THE AGONY AND THE ECSTASY; COMPUTERIZED LITERATURE SEARCHING IN HISTORICAL ARCHAEOLOGY

Linda DeLowry-Fryman

University of Pittsburgh
Pittsburgh, Pennsylvania

and

Maureen Lambert

Johns Hopkins University
Baltimore, Maryland

ABSTRACT

With computerized literature searching, a wide range of data bases is available to the historical archaeologist, and the number of potentially valuable bases is steadily increasing. Some data bases provide bibliographic references, some statistical data, and still others record specialized information such as trademarks, patents, or funding sources. The historical archaeologist will find the information in these data bases useful not only for artifact analysis and identification, but also for historical background research. There are data bases available that cover topics as diverse as historic preservation, bone analysis, demographics, and ceramic trademark identification.

This paper outlines the process of searching, the steps involved in developing an advantageous search strategy, and the selection of appropriate files. The research potential of computerized literature searching can only continue to expand as more and more data bases become available. As soon as one tried to describe the current state of the art, the state changes; but a general understanding of the organization and value of computerized literature searching for the historical archaeologist will provide him/her with the information necessary to minimize the agony and maximize the ecstasy of data base searching.

INTRODUCTION

In recent years, a new tool called computerized literature searching has been added to the arsenal of research techniques that historical archaeologists can utilize in accessing pertinent information in their field. Although data base searching has been available in other fields, especially in the sciences, for more than a decade, only within the past few years have those bases of value to historical archaeology emerged with enough breadth of coverage and chronological depth to make their use as a bibliographic tool viable and worthwhile. In an effort to familiarize the reader with this emerging resource, the authors will describe the process of computerized literature searching and then share the findings of some sample searches that were run as "tests". We hope that the information provided here will help
researchers minimize the agony and maximize the ecstasy of computerized literature searching in historical archaeology.

As Stephen K. Stoan writes in Computer Searching: A Primer for the Uninformed Scholar, "The application of automation to the world of libraries has resulted in considerable improvement in the amount and type of information that can be made quickly available to the library user" (Stoan 1982:10). Like almost all other disciplines, the field of historical archaeology has benefited from the introduction of new technology to library services. Researchers currently have available to them not only the traditional paper indexes and abstracts, which have long been the chief research tools of scholars, but they also now have at their disposal many of the same indexes on line, plus an ever increasing number of new data bases. Historical archaeologists, because of the interdisciplinary nature of their interests, have always ranged widely in the traditional indexes of other fields. Not only do they find pertinent information in the anthropology indexes, but they also use history indexes, indexes to government publications, and even some humanities tools in their research. Since historical archaeology, by its nature, is interdisciplinary, scholars in the field must be familiar with a wide range of research tools that might aid them in their work; consequently, historical archaeologists, on the whole, are flexible and creative seekers of information and relatively knowledgeable about traditional library research techniques. Computerized data base searching is an "...increasingly valuable supplement to the traditional literature search...[thus] understanding what computerized literature searches are, what they can do for you, and what they cannot do for you is important for today's researcher" (Richardson 1984:169).

COMPUTERIZED LITERATURE SEARCHING—THE PROCESS

In library or information science parlance, a data base is merely a collection of discrete records accessed through a computer. The three types of data bases most frequently encountered in library research are statistical, full text, and bibliographical. Statistical bases are self-explanatory in terms of their content—they contain statistics or numbers. Full text bases, the least frequently encountered type, include not only the bibliographic means to identify an item but also the complete text of the article. The third type, the bibliographical data base, is the one we will be highlighting in this presentation. The data bases we will discuss range beyond those directly or narrowly applicable to Ohio Valley urban or historic archaeology, for much of the information gleaned through computerized literature searching would be primarily useful for comparative purposes.

Bibliographic data bases may contain many types of materials, including citations to books, periodicals, reports, newspaper articles, and government publications. They may also contain a wide range of levels, from popular materials to the highly technical or scholarly. Although there have been some strides toward making end-user searching available to scholars who wish to conduct their own computerized literature searches, the most frequent and probably, overall, the most productive type of search is still the interactive search in which a trained data base searcher/librarian (sometimes called search analyst) acts as in the intermediary between the computer and the researcher. As Charles Gilreath points out in Computerized Literature Searching: Research Strategies and Databases, "Ideally the relationship between the search analyst and the end user of information is a synergistic
one, resulting in the best retrieval possible" (Gilreath 1984:19). In any case, the analyst-assisted search is the type most frequently encountered in today's college and university libraries.

Most libraries have a designated area and staff that deal with computerized literature searching. A patron must generally make an appointment with a data base searcher during which the searcher as intermediary works individually with the researcher to understand his information needs, to identify key concepts for the search, to develop a search strategy, and perhaps to identify some pertinent data bases even before going on line to do the actual search. The clearer the researcher can make his information needs to the search analyst, the better his chances of receiving superior search results, i.e., bibliographic citations highly relevant to his research problems/needs.

Although choosing the appropriate data bases is primarily the search analyst's task, the researcher should realize that there are several ways he can help the data base searcher in formulating an effective search strategy. Gilreath writes, "The eclectic nature of social science research produces for the student and scholar a morass of terminology, much of it consisting of common words and phrases to which has been applied special meaning" (1984:53). In historical archaeology, as in the social sciences in general, the fluidity of terminology, coupled with the likelihood of synonymous terms, can best be tackled by the patron working in conjunction with the data base searcher. The researcher can help in defining the search terms by working up a preliminary list of key words, plus their synonyms, before the appointment with the searcher. The researcher, as the subject specialist, is in the best position, by virtue of his training and experience, to determine appropriate terminology and to suggest "contingency planning", i.e., alternative combinations of key words or concepts (Gilreath 1984:20). Even the most expert data base searcher will not be able to produce useful results if the researcher cannot articulate a "well-formulated search question". The computer maxim "garbage in, garbage out" is certainly applicable to on line searching. If the researcher has not clarified the search topic, an on line search is a waste of time and resources (Richardson 1984:171).

When the search analyst, using the search strategy determined by the interview process, finds that the citations being retrieved are relevant, he will provide them to the patron as quickly as possible. Rules concerning the printing of the bibliographic citations vary from library to library, but usually only a limited number can be printed immediately, while the bulk will be printed off line and mailed to the library for patron pick-up.

If the researcher, after evaluating a number of sample hits, determines that the citations are not relevant to his research problem, there are several alternatives that the data base searcher may try. First, the searcher will review the key words to be certain that they truly describe what the patron is seeking. If there has been a breakdown in communication, the search strategy will be reformulated based on the new information. The searcher may also opt to try reverse methodology. With this technique, he will determine the key words or descriptors of one relevant hit and then search on those descriptors in the hope of finding more bibliographic entries that are relevant to the patron's topic. Whatever the process used, the hoped for end is a bibliography of useful and relevant citations.
Once started, the searcher will explore all data bases he determines appropriate to the research topic. To search effectively in historical archaeology, a search analyst must keep abreast of all new bases coming online, even in areas not directly or immediately applicable. In the test searches described below, there were some bases that turned out to be surprisingly useful when, upon first consideration, they did not appear directly relevant. Textile Technology Digest is one example of a potentially useful base that at first glance seemed of dubious value. Clearly, in the rapidly evolving world of computerized literature searching, the search analyst with the most experience, skill, and up-to-date knowledge of available bases will be the most successful intermediary.

Finally, before leaving the general topic of the search process, one other concept of computerized literature searching and its consequences to the researcher in historical archaeology should be mentioned. That concept is the idea of the "false drop" and the difference between a "clean" and a "dirty" search. The ability of the computer to search not only on subject headings but also on title words and words in abstracts is clearly a major improvement over the hierarchical subject heading approach of most paper indexes. As Stoan indicates in his excellent introductory article on data base searching cited earlier in this paper, "...this advantage is sometimes not utilized because of the attitudes of many online searchers" (Stoan 1982:12). These searchers, in an attempt to have as few false drops (irrelevant citations) as possible, will opt to run a"... 'clean' search, limiting the search to descriptors and perhaps to title words", rather than a "'dirty' search, technically called a free-text search, in which the searcher tells the computer to run through the abstracts as well" (ibid.:13).

Since data base searching in historical archaeology involves choosing search terms with multiple meanings and manipulating an often imprecise vocabulary (a problem common to most of the humanities and social sciences), the researcher may wish to specify he prefers a dirty to a clean search, one "...aimed at high recall rather than high relevance" (Stoan 1982:14). This is especially important when the search analyst is working with unusual or non-mainstream data bases which may contain applicable material but which require that the analyst exhibit some creativity or adventurousness in pulling out relevant material.

SAMPLE SEARCHES—METHODOLOGY

We thought that some test searches in a group of data bases would be useful in demonstrating the utility of computerized literature searching for historic archaeology research. Our first step to that end was to brainstorm, and our brainstorming session resulted in a long list of data bases. Initially we included even those of doubtful interest, intending to use the list as a jumping off place for search topics. In compiling this list, we used a catalog of data bases from a vendor called DIALOG which included about 300 different files. We also actually ran a search on one of the DIALOG files known as the "Database of Databases", looking for descriptions of files which contained information on historic archaeology or historic preservation.

Consultation with an historic archaeologist was the next step. We talked with him about topics and asked which data bases he thought might be useful. We also designed the search strategies, which we kept relatively simple in
order to maximize the size of the sets. The topics ranged from very broad to very specific. We then revised our original list and chose several sample files that we wanted to explore. Our samples included searches for specific types of pottery from specific areas, general articles on historic archaeology, historic preservation, research in progress in the area, conservation of textiles, and one on a specific dyeing technique.

SAMPLE SEARCHES—PROCESS AND RESULTS

The first search we performed was on the file Artbibliographies Modern. This file, produced by ABA-Clio, "...contains references to all modern art and design literature in books, dissertations, exhibition catalogs, and some 300 periodicals...covering the fields of art history, biographies of artists, and artistic media such as sculpture, ceramics, printing, etc. since the onset of the 19th century" (DIALOG Information Services, Inc. 1985:11).

Our initial attempt on the file was to locate articles on very specific styles of pottery, such as Lotus Ware, hound-handled pitchers, and a design called "Rebecca at the Well" used on ceramic pitchers. Entering these terms resulted in zero hits. We then tried another specialized term, "Rockingham", which netted four hits, two of which were interesting: "Early Bennington Potteries", in the journal Ceramic Monthly, and "Traditional Wares from Farmhouse and Whelk Stall," from the Antique Dealer and Collector's Guide, a British publication. The latter article, while not of primary use as a source of information for American pottery, would be helpful for comparative research, since it describes the British ware with its distinctive brown glaze, widely imitated by early American potters.

Our final strategy on this file was to look for articles on 19th century pottery, or on pottery that were from the areas of Ohio or Pennsylvania. This attempt resulted in 23 hits, about half of which were of interest. Some sample titles include: "Earthenware Potters along the Great Road in Virginia and Tennessee", "Rookwood: A Cincinnati Art Pottery", and "The Stoneware of Greensboro-New Geneva". The source journals in which the relevant articles appeared included American Art & Antiques, Pennsylvania Folklife, and American Craft.

Next we were eager to try a file called Textile Technology Digest, described in the DIALOG data base catalog as follows: "...international coverage of the literature of textiles and related subjects. Over 650 journals are scanned as well as books, theses, patents, standards, conferences, and directories...Subjects include dyeing, laundering, mill operation, man-made and natural fibers, preservation, home economics, apparel design, marketing, and statistics" (DIALOG Information Services, Inc. 1985:43).

Our original idea was to look for articles on the architecture and construction of historic textile mills. The search strategy employed was probably too broad, as we discovered in looking at the results, which included items like "Present Situation of Japanese Joint Ventures" and "Economic Process Engineering Problems in the Development of Industrial Fibers". We subsequently decided to search for articles on indigo dye, spurred on by memories of a paper on indigo dyeing at the previous year's Symposium on Ohio Valley Urban and Historic Archaeology (Renau 1985). Again, we used a very
broad strategy and netted quite a number of hits, only a fraction of which were applicable to historical archaeology. Examples of relevant items included: "The Indigo Phenomenon: Its History and Application", "Identification of Dyes in Old Oriental Textiles", and "Indigo Dyeing, an Ancient Art". The sources for these items included The Textile Chemist, and Art and Archaeology Technical Abstracts.

The discovery of even a few relevant hits encouraged us to run a very broad search on the terms "preservation" and "conservation". While this resulted in several hundred hits, many of which were false drops, we were able to eliminate those having to do with energy by entering it as a term, then "notting it out" of the set. The resulting set was very helpful, including items such as "Preserving Flags and Banners", a 23 page pamphlet from the Smithsonian, and "Care of Fabrics in the Museum", as well as "Future from the Past", which discussed the conservation of old textile mills.

On the whole, we were very pleased with the results on this file, especially since our initial impression was that it was highly technical and industry-oriented. In addition, we were introduced to a new source we did not previously know, Art and Archaeology Technical Abstracts.

We next included a very general search on some files to discover what was currently happening in the field by selecting two current awareness files, National Newspaper Index and Public Affairs Information Service (PAIS). The former covers several national newspapers (e.g., Wall Street Journal, New York Times, etc.); the latter covers over 800 journals and 6,000 non-serial publications, such as Congressional hearings and government documents. For information on research in progress, we included two files called SSIE Current Research (from the Smithsonian Science Information Exchange) and Federal Research in Progress. SSIE claimed to include research in the social sciences, while Federal Research in Progress seemed from the catalog description to include only the physical sciences. We were therefore more hopeful of success with SSIE.

Articles that were retrieved from the National Newspaper Index and PAIS ranged from popular items such as "Preservation Fights to Preserve Itself", to the more substantial "Historic Preservation Planning in New Jersey". The latter appeared to have an extensive bibliography. Interestingly, the PAIS search also uncovered a membership directory published by the Early American Industries Association, evidence of the wide range of materials to be found in this file.

On the research in progress files, we again employed a very broad strategy, using the phrase "historic archaeology or historic preservation". Federal Research in Progress as expected, did not yield a great number of hits. We did uncover one item called "Historic Preservation Policy: Implications for Urban and Rural Housing". SSIE Current Research was more fruitful, yielding a total of 24 items, including "Neighborhood Discovery", a project by the Baltimore City Commission for Historic and Architectural Preservation, and "Bibliography and Index to Historic Site Archaeology", sponsored by NEH and compiled by J. L. Cotter.

Our historical archaeologist expert pointed out that these last two files would be useful for professional contacts, perhaps in locating reviewers for
Finally, two other data bases have proven to be most useful in the study of historical archaeology: National Technical Information Service (NTIS) and America: History and Life. NTIS "...consists of government-sponsored research...plus analyses prepared by federal agencies, their contractors or grantees" (DIALOG informational cover sheets 1982). This base was searched last year as part of the research for a bibliography on Upper Ohio Valley historic archaeology (DeLowry-Fryman 1985). Of the eight files tested at that time, this one and America: History and Life proved to be the most valuable. A search of NTIS identified a variety of historic archaeology and cultural resource survey reports that had been conducted under contract to the U.S. Army Corps of Engineers. The one drawback with this file is that it includes such reports only if they are government sponsored (contracted) projects. Still, it is a valuable file to check periodically if this kind of current awareness information is useful in your work.

America: History and Life covers the "...full range of U.S. and Canadian history....American studies, ethnic studies, folklore, history...local history, oral history, prehistory,...popular culture" (DIALOG informational cover sheets 1982). Because of the holistic, interdisciplinary character of historical archaeology, this history file seemed a potentially valuable one when the bibliography on the Upper Ohio Valley historic archaeology was being contemplated more than a year ago. A subsequent search of the file bore out the initial impression of applicability, yielding a variety of periodical articles relevant to the topic. It is certainly a data base of importance to any researcher working in historic archaeology.

CONCLUSION

Throughout this paper we have tried to suggest some new and interesting approaches to conducting bibliographic research in historic archaeology. There is clearly a place for computerized literature searching in the subject area. The multi-term search capacity, the enviable speed, and the ability to move rapidly from one type or level of material to another—all capabilities of the computer searching process—are definite advantages of data base searching as a library research tool. Additionally, computerized searching enables one to search an entire abstract as well as descriptors, a considerable advantage over the traditional hierarchical subject heading approach to the paper indexes.

Despite these pluses, one must remember that computerized literature searching is not the only technique the historic archaeologist must keep in his research arsenal. In addition to being more expensive than conducting one's own traditional search in the paper indexes, using only a data base search will inevitably miss some pertinent material because not all relevant journals, books, or documents are online. Because of the relative newness of the searching process itself, there are also chronological limitations to the materials indexed in the online bases.

Finally, poor communication between the search analyst and the researcher or the inexperience of a new searcher can result in a poor search, a problem with which the scholar does not need to deal if he does his own literature
search in the paper indexes. So the computerized literature search cannot be all things to all people; but, as we have tried to show in this paper, it definitely serves a valuable function for the researcher in historic archaeology when used in conjunction with other library research techniques.

ENDNOTES

1. Although some information science visionaries see the end-user search as an ultimate goal, the many problems with its widespread implementation will probably not be overcome in the foreseeable future. System complexity, variability of data base protocols, and search costs are all factors that make end-user searching difficult and costly to implement.

2. The descriptions for NTIS and America: History and Life are taken from the informational cover sheets that accompany each DIALOG search. The cover sheets, which contain an explanation of the file and a sample record which helps the user interpret the results of his search, are copyrighted by DIALOG Information Services, Inc. (1982).

REFERENCES CITED

DeLowry-Fryman, Linda

DIALOG Information Services, Inc.
1985 DIALOG Database Catalog. Palo Alto, California.

Gilreath, Charles L.

Renau, Lynne S.

Richardson, Larry L.

Stoan, Stephen K.