

**WASTE MANAGEMENT DATABASE SEARCHES:
FINDING THINGS AMID THE 70,000-ODD DOCUMENTS AT A DOE FACILITY FOR SPECIAL PURPOSE USE**

Scott Mernitz, Ph.D.
Senior Project Manager
Woodward-Clyde Consultants
Denver, Colorado

Gislinde Engelmann
Environmental Engineer
EG & G Rocky Flats Inc.
Golden, Colorado

ABSTRACT

A new and special approach was needed to integrate RCRA, CERCLA, and varying types of radioactive waste data - according to recent DOE guidance - with National Environmental Policy Act (NEPA) data needs at a DOE facility. Specifically, a scheme was needed to sift through the 70,000-odd records at the Rocky Flats Plant to locate the most relevant 500-5,000 documents for support of NEPA work.

Records at the Plant include countless personnel, health and safety, weapons fabrication, sampling and analysis, and other hard copy files and computer databases; for solid wastes (sanitary landfill), hazardous wastes, mixed wastes, and radioactive wastes. In addition, numerous environmental reports, files, and sample analysis results exist for soil, air, surface water, groundwater, sediment, and other environmental topics, studied here since the early 1950s. Up to 160 operable Units (OUs)—including 174 Individual Hazardous Substance Sites (IHSSs)—were required to be individually and cumulatively addressed by discipline in the hard copy summary report, dubbed the ERTSD (Environmental Restoration Technical Support Document).

First, the other databases that had been assembled for previous special purposes were examined. These were queried with key words and other indicators as appropriate. Next, a system of locating information through the **key informant** and **strings of documents** approach was successful in accessing the 300-odd documents that were finally determined to be needed for the NEPA summary document. Finally, Level I, II, and III documents--based on report utility--were earmarked for later retrieval for this report and other purposes.

The CLARIS-published Macintosh File Maker Pro software was required to be used for the company computer database to be compatible with previous efforts at the Rocky Flats Plant. The report and database were successfully completed and are in use. A training session was held for EG&G personnel to further acquaint them with the software and expedite the use of later customized file searches. The purpose, significance, and conclusions center around this effort as a model for other database searches, and tips to researchers faced with a similar difficult task.

INTRODUCTION

An increasing trend at DOE facilities is the need to integrate CERCLA, NEPA, and RCRA in the study of various environmental media and the cleanup of past contamination. Agency policy to integrate at least two of the three regulatory programs has been in place since 1990 (1). This far-reaching guidance, acknowledging the strong and continuing pressure from the American public to address environmental impacts in the face of issues of national defense and economic well-being, has changed the technical emphasis and budgeting process throughout the Federal government. These three regulatory programs, in point of fact, provide the primary definition for available waste data at Rocky Flats (Fig. 1).

As many of these facilities proceed with their missions, they must also now address waste minimization; alternate process technologies; increased awareness of short- and long-term worker health and local population risks; and the generalized, interdisciplinary environmental impacts of past and present actions. Some facilities--such as the Rocky Flats Plant near Denver--are changing their mission from nuclear weapons production and research to decommissioning and cleanup. These facilities must also, as a secondary mission,

organize their databases for these new purposes. It is within this context that the assignment was conducted.

THE PROBLEM

At Rocky Flats, a new and special approach was needed to sift through the estimated 70,000-odd records to locate the most relevant 500-5,000 documents for NEPA purposes. The research team was initially faced with a monumental task of how to find, organize, pore through, summarize mentally and in writing, and evaluate the utility of each document for NEPA purposes. Duplications were expected (by some) to be a nagging problem, as well. The resulting report, compiled with the methodology described below, is named the Environmental Restoration Technical Support Document (ERTSD).

The ERTSD (2), as a NEPA data update, has several potential uses. Among these, two programmatic EISs are in preparation at Rocky Flats for Defense Programs and Environmental Management. An update to the 10-year-old Site-Wide EIS may also be prepared during the next few years. These are in addition to various special environmental baseline and impacts studies on soils, wildlife, and other topics at the Plant, and EAs for operations such as the new sanitary landfill.

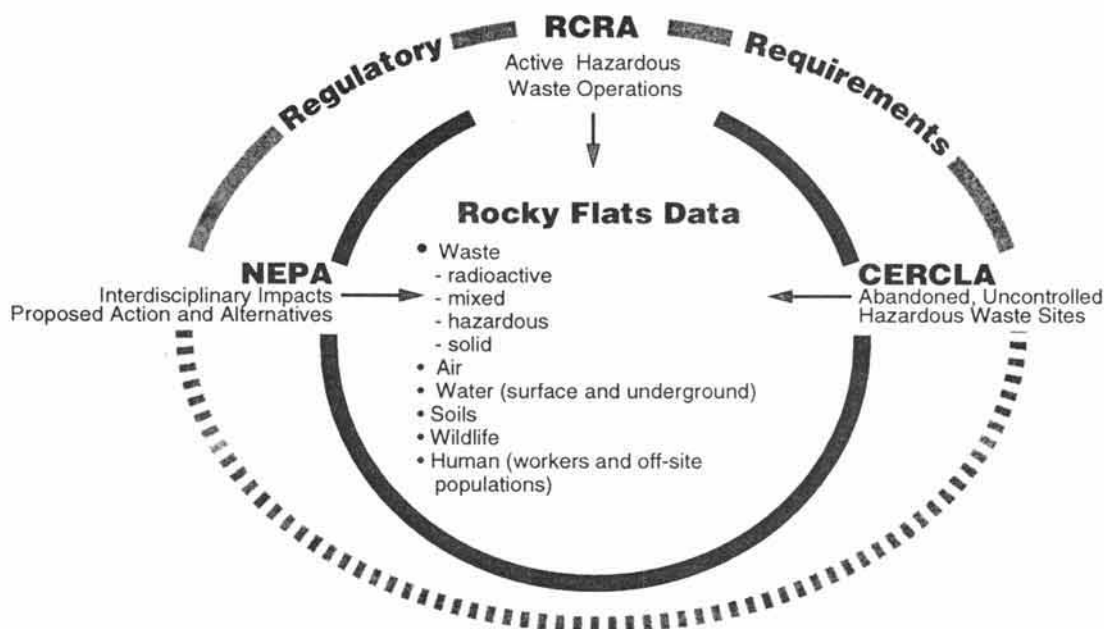


Fig. 1. The circumference of some of the data needs at the Rocky Flats Plant.

The purpose of this paper, therefore, is to suggest a model for database searches at DOE facilities, in cases where the data to be retrieved are a secondary or lesser issue at the facility.

EXISTING DATABASES—SOME PARTICULARS

The research team was faced with a multitude of documents, as previously noted. They were not all unorganized, however. The Rocky Flats staff had compiled databases for special purposes at several times in the past. None were especially for NEPA. The major ones considered and their key characteristics were as follows:

Historical Release Report

- Approximately 4,000 documents
- Documents dated 1950s through 1991
- Generated as a direct requirement of the Inter-Agency Agreement among EPA, DOE, the Colorado Department of Health, and other agencies
- Collected to support a report to completely document historic releases of hazardous substances

Integrated Research File

- Approximately 20,000 documents
- Documents dated 1950s to 1980s
- Mostly technical reports; with some older environmental reports (1960s-1970s)
- Includes some annual Environmental Monitoring Reports

Environmental Master File and Clean Water Subfile

- Most important for NEPA purposes
- Numbers approximately 35,000 documents
- Documents date back to the 1950s

- A cataloging effort was underway, but only about 3,500 of the documents had been organized at the time of this study
- Contains mostly information on water, soil, air, early data collection efforts (1950s to 1970s), spills, and weather
- Present search capability exists for approximately 16,000 documents on the LAN (Local Area Network) system at Rocky Flats

Church-McKay

- One of the first computerized databases at Rocky Flats
- Addressed environmental information from 1970 through 1980
- Consists of about 5,000 documents
- Compiled for special purpose litigation support

Personal DOE/EG&G Staff Files in the Various Environmental Divisions

Research team needed to:

- Conduct interviews to locate reports
- Institute a check-out system
- Provide for guaranteed return

The last data source noted above (personal files) was perhaps the smallest, most challenging, and yet most fruitful. The conditions of technical and professional courtesy were important, and needed to be met.

Other databases were recognized as the study began, including the ChemRisk, SAS, and Weston databases. Some were searched in great detail, like ChemRisk, while others, such as SAS and Weston, were not investigated in detail due to their nature and extent.

One final important source of information was used extensively by the NEPA researchers. This is the Rocky Flats

Reading Room at the Front Range Community College, in north Denver. Since these documents were cleared through security for public review, and since many contained information of interest for various environmental disciplines and members of the public who had special concerns about such disciplines, this proved to be a valuable source of many dozens of relevant documents.

As one might imagine from reviewing the above database information, many details regarding the searches, including techniques which work well, and pitfalls, cannot be presented in this paper. However, it was apparent that the success of the research effort resulted from careful selection and coaching of the appropriate disciplinary professionals to staff the research effort. An attentive sponsor, to smooth the path, expedite appointments, and gain priority and cooperation, is also essential.

THE SOLUTION

As the study began, it thus appeared that a unique solution was needed to access the desired information. The document to be produced had to contain a review of all relevant NEPA documents for all of the environmental disciplines, including soils, ground and surface water, air quality, and perhaps a dozen others. It was finally planned that the ERTSD would present, in an organized and easily used format, a summary of NEPA-type information currently available at RFP for each Operable Unit (OU). This information would include data on OU and Individual Hazardous Substance Site (IHSS) history, environmental setting and characteristics, contaminants present, contaminant levels, dispersion and movement, remediation plans, risk assessment, and other information relevant to an environmental impact analysis of contamination and remediation at the OUs. As suggested above, the ERTSD would be compiled from previous studies at RFP, including published and unpublished reports, data summaries and analysis, plans, databases, memoranda, letters, file information, and other sources. Its objectives would be:

- To provide the NEPA authors and researchers at RFP with a factual basis for discussion and evaluation of each OU and the environmental restoration program (with the goal of providing complete NEPA documentation for those OUs for which sufficient information is available)
- To provide an adequate basis for tiering of future NEPA documentation for the remaining OUs (see CEQ NEPA regulations at 40 CFR 1502.20, (3))

The ERTSD would necessarily include a simple assessment of data utility; analysis of the available data would be left

to the EIS preparers. The technical work scope to produce the ERTSD would have the following key elements (Fig. 2):

1. Use a key personnel and strings of documents (Figs. 3 and 4) research process to identify and locate 5,000 documents - from the 70,000-odd records at RFP and neighboring libraries - which may be relevant for NEPA purposes, handle each of these documents and record basic data (author, title, year) for computer database coding on Part I of a literature evaluation checklist (Fig. 5).
2. Review briefly the 5,000 documents to identify those 1,000 which appear potentially relevant for NEPA purposes and further annotate their contents (Part II, Fig. 5) in a brief 2-3 sentence entry for computer coding.
3. Review further the 1,000 documents to identify the approximately 250 which appear most relevant, and abstract (1-2 paragraphs) their contents for computer coding (Part III, Fig. 5).
4. Mark all 250 abstracted documents for later retrieval.
5. Retrieve the 250 abstracted documents and use computer database sorting to provide relevant subsets to each of 12-15 NEPA discipline specialists (e.g., soils, wildlife, and air quality), so that each specialist may summarize relevant documents for NEPA baseline data purposes.
6. Provide the discipline summaries in the hard-copy report, and catalog all the documents in a computer database for EG&G Rocky Flats staff, to facilitate future NEPA efforts.

Note also on Fig. 2 that Quality Control (QC) checks, to eliminate duplication, check spelling and content, and assure a quality database were conducted two or more times on each entry by an experienced Rocky Flats scientist or engineer.

As the effort was conducted, we did in fact note that the search yielded documents with three levels of utility for the ERTSD effort (Fig. 6), as follows:

- Level III - These documents were determined to have definite utility for the ERTSD effort and represent the most current, comprehensive, and useful documents located during the ERTSD research process.
- Level II - These documents were determined to have possible utility for characterizing current environmental baseline conditions, the nature of contamination, and the status of ongoing studies at the plant site.

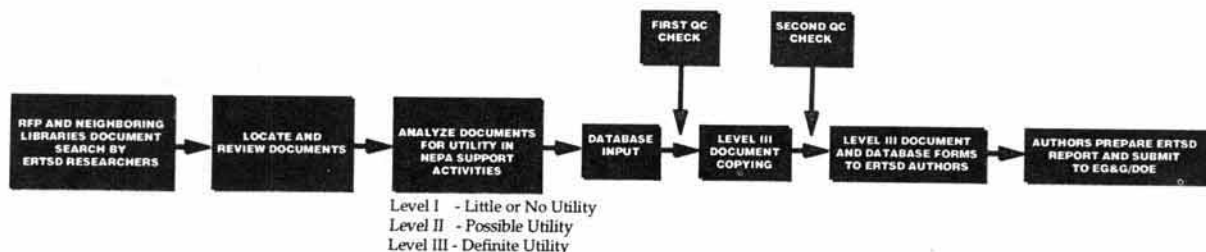


Fig. 2. Generalized ERTSD research process.

- DOE
- E G & G
- Environment and waste management
 - Environmental restoration management
 - Engineering and technology
 - Legal
- WCI
- WCFS
 - WCC
- OTHER CONTRACTORS
- Doty (historical release)
 - Others
- HISTORICAL
- Previous environmental employees

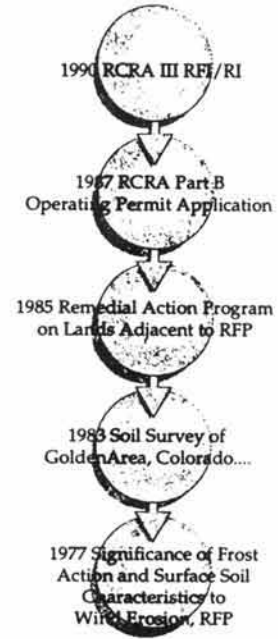


Fig. 3. Key informants at the Rocky Flats Plant.

Fig. 4. An example of a string of documents.

Reviewer _____
Date _____

PART I

AUTHOR(S) _____

PUBLICATION DATE _____ NO. OF PAGES _____ APPENDICES? _____ TABLES? _____ FIGURES? _____ YEARS OF DATA/ACTIVITY COVERED _____

TITLE _____

SOURCE/TYPE OF DOCUMENT _____

PHYSICAL LOCATION (AT RFP OR ELSEWHERE) _____

AVAILABLE AT _____ LOCATION _____

(/) APPLICABLE TO ERTSD _____

IF NOT, WHY NOT? _____

PART II

CIRCLE ONE(S) TO WHICH APPLICABLE ALL 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 NA

CHECK SUBJECT(S) TO WHICH APPLICABLE

_____ CULTURAL RESOURCES	_____ ENERGY	_____ VISUAL RESOURCES
_____ SOILS	_____ VEGETATION	_____ LAND USE, RECREATION
_____ TERRESTRIAL WILDLIFE	_____ SURFACE WATER	_____ AQUATIC WILDLIFE
_____ GEOLOGY/TOPOGRAPHY	_____ AIR	_____ GROUNDWATER
_____ HUMAN RISK	_____ ECOLOGICAL RISK	_____ OTHER (_____)
_____ SITE HISTORY/ NATURE OF CONTAMINATION	_____ TRANSPORTATION	_____ REMEDIATION
		_____ NOISE

KEY WORDS: _____

BRIEFLY ANNOTATE THE DOCUMENT (2-3 SENTENCES ON ITS CONTENTS) _____

_____ LEVEL OF TREATMENT (1=LIMITED, 5=THOROUGH) _____

SECURITY CLASSIFICATION _____

HOW TAGGED/NOTED FOR RETRIEVAL _____

IS THERE A REFERENCE SECTION? _____ HAS IT BEEN CHECKED? _____ ATTACHED? _____

PART III ATTACH AN ABSTRACT (IF AVAILABLE) OR SUMMARIZE CONTENTS OF DOCUMENT IN 1-2 SHORT PARAGRAPHS. _____

Fig. 5. Rocky Flats environmental restoration technical support document literature evaluation checklist.

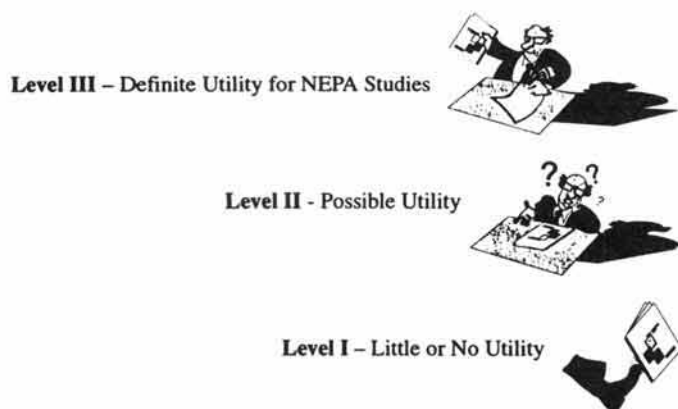


Fig. 6. Document utility.

- Level I - These documents appeared initially promising, but were later determined to have limited or no utility to the ERTSD process.

SIGNIFICANCE OF THIS PROJECT

This project is significant in that the researcher at DOE facility, often faced with a sizeable waste management database, may become confused and overwhelmed when also faced with a special purpose assignment. The "research army" approach--the old tale of 100 monkeys with 100 typewriters in a room for 100 days will produce a classic novel - may be considered. However, often the heavy manpower can be focused by a more conceptual and intuitive model: the key informants and strings of documents approach. This approach was successful in producing the necessary deliverables.

It follows that: 1) a hard copy reference document, with summary maps and tables, and 2) a computerized database specifically developed for the NEPA project are perhaps, in combination, most useful to future NEPA researchers. The reference document is in a loose-leaf, 3-ring binder format, easily updated. Similarly, the computer database can also be updated as new and useful information becomes available. This user-friendly database, on the CLARIS-published Macintosh File Maker Pro software, can sort by author, title, date, keywords, and in several other ways. Easy access and search capabilities for a number of users are provided. The update features of both products help to ensure that both the reference document and the database will not soon become outdated.

SUMMARY AND CONCLUSIONS

The approach and results suggest several tips for future database researchers, confronted with a large volume of information.

1. Estimate the large numbers of documents that may exist, to the nearest thousand, if possible.

2. Estimate the number finally desired for use, in a "manageable size" sense.
3. Get over the initial, overwhelmed feelings quickly--focus on the specific needs of the project.
4. Interview informed persons, those with both current and historical information about the facility.
5. Learn from the documents you already know about; their origin, type, and predecessors.
6. Deduce a way to find similar, useful documents using a focused model, to make the research effort most efficient.
7. Terminate the research, at least temporarily, when the estimated 80% goal is attained.
8. Assess your position, and perhaps be content with an "80% solution" if it meets your needs. Future researchers can find the rest, and new documents will be developed anyway.
9. Record the results in a hard copy reference document and computer database file, both easily used and updated.
10. Prepare many copies of the document, so that all facility users, at least to the division or section level, have a reference copy. In this way, maximum utility and updating, both at this installation and elsewhere, will be facilitated.

We suggest that this approach worked famously. Key NEPA documents in 20 disciplines -- both site-wide, and by OU and IHSS -- were identified, summarized, and remain as key references, with their content already explored, for future NEPA researchers. Unusual NEPA topics such as human and environmental risks, and remediation, were also addressed. This effort was that useful example of scientific inquiry where the research team initially struggled, was overwhelmed, hit upon a solution, and gained the full support of the sponsors for an approach that produced deliverables of acknowledged long-term use at a DOE facility.

REFERENCES

1. Council on Environmental Quality, "Regulations on Implementing National Environmental Policy Act Procedures," 40 CFR 1500-1508, effective November 29, 1978 as amended (1986).
2. EG&G Rocky Flats, "Environmental Restoration Technical Support Document (ERTSD): A NEPA Support Document for the Rocky Flats Plant." Prepared by Woodward-Clyde Consultants. Golden, Colorado (1992).
3. Oak Ridge National Laboratory, "Integrating NEPA and CERCLA Requirements During Responses at DOE Facilities," Env. Sciences Div. Publication No. 3505, Martin Marietta Energy Systems, Inc., for the U.S. Dept. of Energy, ORNL/TM-11564 (1990).