IMPLEMENTING AND OPERATING THE HANFORD ENVIRONMENTAL INFORMATION SYSTEM AND APPLYING IT TO THE CARBON TETRACHLORIDE EXPEDITED RESPONSE ACTION

Paula J. Cowley and George V. Last
PNL*
PO Box 999
Richland, WA USA 99352

Michael R. Schwab and Virginia J. Rohay
WHC**
PO Box 1970
Richland, WA USA 99352

ABSTRACT

To manage waste and perform environmental monitoring and restoration at the 1,450-square kilometer (560-square mile) Hanford Site in southeastern Washington State, vast amounts of scientific and technical data are being generated from sampling. The Hanford Environmental Information System (HEIS), operational for over two years, is a computerized system designed and implemented to manage the Site's environmental sampling data in support of the Hanford mission. HEIS serves as a repository for storing the data and includes tools for efficient data processing, tools for mapping and spatial data analysis, maps of the Hanford Site, and software for entering, viewing, reporting, and depicting data. HEIS continues to evolve as new types of data and capabilities are incorporated in response to the user community.

Lessons learned from putting HEIS into operation include adapting to a data management culture, data validation and defensibility, benefits of electronic data loading, and garnering management support.

HEIS is being applied to the carbon tetrachloride expedited response action being performed at the Site. Data collection and analysis is designed to maximize use of field screening methods and existing data to meet the accelerated cleanup schedule.

DESIGNING AND IMPLEMENTING HEIS

The Hanford Environmental Information System (HEIS) has been designed and implemented to store, manage, and aid in the analysis of Hanford's environmental sampling data, which are crucial to Hanford's mission of environmental cleanup and restoration. HEIS is a repository for storing the data and includes tools for efficient data processing, tools for mapping and spatial data analysis, maps of the Hanford Site, and software for entering, viewing, reporting, and depicting data. Major HEIS components have been operational for over two years. Data are being added at an increasing rate, with approximately 350,000 chemical and physical property records being entered during calendar year 1992. HEIS continues to evolve as new types of data and capabilities are incorporated in response to the user community.

Figure 1 shows the HEIS hardware and network configuration. The HEIS database runs on a Sequent multiprocessor computer that is dedicated to support the HEIS database. The Sequent is available either through the Hanford local area network on individual desktop computers or via phone-in modems. Therefore, the database is accessible to authorized persons working at the Hanford Site as well as off-site users. HEIS also includes a Geographic Information System (GIS) to display and analyze data on a map. The GIS software runs on Sun workstations. Maps are stored on the Sun's local hard disk, which renders map data readily available and able to be drawn quickly without overloading the network.

Early in the design process, it was apparent that no single commercial software product would satisfy HEIS' diverse requirements. Off-the-shelf commercial software was selected and integrated with a Hanford-specific design based on user requirements. Commercial software being used includes ORACLE as the relational database management system for storing large volumes of sampling and analytical results data, ARC/INFO as the GIS, Uniface as the database user interface for data entry and data viewing, SQR Workbench for reporting and data extraction, and GNUPLOT for graphics. The HEIS developers have also written some software in the "C" programming language. The challenge for software developers has been to integrate these diverse products into an environment that can meet user needs.

The system currently includes ground-water, geologic, biota, and well construction data and has capabilities for soil gas, atmospheric, survey, and geophysical data. Subsystems have been implemented for scheduling and tracking samples. New capabilities for storing tank characterization data (i.e., chemical and physical properties of samples taken from waste tanks) and gaseous effluent data are also being implemented. For each type of data added to the database, the HEIS software development team conducts a user requirements analysis, a formal data modeling process, an implementation and documentation process, and internal and independent testing phases before the software is made available to the user community. As the system grows and expands, we are continually faced with database and GIS design decisions. For

* Pacific Northwest Laboratory is operated for the U.S. Department of Energy by Battelle Memorial Institute under Contract DE-AC06-76RLO 1830.
** Westinghouse Hanford Company is the Hanford Operations and Engineering Contractor for the U.S. Department of Energy under Contract DE-AC06-87RL10930.

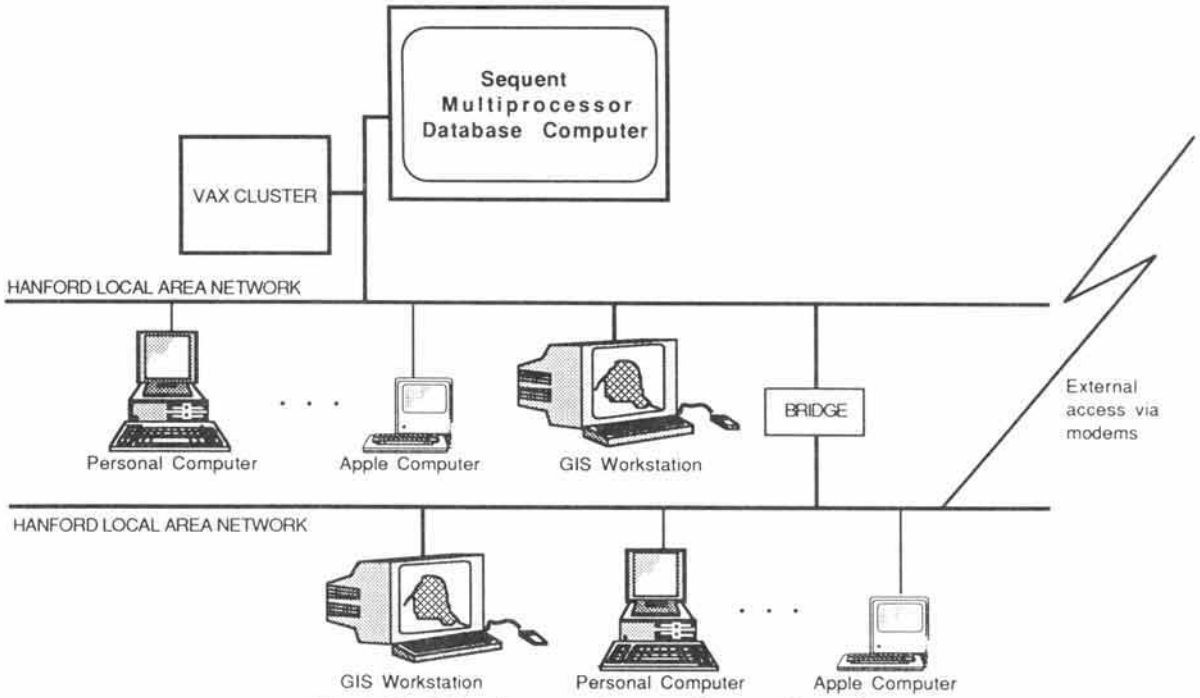

Fig. 1. HEIS hardware and networking configuration.

example, we are currently considering whether raw or interpreted data from geophysical summary logs should be included in the database and where to store information about spatial data.

HEIS' geographic information system represents a major tool for the display and analysis of HEIS data. Besides utilizing the capabilities supplied by ARC/INFO, the HEIS team has developed a graphical user interface that provides GIS users with access to many powerful capabilities without requiring them to become GIS experts. The user interface provides user-friendly capabilities for building custom maps, selecting and subsetting data, generating surfaces and contours, importing data, creating output products, and accessing the HEIS database.

The core HEIS software development team consists of Pacific Northwest Laboratory (PNL) computer scientists, some of whom have extensive experience in environmental information management. These developers work closely with scientists and engineers in the various disciplines covered by HEIS. Significant software or data structure changes proposed by either the user community or software developers are presented to the HEIS Configuration Control Board (CCB) for approval. The board consists of senior staff from the Westinghouse Hanford Company (WHC) and PNL who represent the various programs involved in HEIS. Invariably, the proposals emerging from the CCB are stronger than they had been when presented. We believe that one of the strengths of *HEIS* is teaming a strong computer science capability with strong technical expertise in various subject areas.

OPERATING HEIS

The HEIS database has been in operation for more than two years. Ground-water site-wide environmental surveillance and *Resource Conservation and Recovery Act of 1976* (RCRA) monitoring data from an earlier database were converted to HEIS. The converted data cover monitoring efforts since the 1950's, including ground-water scheduling and tracking software. New capabilities to handle *Comprehensive Environmental Response, Compensation and Liability Act of 1980* (CERCLA) ground-water, geologic, biota, and well construction data were added. During the first two years, the amount of ground-water data in HEIS nearly doubled, primarily due to an aggressive CERCLA site characterization effort.

Our Data Management Vision

Our vision is to have HEIS be a resource that provides the right data in the right form at the right time to the right place at the right cost. We have found HEIS to be the most effective means of implementing this data management vision. A number of lessons learned have emerged during the two years of operation. In the rest of this section, we discuss some of those lessons in the hope that other data management organizations may avoid similar problems.

Adapting to a Data Management Culture

We did not anticipate the huge culture change that data management would mean to the Hanford Site. Scientists and engineers at Hanford were accustomed to gathering data, analyzing and preparing reports about the data, and then filing the data on hard copy. If another person wanted to use the same data, that person was responsible for finding who currently had the data and then quizzing the current file owner on the assumptions made when the data were gathered. This led to an environment in which it was often easier and even more cost-effective to collect new samples and generate new data. Because of the secrecy associated with Hanford's production missions, data sharing was not encouraged. Now that Hanford has a single, open mission of environmental restoration, data sharing is encouraged and even expected. However, it has been difficult among Hanford scientists and engineers to implement the philosophy of valuing and sharing data instead of treating data as single-use, disposable commodities. We have made significant progress, but it has proven more

challenging than anticipated to effect change in the way people work to accommodate a data management infrastructure.

Implementing automated sample and data tracking in HEIS is a case where we are making progress in changing the culture. When only a few hundred samples and their associated results were being tracked, a project manager could track the data manually or with a simple spreadsheet. Once thousands of samples and data needed to be tracked, the task became overwhelming. Now sample information can be put into HEIS when the sample is taken and sent to the laboratory. When the results are returned from the lab and entered into the database, HEIS automatically marks the expected results as being received so that it is easy to determine what is still outstanding. Results data can also be used to determine laboratory performance, such as whether the results have been returned on schedule. This sample tracking lesson was learned the hard way by many project managers who had no idea where their samples and results were.

HEIS has the ability to generate unique identifying numbers for each sample. Users can request a block of sample numbers for an activity and use those sample numbers as needed. Unused numbers can be returned for reassignment. Initially there was resistance to using the HEIS numbers because staff were accustomed to assigning their own numbers. However, significant confusion arose when the staff were unable to maintain unique numbers. Use of unique HEIS sample numbers has now become a standard way of doing business on the Site. Even some programs that don't put their data into HEIS are considering using HEIS sample numbers.

Tracking missing and/or incomplete data is a difficult and time consuming process. We have been working to change the culture so that it becomes part of the staff's normal process to deliver data taken in the field to the responsible organization for entry into HEIS. To track incomplete data, we log in all data received and maintain an "open items" log of data that are not quite ready to be entered into HEIS.

Electronic versus Manual Data Loading

Most laboratories Hanford uses have the ability to deliver analytical results in electronic data files on diskette. Only 10% of the data we receive are not in electronic form, but processing the data when only paper is received takes roughly 70% of the data processing staff's time. Our experience has shown that it takes 20 times the staff resources to key in and verify non-electronic data compared to the time required to load and verify data received electronically. Besides the extra time needed to manually key the data into the system, we have found a 100% verification of manually entered data to be required since 1-5% of the records entered manually contain errors. With electronic data loading, the only verification required is to ensure that the process ran to completion satisfactorily.

Issues of Data Validation and Defensibility

The HEIS user community includes the Washington State Department of Ecology (Ecology) and the U.S. Environmental Protection Agency (EPA). These two organizations are the "regulators" who perform an oversight function to ensure that the U.S. Department of Energy (DOE) and its contractors are working in accordance with the *Hanford Federal Facility Agreement and Consent Order* signed by DOE, Ecology, and EPA. The regulators have access to the HEIS database and use it as a tool in performing their oversight function. The database is configured so that all users have access to verified and validated data, while only those responsible for the data have the ability to change the data and have access to unverified, unvalidated data. A special security scheme allows those who are responsible for the data to determine when the regulators will be able to view the data. Measures are under way to accelerate the validation process so that data can be made available to the regulators in a more timely manner. The validation process is performed outside of HEIS, although we are investigating how HEIS could be used to expedite the process.

We are still struggling with the issue of what constitutes "legally defensible" data. We have tried to design the system so that the data can be used in a court of law. We know that a decision to "close" a particular site and consider it cleaned up will be largely based on the data. But what is sufficient and necessary? What extra steps must be taken when the data must remain available throughout the 30-year Hanford cleanup? We know the data must be of known quality and we include a significant amount of "metadata" (i.e., data about the data) that describe how the sample was analyzed. We know that data must be traceable to the source. Once again, metadata have been included to support this requirement. In addition, "change records" are generated to support traceability whenever a result record is modified. The database will undoubtedly evolve as we come to understand the legal issues better.

Gaining Management Support

A system as large and complex as HEIS with its Configuration Control Board, its extensive user and operator manuals, its requirements analysis, and its software development process is expensive to establish and maintain. The system itself grows as more user requirements are identified. Hardware and software maintenance costs are ongoing. There are also operational costs of keeping the HEIS-specific software up to date as advances are made in the computer field. Strong management support is required if a system like HEIS is to be successful. Management must be made cognizant of HEIS' role in cleaning up the Hanford Site. HEIS' multi-contractor project team is constantly working to gain management support and ensure that the resources we have are being used effectively to benefit the entire HEIS community.

APPLYING HEIS TO HANFORD'S CARBON TETRACHLORIDE EXPEDITED RESPONSE ACTION

The 200 West Area Carbon Tetrachloride Expedited Response Action (ERA) is an environmental restoration project to remove carbon tetrachloride vapor from unsaturated soils using a vapor extraction system. The ERA is a CERCLA provision that allows accelerated cleanup activities to be undertaken prior to completion of a remedial investigation/feasibility study and the record of decision. Site characterization and performance baseline data are being collected to optimize the removal of carbon tetrachloride. Data collection and analysis are designed to maximize use of field screening methods and existing data to meet the accelerated cleanup schedule.

Field screening methods produce real-time data that can be used in the field to guide sampling. For example, real-time data on soil gas concentrations can be used to guide selection of the next sampling location. We are investigating how rapid display of the map of soil gas concentrations can aid decisions regarding sampling point density. The definition of the soil gas

plume is critical for guiding remediation decisions, such as where to place new extraction wells.

HEIS is being used to produce soil gas plume maps for use by the ERA through the Volatile Organic Compounds - Arid Integrated Demonstration (VOC-Arid ID) program, which is funded by the DOE Office of Technology Development to provide improvements to the current methods of conducting cleanup of volatile organic compounds such as carbon tetrachloride. HEIS generates diagrams of well site geology, displays maps showing plumes of contamination in ground-water, and depicts calcium carbonate percentages and other geologic information on maps.

We are investigating how a field computer system can be used for well site geologic data documentation and entry directly into HEIS. This system would streamline the availability of the data and would reduce the opportunities for errors during database entry from paper records. We are also evaluating how HEIS sample numbers can be provided electronically to field personnel using the same field computer system.

CONCLUSIONS

The environmental restoration mission at Hanford is projected to require 30 years to complete. Over the years, HEIS hardware and software will continue to evolve as computer technology advances and additional user needs are identified. We estimate that we will create from 10 to 20 million records consisting of 50 to 150 gigabytes (one gigabyte equals 1,000 megabytes). HEIS has been designed to support this huge data management task by integrating diverse software tools and data structures into a computer system that helps to meet the challenges of Hanford's environmental mission.

Our primary focus continues to be managing the data effectively. This has been a learning process for everyone involved. In addition, we are beginning to focus on new ways to use the data we have gathered. For example, HEIS allows us to take data in a variety of forms and utilize those data in an integrated way. This paper has presented some of our early attempts to use the HEIS as an analysis tool and guide for cleanup decisions. We expect the HEIS database and GIS to be indispensable partners in the process.

REFERENCES

1. T. WILLIAMS and C. KELLEY, "GNUPLOT, Version 3.2," Free Software Foundation, Inc.(1992).
2. "Resource Conservation and Recovery Act of 1976," 42 USC 6901 et seq.
3. "Comprehensive Environmental Response, Compensation and Liability Act of 1980," 42 USC 9601 et seq.
4. ECOLOGY, EPA and DOE, "Hanford Federal Facility Agreement and Consent Order," 2 vols., Washington State Department of Ecology, U.S. Environmental Protection Agency, and U.S. Department of Energy (1990).
5. WHC, "Hanford Environmental Information System (HEIS) User's Manual," Vol. 1, Rev. 1, WHC-EP-0372, Westinghouse Hanford Company (1991).

TRADEMARKS

1. Sequent is a trademark of Sequent Computer Systems, Inc.
2. Sun Workstation is a trademark of Sun Microsystems, Inc.
3. ORACLE is a trademark of Oracle Corporation.
4. ARC/INFO is a trademark of Environmental Systems Research Institute, Inc.
5. Uniface is a trademark of Uniface B.V., Amsterdam, The Netherlands.
6. SQR Workbench is a trademark of SQ Software.
7. VAX is a trademark of Digital Equipment Corporation.
8. Apple is a trademark of Apple Corporation, Inc.