

Final Report--Objective B, Task 1

December 1986

A RESOURCE LIBRARY FOR PARAPSYCHOLOGICAL LITERATURE AND RELATED DISCIPLINES

By: G. SCOTT HUBBARD
SRI International

RHEA A. WHITE
Parapsychology Sources of Information Center

Prepared for:

PETER J. McNELIS, DSW
CONTRACTING OFFICER'S TECHNICAL REPRESENTATIVE

333 Ravenswood Avenue
Menlo Park, California 94025 U.S.A.
(415) 326-6200
Cable: SRI INTL MPK
TWX: 910-373-2046



SRI International

*Final Report--Objective B, Task 1
Covering the Period 1 October 1985 to 30 September 1986*

December 1986

A RESOURCE LIBRARY FOR PARAPSYCHOLOGICAL LITERATURE AND RELATED DISCIPLINES

By: G. SCOTT HUBBARD
SRI International

RHEA A. WHITE
Parapsychology Sources of Information Center

Prepared for:

PETER J. McNELIS, DSW
CONTRACTING OFFICER'S TECHNICAL REPRESENTATIVE

SRI Project 1291

Approved by:

ROBERT S. LEONARD, *Executive Director*
Geoscience and Engineering Center



333 Ravenswood Avenue • Menlo Park, California 94025 • U.S.A.
(415) 326-6200 • Cable: SRI INTL MPK • TWX: 910-373-2046

ABSTRACT

The organization and maintenance of an extensive data-base facility for parapsychological literature is described. During the first year of a multi-year effort, the Parapsychology Sources of Information Center (PSIC) purchased the necessary hardware and software, and then implemented the first working version of the data base. In addition, the PSIC has entered approximately 16 years of abstracts for each of the major parapsychological journals—ending with the current issue. The overall goal is to eventually include abstracts of bibliographic information of the entire literature of parapsychology in the data base. Although the core of the data base will consist of English-language parapsychological journals, it will also include books, chapters, dissertations and theses, separate reports, and conference proceedings. In addition, the data-base will include abstracts (in English) of parapsychological literature not originally written in English, as well as articles on parapsychology in the journals of other disciplines (e.g., psychology, psychiatry, physics, etc.).

TABLE OF CONTENTS

ABSTRACT	ii
I INTRODUCTION	1
II BACKGROUND	2
III METHOD OF APPROACH	3
IV DISCUSSION	4
APPENDIX	5
Summary	6
Overview	7
Contents of PsiLine	7
Contents of Each Record	8
PsiLine Search Capabilities Using TEXTBANK	10
Data-Base Display Capabilities	11
Hard-Copy Records	12
Recommendations for FY 1987	13
Access Procedures	13
Fees	13
Development of Standardized Vocabulary	13

I INTRODUCTION

The Parapsychology Sources of Information Center (PSIC) has been asked to provide, and maintain, an extensive data-base facility for parapsychological literature.* Although the program will be evaluated and the contract renewed on a yearly basis, it is anticipated that the data base will be enlarged and maintained through Phase III of the prime contract, on SRI International's major research effort.

This report details the progress of this effort for FY 1986.

*This report constitutes Objective B, Task 1, detailing the establishment of a resource library of parapsychology literature.

II BACKGROUND

Since the founding of the Society for Psychical Research (London, 1882), there has been extensive worldwide publication of journal articles, books, abstracts, popular articles, and conference proceedings about parapsychology that have been written in many different languages. For over 20 years, Ms. Rhea A. White, Director of the PSIC, has been following this literature from a technical and organizational perspective. Her hard-copy records, prepared primarily by hand, contain a nearly complete cross-referenced and abstracted accounting of this literature.

Other areas of science, such as those represented by the American Physical Society or the American Psychological Association, have massive amounts of literature that have been abstracted and categorized, and are available through a number of commercial library retrieval organizations. While some of these commercial services contain some material on parapsychological literature, it is often either out-of-date and/or incomplete. Given the estimated near-time demand, it was inappropriate to obtain the services of one of the commercial organizations and require them to add the parapsychological literature to their data bases.

Because Ms. White's expertise is in library science, and because she is an active researcher and is completely familiar with the parapsychological literature, she was the obvious and appropriate choice for the performance of this task.

III METHOD OF APPROACH

As detailed in the Interim Report dated 31 March 1986,* the PSIC has already completed the necessary hardware and software choices to implement the first working version of the data base. In order to carry out the second task for FY 1986, the PSIC began the actual work of writing abstracts and entering them into the data base.

A detailed report of the progress to date is contained in the Appendix. A listing of the volumes entered from the principal journals is given below:

- *European Journal of Parapsychology (1975-date)*
- *Journal of the American Society for Psychical Research (1970-1985)*
- *Journal of Parapsychology (1964-date)*
- *Journal of the Society for Psychical Research (1970-1985)*
- *Proceedings of the American Society for Psychical Research (1970-1974)*
- *Proceedings of the Society for Psychical Research (1970-1985).*

The PSIC is continuing the task of writing abstracts, checking references, and entering data.

*"Report on Design of the PsiLine Database System," R. A. White, PSIC, New York.

IV DISCUSSION

The Parapsychology Sources of Information Center has already satisfied the initial requirement for a complete design specification for PsiLine--an on-line bibliographic retrieval system for the parapsychological literature. The enormous task of completely abstracting and coding all necessary parapsychological literature will obviously require sustained effort.

The scientific literature on parapsychology research was initiated in the Nineteenth Century. As a result, it is estimated that the PsiLine data base will contain approximately 50,000 documents when brought up to date; approximately 1,000 documents per year will be added to maintain currency.

As outlined in the final report from the PSIC for FY 1986 (see Appendix), two outstanding issues need to be resolved for FY 1987 in order to make PsiLine available to the general community. Those issues are (1) access procedures and (2) fee structure. It was obvious early in the contract year that PsiLine could not be made directly available to researchers. Manipulating a large computerized data base requires specialized knowledge and experience to achieve the desired results. Therefore, queries will be directed to the PSIC, where Ms. White will perform the actual search. The fee structure will most likely be based on the rates charged by commercial services such as Dialog. Once these issues are fully resolved, PsiLine will be made available to the parapsychology community.

Appendix

FINAL REPORT FOR 1985-1986 ON PSILINE DATA BASE

Rhea A. White

Parapsychology Sources of Information Center

2 Plane Tree Lane, Dix Hills, New York 11746

SUMMARY

During Fiscal Year 1986, a data-base program, TEXTBANK, was selected for PsiLine, the data base of the parapsychological literature. As of September 30, 1986, PsiLine consists of nearly 3.5 megabytes, composed of 2,000 records. Each record consists of an individual journal article, chapter, dissertation, letter, or other published record. PsiLine is mounted on an IBM compatible system operating at 7.22 MHz. TEXTBANK is a text retrieval program having variable record and field sizes whose capacity is limited only by the hardware or the operating system. It searches every word or number in the data base.

OVERVIEW

CONTENTS OF PSILINE

PsiLine currently covers documents published in the major parapsychological journals from 1970 to date, with some exceptions. The coverage of the *Journal of Parapsychology* begins with 1964; coverage of the *International Journal of Parapsychology* begins with Volume 1, 1959 and proceeds through the last volumes in 1968; and over 100 articles from non-parapsychological journals that were published prior to 1970 are also included.

The other principal parapsychological journals in the data base are the *Journal of the American Society for Psychical Research* (1970-1985), the *Journal of The Society for Psychical Research* (1970-1985), the *Proceedings of the American Society for Psychical Research* (1970-1974) [(latest)], and the *Proceedings of the Society for Psychical Research* (1970-1984) [(latest)].

The full runs of some of the newer parapsychological journals are included. In this group are *Anabiosis: The Journal of Near-Death Studies* (1982 to date), the *European Journal of Parapsychology* (1975 to date), and *Parapsychological Monographs* (all numbers published to date are included, as well as other monographs, including the reports of the Princeton Engineering Anomalies Research Laboratory).

The entire contents of these major parapsychological journals are included, except for book reviews, letters about book reviews, notices, and glossaries. In other words, all articles, research notes, papers, articles about books, and letters to the editor are included.

Also included are approximately 250 articles from non-parapsychological journals, 50 from general interest magazines, 50 dissertations, 80 abstracts of Parapsychology Foundation conference papers, 50 chapters, and 60 books.

CONTENTS OF EACH RECORD

A record consists of an individual journal article, book chapter, dissertation, or other published record. There are 16 fields (TEXTBANK calls them zones) in each record (TEXTBANK calls them documents). These fields are:

- **Record Number**--A unique number that enables a document to be called up quickly. It is also used to indicate cross-references from one abstract to another.
- **AU**--Author's name (including names of all joint authors; given as they appear in the article).
- **TI**--The full title of the document.
- **SO**--The source (generally the journal title, year and month of publication, volume and issue number, inclusive pagination, place of publication, and publisher).
- **AD**--Additional items of interest such as the number of figures, illustrations, tables, references, etc.
- **YR**--The year of publication (listed separately to enable sorting by year of publication only).
- **PS**--The publication source, or the organ or organization responsible for the publication.
- **AB**--The abstract or summary of the document.
- **LA**--The language of the original document.
- **DI**--The discipline characterizing the source of the document, e.g., general science, psychology, parapsychology, education.
- **DT**--The document type, e.g., journal article, book, dissertation, chapter.
- **DF**--The document form, such as an experimental report, letter, research note, book review article, editorial.
- **DE**--Descriptors or subject headings.
- **ID**--Identifiers, which in PsiLine means proper name headings, and the affiliations of the authors. In this field or zone, order is important.

First is the name of the country where the initial author wrote the document followed by his or her institutional affiliation. If a second author has a different country or affiliation, it is given next, and so forth. Last, proper name subject headings (such as names of psychics or mediums or organizations covered in the document) are identified.

- **RL**--Reading level (1 indicates no background is required to comprehend the document; 2 means some background is required; and, 3 means the document is technical in nature and aimed at professionals). Most of the documents in the parapsychological journals--certainly the experimental reports--receive a rating of 3.
- **DN**--Document number; this refers to the number of the document if it is abstracted in *Parapsychology Abstracts International* (PAI). It is included because a search consisting of document numbers only could be obtained.

If a heading in the Descriptor or Identifier field is followed by a space and an asterisk, this indicates that a major portion of the document is devoted to that subject. If a title is followed by three asterisks, it indicates that the document is significant for some reason: outstanding quality, uniqueness of coverage, possible notoriety, frequently cited by other authors, and so forth.

All of the fields in the data base have been filled in (when available) except for the descriptor and document number fields. The latter can only be filled in if the document is in *PAI*.

The descriptors are a separate problem. Because a standardized vocabulary and thesaurus of terms for PsiLine has not yet been developed, not all records have descriptors; assigning a descriptor for some records, for example, would require a rereading of a journal article and thus would be too time consuming at this stage in the development of PsiLine (see "Recommendations for FY 1987" at the end of this Appendix). However, even without a descriptor, a record can be accessed by any word in its title or abstract. Currently, four different kinds of display are available for the Descriptor field: (1) for many items there are no descriptors; (2) sometimes a few descriptors have been assigned that are quite major and obvious; (3) some of the records have been completed, with descriptors typed in lower case except for the first letter of the first word; and (4) using the program David Altman wrote, we were able to transfer some of the old dBASEII headings to TEXTBANK; these headings were typed all in caps.

PSILINE SEARCH CAPABILITIES USING TEXTBANK

TEXTBANK permits Boolean searches through use of the logical connectors between terms of OR, AND, NOT, or XOR. For an OR search, at least one of the two terms must occur. For an AND search, both terms must occur; for a NOT search, we look for the first term named but not the second; and for an XOR search, only one or the other of the terms may occur, but not both.

When doing complex searches by using logical connectors to join the answer sets of previous searches, TEXTBANK automatically stores the last five answer sets so that they can be searched in place of the former query. It is also possible to save earlier answer sets until the end of the search.

PsiLine is capable of searching the entire document but if we are interested in only one field (e.g., SRI International not as an affiliation of an author but as a subject), then we specify that only the abstract field be searched. To search for J. B. Rhine as an author but not as a subject, we specify that only the author field be searched.

Adjacency may also be specified. If we want terms to occur within a limited proximity of one another, we can specify within the entire document or within a specific field that a term occur within a specified number of characters or words or sentences or paragraphs of one another. (However, if sentences, or words or paragraphs are specified, and a specific field is selected to be searched, it must be one that contains paragraphs, sentences, or words.) If direct adjacency is required, we simply search the terms we want to find together, such as "random event generator." If we search a long phrase, we must take care that certain words reserved by TEXTBANK are not included, such as "and, is, in." In such cases, the words can be searched simply by entering the phrase enclosed in quotation marks.

TEXTBANK also allows for range searching, so that if we are interested in RNG experiments published after 1983, say, we would do an RNG or REG (random number generator or random event generator) search only in those documents where the year field is greater than 1982.

It is possible to save specific queries for future use. For example, if we wish to keep up to date on a specific subject, say, precognition experiments employing dreams as the response mode, we could do so by requesting that this query be run each time new data are loaded in the data base, or if preferred, at a regular interval, such as once a month.

DATA-BASE DISPLAY CAPABILITIES

It is possible to specify which fields we would like to see displayed on the screen or printed. As a general rule, three types of displays will be available: (1) author, title, and source; (2) author, title, source, and abstract; and (3) entire record. Before displaying or printing the answers, they can be sorted on up to six fields in ascending or descending order. Thus, they could be sorted chronologically, and within that by author, and within that by title.

HARD-COPY RECORDS

The parapsychological journals that have been abstracted are stored on bookshelves at the PSIC. Articles from journals not owned by the PSIC, separate reports, monographs, etc., are stored in file cabinets by the record number used in PsiLine. Thus, it is easy to locate the original of any record if a full copy should be desired.

The unit we use for loading purposes is usually a full volume of a given journal. The printout at various stages of editing, as well as the handwritten abstracts and any relevant notes concerning that volume, are kept in a manila folder--one per volume--marked with the load number for the data base. Thus, if anything goes wrong or something needs to be checked, the input data are always easy to find.

We also maintain a listing of what was loaded into the data base each time, including the numbers of the records, the number of characters, the total number of records, and a notation identifying the floppy disks. We keep the information on two floppy disks per load: one is on the disk of the inputter, and the other is on another set of disks arranged by journal. For example, the file for the *Journal of Parapsychology* for 1971 is located on the I.G. No. 7 disk (the 7th disk containing files keyboarded by Inga Goldhammer) and the *Journal of Parapsychology* disk No. 4.

RECOMMENDATIONS FOR FY 1987*

ACCESS PROCEDURES

It is recommended that the PSIC perform all search and display for PsiLine. A client would write, telephone, or contact the PSIC via modem to communicate a search request. The PSIC would conduct the search and send the results in whatever display the client preferred.

FEES

During FY 1987, the fee structure and distribution of revenue (if any) will be determined. For example, the fee structure for PsiLine may be similar to commercial rates, such as Dialog, Psychinfo, Medline, and Dissertation Abstracts International. PSIC proposes that the charge for searching PsiLine would be \$30 per hour, or 50 cents per minute. This fee would be halved for members of the Parapsychological Association (PA). After the PSIC upgrades to a faster microprocessor, the charge would be higher because more work could be done per minute. For printouts, the charge would be 20 cents per item for the No. 1 display choice (bibliographic citation plus abstract) and 60 cents for the No. 3 display (full record). These same rates would apply for members of the PA.

DEVELOPMENT OF STANDARDIZED VOCABULARY

An essential feature of PsiLine will be a standardized vocabulary and thesaurus of terms. It is recommended that development of such a vocabulary be delayed until after the bulk of the bibliographic information and abstracts for the parapsychological literature have been entered.

*Note: Above recommendations are from Subcontractor (PSCI) to SRI International.