine together with RDB will be a good help. [J. Rivera]
**ORACLE**

Name of the manufacturer: Oracle Corp.
Name of the vendor: Richard Franceschini
Cost: Relational DB

Summary: Oracle does not support bibliography data nor text retrieval capability well within the DBMS. Oracle has a software called Context that does not require the Oracle server in the current version. The next version of Context will require the proprietary server. Context is stand-alone software that needs a program (C language) to send the information. Context has many features. Context is a natural language processing technology that identifies themes and content in English text. Evaluation terminated, because original programming effort is beyond our scope and due to procurement cost.

**ESQL**

Name of the manufacturer: Seekersoft
Name of the vendor: Seekersoft
Cost: Relational DB
Type: WWW (CGI) interface to DBMS from HTML documents.

Summary: ESQL is a interface between HTML and a DBMS that uses SQL. The SQL statement can be used inside the HTML code.

**Glimpse 3.0**

Name of the manufacturer: Udi Manber and Burra Gopal
Name of the vendor: University of Arizona and National Chung- Cheng University
Cost: Relational DB
Type: Global Implicit Search (Search engine)

Summary: Glimpse is an indexing and query system that allows you to search through all your files very quickly. Glimpse supports most agrep’s options including approximate matching, Boolean queries, and even some limited forms of regular expressions. Does not offer an organized data repository.

**AIRS II**

Name of the manufacturer: Arachnae Management Limited
Name of the vendor: Arachnae Management Limited
Cost: Server $5000, each concurrent user/client $100
Type: Full-text DB server

Summary: At the present, there is not enough information on this product, but it should be considered in the future for its important features that meet with some of our demands. Although the people in charge of this product were contacted by email and phone, we were not able to receive enough information as to clear up our difficulties found in the criteria evaluation.
**Fulcrum Search Server**

Name of the manufacturer: Fulcrum Technologies Inc.
Name of the vendor: Fulcrum Technologies Inc.
Cost:  
Type: Search Engine

Summary: Fulcrum Search Server-a high performance, multiplatform, indexing and retrieval server engine. It is your key to retrieving and publishing information across the enterprise information critical for effective decision-making. Does not support data repository.

**MORE- Multimedia Oriented Repository Environment**

Name of the manufacturer: David Eichman, Terry McGregor and Dann Danley
Name of the vendor: University of Houston
Cost: Free  
Type: Metadata-based repository system

Summary: Terminated evaluation on lack of merits:  
1) Absence of support  
2) Absence of specific information  
3) MORE is a metadata base repository- the information in its underlying DBMS isn't the objects themselves, but rather information concerning the object, which is stored using other mechanisms.

**Spires**

Name of the manufacturer: Stanford University
Name of the vendor: Stanford University/SLAC
Cost: Pending  
Type: full bibliographic DBMS with network access

Summary: Candidate for further consideration.

**BRS/SEARCH/Netanswer**

Name of the manufacturer: Dataware Technologies
Name of the vendor: Dataware Technologies
Cost: $15,000  
Type: full bibliographic DBMS with network access

Summary: Candidate for further consideration.

**ISITE/ISEARCH**

Name of the manufacturer: Center for Networked Information Discovery and Retrieval (CNIDR)
Name of the vendor: Same
Cost: freeware  
Type: still in beta
# Appendix C

## Table of Products Evaluated
*(interim)*

<table>
<thead>
<tr>
<th>Data Base</th>
<th>Evaluator</th>
<th>Evaluation Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFW</td>
<td>Jose</td>
<td>a- Search Engine only</td>
</tr>
<tr>
<td>Glimpse 3.0</td>
<td>Madel</td>
<td>a- Search Engine only</td>
</tr>
<tr>
<td>1q-text</td>
<td>Jose</td>
<td>a- Search Engine only</td>
</tr>
<tr>
<td>MG</td>
<td>Jose</td>
<td>a- Search Engine only</td>
</tr>
<tr>
<td>qt (Query Text)</td>
<td>Jose</td>
<td>a- Search Engine only</td>
</tr>
<tr>
<td>SMART</td>
<td>Jose</td>
<td>a- Search Engine only</td>
</tr>
<tr>
<td>Topic Internet Server</td>
<td>Yamilet</td>
<td>a- Search Engine only</td>
</tr>
<tr>
<td>WAIS</td>
<td>Yamilet</td>
<td>a- Search Engine only</td>
</tr>
<tr>
<td>ORACLE</td>
<td>Jose</td>
<td>b- Doesn't support text Processing</td>
</tr>
<tr>
<td>ORACLE with CONTEXT</td>
<td>Jose</td>
<td>c- Resource Intensive</td>
</tr>
<tr>
<td>MORE</td>
<td>Madel</td>
<td>d- Non Responsive</td>
</tr>
<tr>
<td>Postgres 95</td>
<td>Yamilet</td>
<td>d- Non Responsive</td>
</tr>
<tr>
<td>System II</td>
<td>Yamilet</td>
<td>d- Non Responsive</td>
</tr>
<tr>
<td>ESQL</td>
<td>Jose</td>
<td>e- RDB Support Tools</td>
</tr>
<tr>
<td>AIRS II</td>
<td>Madel</td>
<td>f- Good Candidate</td>
</tr>
<tr>
<td>Basis Plus</td>
<td>Yamilet</td>
<td>f- Good Candidate</td>
</tr>
<tr>
<td>BRS/Search</td>
<td>Madel</td>
<td>f- Good Candidate</td>
</tr>
<tr>
<td>Fulcrum Search Server</td>
<td>Madel</td>
<td>f- Good Candidate</td>
</tr>
<tr>
<td>Microsoft Access</td>
<td>Yamilet</td>
<td>f- Good Candidate</td>
</tr>
<tr>
<td>Spires</td>
<td>Madel</td>
<td>f- Good Candidate (Selection)</td>
</tr>
<tr>
<td>RDB</td>
<td>Jose</td>
<td>g- RDB without indexing capabilities</td>
</tr>
</tbody>
</table>
Appendix D

Selected Screens from WWW Access to CEDR Structured Documentation
CEDR - Comprehensive Epidemiologic Data Resource

Information Retrieval

DFS Database:

- Browse Data File Set Structured Documentation
- Search Data File Set Structured Documentation

File Database:

- Browse File Structured Documentation
- Search File Structured Documentation

Var Database:

- Browse Variable Structured Documentation To browse structured documentation for variables, browse via Data File Set for File level structured documentation.
- Search Variable Structured Documentation

Code Database:

- Browse Code Set Structured Documentation
- Search Code Set Structured Documentation

ICD Summary Database (Death Tables):

- Browse ICD Summaries
- Search ICD Summaries

This CEDR information retrieval facility developed by Madel Perez, Jose Rivera and Yamilet Rodriguez.

We welcome your comments and questions concerning the CEDR project, our information retrieval products and Customer Service. If you would like further information about CEDR, please contact us at:

Allan M. Konrad
email: konrad@sims.berkeley.edu
phone: (510) 486-5458
mail: Mailstop 50B-3238
Lawrence Berkeley National Laboratory
Berkeley, CA 94704
**CEDR DFS Database Search Form**

Fill in the entries you know, leave the others blank.

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset:</td>
<td></td>
</tr>
<tr>
<td>Name:</td>
<td></td>
</tr>
<tr>
<td>AltName:</td>
<td></td>
</tr>
<tr>
<td>Description:</td>
<td></td>
</tr>
<tr>
<td>Races:</td>
<td></td>
</tr>
<tr>
<td>Sexes:</td>
<td></td>
</tr>
<tr>
<td>Diseases:</td>
<td></td>
</tr>
<tr>
<td>Exposure-Agent:</td>
<td></td>
</tr>
<tr>
<td>Sites:</td>
<td></td>
</tr>
<tr>
<td>Contact-Name:</td>
<td></td>
</tr>
<tr>
<td>Subject:</td>
<td></td>
</tr>
<tr>
<td>Added-Keywords:</td>
<td></td>
</tr>
<tr>
<td>CEDRQAQC:</td>
<td></td>
</tr>
</tbody>
</table>

[Submit Query] [Clear Fields]
CEDR File Database Browse Form

FRC94A02 DLG4014 Fernald demographics and work history
FRC94A02 NLOEXTL External dosimetry analysis file
FRC94A02 NLOINTL Fernald internal doses
HFC7BA01 CEDR78_1HF CEDR78 F#1: Vital Stat. & Int. Depos.
HFC7BA01 CEDR78_2Hanford CEDR78 File 2: External Dosim.
HFC7BA01 CEDR78_3Hanford CEDR78 File 3: Additional Vars
HFI89A01 IARC89_1HF IARC89 F#1: Vital. Stat. & Int. Depos.
HFI89A01 IARC89_2Hanford IARC89 F# 2: External Dosimetry
HFI89A01 IARC89_3Hanford IARC89 File 3: Additional Vars
HFMCCA02 MALCCA_1
HFMCCA02 MALCCA_2
HFMCCA02 MALCCA_3
HFMCCA02 MALCCA_4
HFMCCA02 MALCCA_5
HFMPYA02 MALPYA1
HFMPYA02 MALPYA2
HFMPYA02 MALPYA3
HFMPYA02 MALPYA4
HFMPYA02 MALPYA5
HFSRCW01 INDEHanford internal deposition
HFSRCW01 OHH88_COHanford job histories - construction
HFSRCW01 OHH88_OPHanford job histories - non-construction
HFSRCW01 ORE_44Hanford external dosimetry 1944-1982
HFSRCW01 ORE_83Hanford external dosimetry 1983-1989
HFW89W01 ADD89_1
HFW89W01 ADD89_2
HFW89W01 ADD89_3
HFW89W01 DOS89Hanford DOS89: External Dosimetry
HFW89W01 INT89Hanford INT89: Internal Deposition
HFW89W01 JOB89Hanford JOB89: Job Histories
LAFEMh01 LAEEFILELANL females 1987 analytic file
LAMENA03 LAMEFILELANL males 1993 analytic file
LASUIA02 LASUFILELANL females suicide analytic file
LAUPUA01 UPPU8926 Plutonium-Exposed LANL Workers--25-yr
MCD94A01 ELL2542 Analysis file for mcd94a01
MDFEXTA02 ANFILE mound JOM91 Paper Analytic File
MFD94A01 CASE45Cumulative (case?) internal dose info.
MFD94A01 CHRONSMK Smoking hist. codes and dates (CHRONSMK)
MFD94A01 CONT45Cum. (control?) int. dose info. (CONT45)
MFD94A01 DASSEXT Average daily ext exp, days wkd (DASSEXT)
MFD94A01 DAYSINT Average daily int exp, days wkd (DAYSINT)
MFD94A01 DEMGREM Dose, work, med, and smoking (DEMGREMP)
MFD94A01 EXPCODE Thorium, radium, and radon expos (EXPCODE)
MFD94A01 MEDREC Detailed medical data (MEDREC)
MFD94A01 NEWEXT Film badge, doses, and lag cutoff (NEWEXT)
MFD94A01 NEWINTCalc'd lung doses and lag cuts (NEWINT)
MFD94A01 PAYCODES Pay code and job title data (PAYCODES)
MFD94A01 SMK1580 Smoking codes and dates by fac (SMK1580)
MFP94A01 ELL31455 rem case file used in mortality analys
MFT93A01 HPIARC_1 Identification and internal exposure inf
MFT93A01 HPIARC_2External doses information
MFT93A01 HPIARC_3Workers with internal exposure
MFT93A01 HPIARC_4Additional variables
MFT93A01 ORIARC_1ORNL IARC general
**CEDR File Database Search Form**

Fill in the entries you know, leave the others blank.

<table>
<thead>
<tr>
<th>Field</th>
<th>Input Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>FILE:</td>
<td></td>
</tr>
<tr>
<td>Dataset:</td>
<td></td>
</tr>
<tr>
<td>File-Name:</td>
<td></td>
</tr>
<tr>
<td>Description:</td>
<td></td>
</tr>
<tr>
<td>Analysis-Unit:</td>
<td></td>
</tr>
<tr>
<td>Subject:</td>
<td></td>
</tr>
<tr>
<td>Added-Keywords:</td>
<td></td>
</tr>
</tbody>
</table>

[Submit Query] [Clear Fields]
Appendix E

Selected Screens from WWW Access to Epidemiology Guides
(site archive records)
Epidemilogic Guides

Prepared for the U.S. Department of Energy
Office of Epidemiology and Health Surveillance

By
History Associates Incorporated
Rockville, Maryland

Browsing Epidemilogic Guides

- Oak Ridge National Lab and Oak Ridge Operations Office: Records Relating to RaLa, Iodine-131, and Cesium-137 at the ORNL and the Oak Ridge Operations Office

Searching in Epidemiologic Guides

Search Epiguides

Searching Epi Data Element Database:
Search Epide

Record Series for entire DOE sites
Record Series for DOE sites (not confined to epidemiology)
This CEDR information retrieval facility developed by Madel Perez, Jose Rivera and Yamilet Rodriguez.

We welcome your comments and questions concerning the CEDR project, our information retrieval products and Customer Service. If you would like further information about CEDR, please contact us at:

Allan M. Konrad
e-mail: konrad@sims.berkeley.edu
phone: (510) 486-5458
mail: Mailstop 50B-3238
Lawrence Berkeley National Laboratory
Berkeley, CA 94720

Updated 05 September 1996.
### CEDR Epidguides Database Search Form

Fill in the entries you know, leave the others blank.

<table>
<thead>
<tr>
<th>Control:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TocSubElem:</td>
<td></td>
</tr>
<tr>
<td>Introduction:</td>
<td></td>
</tr>
<tr>
<td>Series-Type:</td>
<td></td>
</tr>
<tr>
<td>Location:</td>
<td></td>
</tr>
<tr>
<td>OtherID:</td>
<td></td>
</tr>
<tr>
<td>Originating:</td>
<td></td>
</tr>
<tr>
<td>Program:</td>
<td></td>
</tr>
<tr>
<td>Finding-AIDS:</td>
<td></td>
</tr>
<tr>
<td>Bibliography:</td>
<td></td>
</tr>
<tr>
<td>Disposition:</td>
<td></td>
</tr>
<tr>
<td>Series-Desc:</td>
<td></td>
</tr>
<tr>
<td>Data-Element:</td>
<td></td>
</tr>
<tr>
<td>Computing-Env:</td>
<td></td>
</tr>
<tr>
<td>Directory:</td>
<td></td>
</tr>
<tr>
<td>File-Name:</td>
<td></td>
</tr>
<tr>
<td>File-Type:</td>
<td></td>
</tr>
</tbody>
</table>

[Submit Query] [Clear Fields]
HANFORD SITE: A GUIDE TO RECORD SERIES
SUPPORTING EPIDEMIOLOGIC STUDIES CONDUCTED FOR THE DEPARTMENT OF ENERGY

INTRODUCTION
Overview
The Epidemiologic Records Inventory Project
Role of HAI

BACKGROUND
History of the DOE
more
more
Site History
Organization
Pacific Northwest Laboratory (PNL)
Hanford Environmental Health Foundation (HEHF)
Occupational Epidemiologic Studies
Breast Cancer Incidence Study
Epidemiologic Surveillance System (ESS)
DOE Health and Mortality Studies
AEC Feasibility Study
AEC Pilot Study
AEC Health and Mortality Study
more
Lung Cancer, Ionizing Radiation, and Tobacco Smoking Among Males at Hanford Case-Cohort
Goodtime, Trauma, and Timeout Studies

ACCESS
Overview
Records Locations
Records Retention and Disposition Schedules
DIDS Database
RHA-NIS Database
NARA Standard Form 135a

METHODOLOGY
Data Elements

SCOPE
more
Administrative Contacts
U.S. Department of Energy
U.S. Department of Energy
Seattle Federal Records Center
National Archives and Records Administration - Pacific Northwest
National Archives and Records Administration

ARRANGEMENT
I. Administrative and General
II. Database Management
III. Data Collection
IV. Data Analysis
V. Electronic Record Series
more

DATA ITEMS IN RECORD SERIES DESCRIPTIONS
Record Series Descriptions
Access Restrictions
more
Accession/Other Identification Number
more
Arrangement
Condition
Container Number
Data Elements
Disposition Authority
Electronic Record Series Descriptions

Title and Inclusive Dates

Atomic Energy Commission (AEC) Health and Mortality Study (HMS) Progress Reports, 1962
Hanford Environmental Health Foundation (HEHF) Annual Reports, 1958-1971
Hanford Environmental Health Foundation (HEHF) Monthly Reports, 1974-1989
Hanford Exposure Project (HEX) Log, 1967-1968
Hanford Exposure Project (HEX) Records, ca. 1940-1974
Hanford Health and Mortality Study (HMS) Administrative Files, 1964-1989
Hanford Health and Mortality Study (HMS) Advisory Committee Meeting Records, 1982-1989
Industrial Medical Services Section Procedures, 1945, 1946, 1955-1958
Inspector General Investigation Records, 1977
Joint Epidemiology Group Meeting Minutes, 1984-1990
Mancuso Correspondence and Study Records, 1965-1978
Occupational Program Records, 1991
Pacific Northwest Laboratory (PNL) General Records, 1979-1992
Skin Cancer Project Study Records, 1987
Timeout Study Records, 1986-1989
Trauma Study Planning Records, 1989
Analysis (ANAL.89) File Edit Records, 1991
Center for Epidemiologic Research (CER) Database Correlation Records, 1993
Comprehensive Epidemiologic Data Resource (CEDR) Database Correlation Records, 1993
Goodtime Study Database Code Records, 1989
Hanford Health and Mortality Study (HMS) Database Documentation Records, ca. 1989-1995
Hanford Mortality (HMO) File Documentation Records, 1979, 1984-1994
Health Surveillance System (HSS) Query Records, 1990-1992
International Agency for Research on Cancer (IARC) Database Correlation Records, 1991
Key File Edit Records, 1992
Mortality and Occupational Exposure (MOCX) File Documentation Records, 1987
Mortality Study Database (MORX) Modification Records, 1991-1993
Timeout Study Database Records, 1987-1989

E-6
Benton and Franklin Counties Monthly Mortality Reports, 1988-1993
Benton County Voter Report, 1981
Blood Type Record Cards, 1964
Breast Cancer Study Medical Records, 1984-1989
Building Lists, 1967
Construction Employees' Roster, 1985-1987
Contractor Employee Roster Records, 1950-1953
Cooperative Program Personnel Questionnaires, 1992
Cumulative Radiation Exposure Records, 1969
Death Certificate Records, 1944-1994
External Exposure Data Records, 1983-1989
Goodtime Study Alcohol, Tobacco, and Exercise Data Abstract Forms, 1984-1992
Hanford Area Autopsy Records, undated
Hanford Atomic Power Operations (HAPO) Personnel Data, 1964 LOCI=RHA, 712 Building, F
Hanford Environmental Health Foundation (HEHF) Death Certificate Retrieval Records, 1
Health Medical Examination (HME) Report Records, 1944-1976
Health Surveillance System (HSS) Health Event Audit Records, 1984-1993
Health Surveillance System (HSS) Health Event Monthly Reports, 1984-1990
Health Surveillance System (HSS) Health Event Records, 1987-1993
Injury and Accident Reports, 1953-1960
Internal Exposure Records, 1974-1976
Lung Cancer Study Records, 1983-1994
Master File Records, 1988
Medical Chart Logs, 1966-1967
Mortality Data Validation Report, 1993
Mortality Record Cards, 1981
Occupational Health History Master Files, ca. 1944-1979
Operations and Construction Workers Roster Records, undated
Organizational Charts and Directories, 1944-1986
Organizational Code Directory, 1969
Public Health and Social Services Records, 1963
Social Security Administration (SSA) Employee Data Records, 1967-1972
Timeout Study Alcohol, Tobacco, and Exercise Data Abstract Forms, ca. 1989
Trauma Study Alcohol, Tobacco, and Exercise Data Abstract Forms, 1989
Tumor Registry Project Records, 1987
United States Death Rate Records, 1985, 1988
X-ray Exposure Study Records, 1968, 1982
Health Surveillance System (HSS) Basic Analysis Reports, 1985-1991
Health Surveillance System (HSS) Hanford Roster Analysis, 1987-1992
Leukemia Data Correlation Records, 1992
Medical Department Monthly Reports, 1948-1953
Mortality and Occupational Exposure (MOX) and Monson Reports, 1988-1991, 1993
Plutonium Finishing Workers Profile Records, 1991
Radiation Exposure Data Records, 1990
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Master Files
Personal Information Files

Work History Files

Occupational Radiation Exposure Files

Vital Status Files