BNWL-1478
UC-32

PACIFIC NORTHWEST LABORATORY
INFORMATION RETRIEVAL SYSTEM

R. L. Reynolds and B. M. Cole

June 1970

AEC RESEARCH &
DEVELOPMENT REPORT

BNWL-1478

BATTELLE

BATTELLE MEMORIAL INSTITUTE

NORTHWEST

PACIFIC NORTHWEST LABORATORIES

3000 STEVENS DRIVE, P.O. BOX 999, RICHLAND, WASHINGTON 99352

LEGAL NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Atomic Energy Commission, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or 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.

PACIFIC NORTHWEST LABORATORY

RICHLAND, WASHINGTON

operated by

BATTELLE MEMORIAL INSTITUTE

for the

UNITED STATES ATOMIC ENERGY COMMISSION UNDER CONTRACT AT(45-1)-1830

3 3679 00061 7227

BNWL-1478

UC-32, Mathematics & Computers

PACIFIC NORTHWEST LABORATORY
INFORMATION RETRIEVAL SYSTEM

By

R. L. Reynolds

B. M. Cole

June 1970

BATTELLE MEMORIAL INSTITUTE
PACIFIC NORTHWEST LABORATORIES
RICHLAND, WASHINGTON 99352

Printed in the United States of America
Available from
Clearinghouse for Federal Scientific and Technical Information
National Bureau of Standards, U.S. Department of Commerce
Springfield, Virginia 22151
Price: Printed Copy \$3.00; Microfiche \$0.65

PACIFIC NORTHWEST LABORATORY
INFORMATION RETRIEVAL SYSTEM

R. L. Reynolds and B. M. Cole

ABSTRACT

The Pacific Northwest Laboratory Information Retrieval System is a general purpose computer system for the storage, search, and retrieval of almost any type of data. The system will store and search both numeric and narrative data of fixed or variable length, and there are no limits on the amount of information which can be stored except those imposed by hardware.

The file structure of the system is described, and brief instructions for its use are included. Two current applications are discussed with examples of the computer output included.

TABLE OF CONTENTS

INTRODUCTION	1
APPLICATIONS	2
Steam Electric Plant Files.	2
Personnel Capabilities File	2
DATA STORAGE AND RETRIEVAL	6
USE OF THE SYSTEM.	13
STARTUP INSTRUCTIONS FOR DATA ENTRY	15
UPDATING AND CORRECTING DATA	19
ACKNOWLEDGEMENTS	24
REFERENCES	24
DISTRIBUTION	25

LIST OF FIGURES

1. Query Output - English Language Version	4
2. Query Input - Checklist Version	7
3. Query Output - Checklist Version	8
4. Startup Output - Checklist Version	18
5. Logic Flow Diagram	20

PACIFIC NORTHWEST LABORATORY
INFORMATION RETRIEVAL SYSTEM

R. L. Reynolds and B. M. Cole

INTRODUCTION

Any analysis which is based on large amounts of statistical data can be very laborious and oftentimes incomplete. The use of a computer to store and retrieve information is the most effective means to reduce the amount of hand calculations involved in evaluating data and to ensure more accurate and complete results from these evaluations.

The PNL Information Retrieval System is a general purpose computer system for the storage, search, and retrieval of almost any type of data. (1) It was designed to be usable without modification for most applications; however, the modular construction permits additions for any particular application with relative ease. The development was sponsored by the Atomic Energy Commission and private industry.

This system will store and search both numeric and narrative data of fixed or variable length. There are no limits on the amount of data which can be stored except those imposed by the hardware available at any particular computer installation.

Convenience for the user and flexibility in the formulation of inquiries were primary considerations in the design of the system. Efficiency of the search and retrieval operations is good, but was considered secondary to user convenience and generality in application.

The general system provides two forms of inquiry, a check-list form and an English language form. Two current applications of this system are described, each using an extension of the general system with modifications to meet specific requirements.

APPLICATIONS

This system has been used during the past year for several applications. Two of these, one using the English language version and the other the check-list version, are described here. Both are shown on the block diagram of logic flow (Pages 20-23).

Steam Electric Plant Files

Data for 928 major steam electric generating plants within the continental United States was gathered from published statistical sources. (2,3,4, 5,6, & 7) A list of the type of information recorded and loaded into the information system files can be found in Table I. The power plant files have been used to generate fossil fuel cost distributions, within each Federal Power Supply area, for the conceptual reactor design and evaluation studies being conducted by Battelle-Northwest for the U. S. Atomic Energy Commission.

There are many analyses and evaluations which can be derived from these data. Typically such analyses include: capital costs, production costs, fuel costs as a function of plant size, year of startup, geographic location, the amount of heat rejected to any given river system, generation as a function of plant size, or age. Samples of the type and form of the information retrievable from these files is shown in Figure 1.

Personnel Capabilities File

The BNW Personnel Capabilities File is a rapid and thorough means of indicating whether past experience or capabilities in specified fields is among the qualifications of members of the professional staff.

Questionnaires filled out by staff members described areas in which they felt they were competent. Selected data were added to the system from existing computerized personnel files.

Inquiries to retrieve the names of personnel meeting specified requirements are prepared on a check-list form.

TABLE I

File No.	File Description
1	Region, FPC
2	Area, FPC
4	Year of Data
5	Plant Type
6	Utility Type
7	Plant Name
8	Utility Name
9	Capacity, MWe
10	Net Generation, 1967
11	Year of Startup
14	Percent of Coal Consumed, 1967
15	Percent of Oil Consumed, 1967
16	Percent of Gas Consumed, 1967
18	Coal Consumed, Thousands of tons, 1967
19	Coal Heat Value, Btu's per pound, 1967
20	Oil Consumed, Thousands of barrels, 1967
21	Oil Heat Value, Btu's per gallon, 1967
22	Gas Consumed, Millions of cubic feet, 1967
23	Gas Heat Value, Btu's per cubic foot, 1967
24	Plant Cost, Thousands of dollars
25	Non-Fuel Production Cost, Thousands of dollars, 1967
26	Fuel Production Cost, Thousands of dollars, 1967
28	Coal Cost, Cents per million Btu as burned, 1967
29	Oil Cost, Cents per million Btu as burned, 1967
30	Gas Cost, Cents per million Btu as burned, 1967
31	Fuel Cost, Cents per million Btu as burned, 1967
32	Condenser Cooling Supply
33	State
34	City or Town
35	Longitude, degrees.minutes west
36	Latitude, degrees.minutes north
39	Number of employees

```
*****
IF STATE IS WASH THEN LIST PLANT NAME $
*****
PLANT NAME: SHUFFLETON
.
PLANT NAME: GEORGETOWN
.
PLANT NAME: LAKE UNION
.
PLANT NAME: STEAM PLANT NO1
.
PLANT NAME: STEAM PLANT NO2
.
PLANT NAME: HANFORD
```

.
.
.
.

6 DETAILS SATISFIED THE PRECEDING QUERY

```
*****
IF STATE IS WASH AND LONGITUDE GT 122. AND LATITUDE GE 47. AND
*****
CAPACITY LT 50, THEN LIST TOTAL CAPACITY $
*****
```

.
.

```
          CALCULATED SUM
CAPACITY* MWE: 6.0000+01
```

.
.
.
.

3 DETAILS SATISFIED THE PRECEDING QUERY

FIGURE 1. Query Output - English Language Version

```

*****
IF STATE IN WASH AND AREA IS 44 THEN LIST PLANT NAME AND CITY AND
*****

*****
STATE AND CAPACITY $
*****
PLANT NAME: LINCOLN STREET
CITY OR TOWN: PORTLAND
STATE: ORE.
CAPACITY, MWE: 35.5000
.
.
PLANT NAME: STATION L
CITY OR TOWN: PORTLAND
STATE: ORE.
CAPACITY, MWE: 75.5000
.
.
.
.
.
.
.
2 DETAILS SATISFIED THE PRECEDING QUERY

*****
IF LONGITUDE GT 85. AND LONGITUDE LT 90. AND LATITUDE GE 38. AND LATITUDE LT
*****

*****
42. AND CAPACITY GT 100. AND CAPACITY LE 200. AND YEAR OF STARTUP GE 1960
*****

*****
THEN LIST AVERAGE PLANT COST $
*****
.

CALCULATED AVERAGE
PLANT COST, THOUSANDS OF DOLLARS: 3.234404
.
.

```

FIGURE 1. (Contd.) Query Output - English Language Version

The sample check-list form, which is shown in Figure 2, is filled out to search for employees who mentioned specified words or phrases in their capabilities statements. The results of this search on a small system of fictitious entries is shown in Figure 3, as is the response to a request to print all available data about a particular fictitious person. This small demonstration system is very similar to the operating BNW system, which includes more than 2,700 employees.

DATA STORAGE AND RETRIEVAL

Information entered into the system is stored in files which are defined and named during the initial data entry startup. Each of these files contains data of a particular type, such as annual salaries, powerplant capacities, geographical locations, or résumés.

Data is stored in the system for unique items such as individuals, manufacturing plants, cities, or any other group of clearly defined items for which large amounts of data are known.

Each unique item will have data referring to it stored in some or all of the files which are set up, so that conceptually the data forms a figurative two-dimensional matrix of files (types of data) against unique entities (persons or things).

Data in any of the files can be searched to find which entries meet any specified requirements, and the results of searching many files can be combined by logical "and/or" relationships to select those entities which meet all of the requirements that have been imposed.

The information which is in specified files for those entities which have met the imposed requirements will be printed out. The files which contain data to be reported may include those which were searched, but are not required to.

At least one of the specified files is searched for each unique entity. After the search, however, random access pointers are used to retrieve the requested information.

BATTELLE-NORTHWEST RICHLAND WASHINGTON	INQUIRY FOR THE BNW PERSONNEL CAPABILITIES FILE		
CHECK IF YOU WANT YOUR QUESTIONS PRINTED ☒			
YOUR NAME <u>MARKETING</u>	BUILDING _____		
DELIVER TO: CAPABILITIES FILE, MATHEMATICS BUILDING, 3000 AREA.			
FIRST SHEET ONLY			
CHECK BLACK BOX IF CONTINUED ON BACK			
I REQUEST INFORMATION ON EMPLOYEES WITH THE FOLLOWING NAMES OR PAYROLL NUMBERS			
A (_____) (_____) (_____) (_____) (_____) (_____) (_____) (_____) (_____) (_____) (_____) (_____) (_____) (_____) (_____)			
OR INFORMATION ON THOSE PERSONS WITH THE FOLLOWING : CAPABILITIES IN			
(<u>ARMY</u>) B(<u>COMPUTER</u>) B(_____) OR (<u>NAVY</u>) OR (_____) OR (<u>AIR FORCE</u>) AND OR (_____) OR (_____) OR (_____) AND OR (_____)			
C A COLLEGE DEGREE IN			
D (<u>CHEMICAL ENGINEERING</u>) AND D(_____) OR (<u>CIVIL ENGINEERING</u>) OR (_____) OR (_____) OR (_____)			
E WHO HAS A ☐ BS, ☐ MS, ☐ PHD, ☐ BS OR BETTER, ☒ MS OR BETTER, ☐ NO DEGREE			
F WHO ATTENDED _____ UNIVERSITY			
G WHO IS ☐ EXEMPT, OR ☐ NON-EXEMPT			
H WITH A SERVICE DATE ☐ BEFORE, ☐ IN OR AFTER THE YEAR (_____)			
J WHOSE BIRTHDATE IS ☐ BEFORE, ☐ IN OR AFTER THE YEAR (_____)			
K WHO HAS A JOB TITLE OR CODE OF (_____) OR (_____)			
L AN ORGANIZATION CODE OF (_____) OR (_____)			
M A REPORTING LEVEL NUMBER ☐ LESS THAN ☐ EQUAL TO ☐ GREATER THAN _____			
N WHO WORKS IN THE (_____) AREA WHO WORKS IN (_____) BUILDING WHO LIVES IN THE CITY OF _____			
S ☐ WHO HAS A CAPABILITIES ENTRY IN THE FILE			
I WANT THIS INFORMATION TO INCLUDE:			
☐ COMPLETE FILE FOR THESE PERSONS WORK ADDRESS AND TELEPHONE	☐ PERSONNEL RECORD ☐ SOCIAL SECURITY NUMBER	☒ CAPABILITIES ☐ PAYROLL NUMBER	☒ NAME
☐ HOME ADDRESS AND TELEPHONE	☐ ORGANIZATION CODE; EXEMPT OR NON-EXEMPT; REPORTING LEVEL; AND JOB TITLE	☐ CONTINUOUS SERVICE DATE; DUPONT DATE; GE OATE; BNW DATE	
☐ ORGANIZATION CODE	☐ DATE OF LAST CHANGE TO PERSONNEL RECORD	☒ COLLEGE AND DEGREE	
JUST GIVE THE TOTAL NUMBER OF PERSONS		☐ ALSO PRINT MAILING LABELS	

BB-1510-010 (2-70)

, FIGURE 2. Query Input - Checklist Version

29 JUN 70

D062970 5040470

RETURN THIS PRINTOUT TO

MARKETING

12 13 21 E5 B(ARMY)(NAVY)(AIR FORCE) B(COMPUTER) B(SIMULATION) *
D(CHEMICAL ENGINEERING)(CIVIL ENGINEERING)

THIS INQUIRY SEARCHED FOR THE FOLLOWING ITEMS

AND	DEGREE CODE	GREATER OR EQUAL	MS
OR	CAPABILITIES	EQUALS	ARMY
OR	CAPABILITIES	EQUALS	NAVY
AND	CAPABILITIES	EQUALS	AIRFORCE
AND	CAPABILITIES	EQUALS	COMPUTER
AND	CAPABILITIES	EQUALS	SIMULATION
AND	MAJOR CODE	EQUALS	CHEMICAL ENGINEERING
OR	MAJOR CODE	EQUALS	CIVIL ENGINEERING

COLLEGE CODE: 1188 CALIFORNIA* UNIVERSITY OF, BERKELEY
MAJOR CODE: 118 CIVIL ENGINEERING
YEAR OF 1ST DEGREE: 46
PG COLLEGE CODE: 1188 CALIFORNIA, UNIVERSITY OF, BERKELEY
PG MAJOR CODE: 174 SYSTEMS ENGINEERING
DEGREE CODE: 2 M.S.
FIRST NAME: HAROLD
LAST NAME: WILLIAMS
CAPABILITIES: 25760 HAROLD S WILLIAMS

06/29/70

* CIVIL ENGINEERING
S, 1946 UNIVERSITY OF CALIFORNIA

* SYSTEMS ENGINEERING
S, 1952 UNIVERSITY OF CALIFORNIA

* COMPUTER PROGRAMMING
TECHNICAL PROGRAMMING FOR LARGE SCALE COMPUTERS SINCE 1954. TAUGHT COMPUTER

FIGURE 3. Query Output - Checklist Version

SCIENCE COURSE AT UNIVERSITY OF CALIFORNIA IN 1954 THROUGH 1960.

* AVIATION

MR. WILLIAMS HAS A PILOTS LICENSE TO FLY BOTH SINGLE AND TWIN ENGINE AIRCRAFT

* REACTORS

HAS PARTICIPATED IN STUDIES OF REACTOR MODEL SIMULATION.

* MILITARY

HAS SERVED 3 YEARS AS AN OFFICER IN THE U.S. AIR FORCE.

COLLEGE CODE:
MAJOR CODE:
YEAR OF 1ST DEGREE:
PG COLLEGE CODE:
PG MAJOR CODE:
DEGREE CODE:
FIRST NAME:
LAST NAME:
CAPABILITIES:

6192 PENNSYLVANIA, UNIVERSITY OF, PHILADELPHIA
116 CHEMICAL ENGINEERING
44
4640 MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE
116 CHEMICAL ENGINEERING
2 M.S.
WILLIAM
~1~GENS
24360 WILLIAM S ROGERS

06/29/70

* CHEMICAL ENGINEERING

MS, M.I.T. 1949, BS, UNIVERSITY OF PENNSYLVANIA 1944. EXPERIENCE IN CHEMICAL PROCESSING, ESPECIALLY IN THE FIELD OF PLASTICS. PRESENTED TWO PAPERS AT AMERICAN INSTITUTE OF CHEMICAL ENGINEERS SYMPOSIUM. PUBLISHED 35 TOPICAL REPORTS AND 4 ARTICLES IN PROFESSIONAL JOURNALS. REGISTERED PROFESSIONAL CHEMICAL ENGINEER OF THE STATE OF WASHINGTON.

* COMPUTER SYSTEMS

TECHNICAL ASSISTANCE FOR LARGE SCALE COMPUTERS. MEMBER SIMULATION COUNCIL.

* MILITARY

HAS SERVED 3 YEARS IN THE U.S. NAVY.

2 ENTRIES SATISFY THIS QUERY

*** ALL QUERIES HAVE BEEN PROCESSED ***

FIGURE 3. (Contd.) Query Output - Checklist Version

BNWL

29 JUN 70

DO62970 5040470

RETURN THIS PRINTOUT TO

ADMINISTRATION

PAYROLL NUMBER: 53669
FIRST NAME: JAMES
FST MIDDLE INITIAL: R
SND MIDDLE INITIAL:
LAST NAME: JONES
SOCIAL SECURITY NUMBER: 559286320
BUILDING ADDRESS: MATH
AREA ADDRESS: BNW
PLANT PHONE: 61212
STREET ADDRESS: 39 TAYLOR
CITY: RICHLAND
STATE: WASH
HOME PHONE NUMBER: 9643011
PAY STAU: EXEMPT
ORGANIZATION CODE: 7R20
REPORTING LEVEL: 4
JOB CODE: 2030 SENIOR R+D MANAGER
SEX: MALE
MARITAL STATUS: MARRIED
BIRTH DATE: 07/27/36
BNW DATE: 05/02/67
GE STARTUP DATE: /
DUPONT DATE: /
CS STARTUP DATE: /
COLLEGE CODE: 1179 CALIFORNIA* UNIVERSITY OF* AT LOS ANGELES
MAJOR CODE: 390 OTHER NON-TECHNICAL
YEAR OF 1ST DEGREE: 58
PG COLLEGE CODE: 1179 CALIFORNIA* UNIVERSITY OF* AT LOS ANGELES
PG MAJOR CODE: 305 BUSINESS ADMINISTRATION
DEGREE CODE: 2 M.S.
DATE**: 06/29/70
CAPABILITIES: 53669 JAMES P JONES

06/29/70

* ECONOMIST
BA UCLA 1958. MEMBER OF AMERICAN ECONOMICS ASSN., GOOD BASIC CAPABILITY IN

-10-

BNWL-1478

FIGURE 3. (Contd.) Query Output - Checklist Version

PRODUCT DIVERSIFICATION IN SUCH FIELDS AS INDUSTRIAL CARBONS, PETROLEUM, FOOD
PROCESSING. 3 ARTICLES PUBLISHED IN AMERICAN ECONOMIC REVIEW.

* MANAGEMENT

BA, UCLA 1960, MAJOR IN MANAGEMENT

* BUSINESS

BA, UCLA 1960

* SMOG CONTROL

HAS BASIC UNDERSTANDING AND HAS PARTICIPATED IN RESEARCH DEALING WITH SMOG
CONTROL.

* RECREATION

HAS DONE EXTENSIVE RESEARCH IN THE FIELD OF RECREATION.

* DRAMA

HAS BEEN ACTIVE IN LITTLE THEATER GROUPS IN CALIFORNIA AND HAS APPEARED SEVERAL
TIMES IN RICHLAND PLAYERS PRODUCTIONS.

* MILITARY

HAS SERVED IN THE U.S. ARMY FOR TWO YEARS.

1 ENTRIES SATISFY THIS QUERY

FIGURE 3. (Contd.) Query Output - Checklist Version

Most files are designated as being one of three major types, depending on the type of data to be stored. The three major file types are called single vector (SV), special code (SC), and data file (DF).

Single vector (SV) files store data of a consistent and fixed length. Numeric data, which will always occupy one computer word for single-precision values, is stored in SV files and can be integer or real. Narrative (alphanumeric) data of fixed length, such as 80 characters, can also be stored conveniently in this type of file.

Searches of numeric SV files will retrieve entries which have a specified logical relationship (equal, not equal, less than, less or equal, greater than, greater or equal) to a value which is also specified in the question which initiates the search.

Narrative SV file searches can be based on a complete match, or mismatch, either of the entire entry and a sequence of characters specified by the inquiry, or of the initial part of the entry and the specified characters.

One of the files is specified as a control word (CW) type, and is a special version of the SV files. CW entries in the system must be unique (such as identification numbers, the names of states, isotopic symbols, or different letters of the alphabet) with no more than six characters, and are used to identify items of data which are entered into the system later. Another file can be designated as an identification (ID) type, if longer identifying titles are required.

Data files (DF) store narrative data of variable length, and neither consistency nor pre-determination of the length is required.

Searches of DF files will compare a string of characters (letters, words, phrases, or entire sentences) specified by the inquiry to the entire contents of each DF entry. If the specified characters occur in the same sequence anywhere in a DF entry, ignoring blanks and punctuation, a match will be indicated for that entity,

Special code (SC) files contain abbreviations or alphanumeric codes which stand for more lengthy narrative descriptions used by many control word entities. The descriptions are contained in common block (CB) files, and can be of any length up to 80 characters. A personnel file, for example, could have an SC file called "Job Title," with the job code "014" entered as data for a hundred individuals. The corresponding CB file would list all of the job titles, and would correlate the code "014" with the title "Vice-president for Marketing." A search to retrieve the names of these one-hundred employees could be made by comparing file entries with "Vice-president for Marketing," "Vice-President," or "014."

SV and SC files require computer memory space for each **entity**, whether data is entered or missing. DF and CB files require space only for data which has actually been entered.

USE OF THE SYSTEM

Information previously stored for one version-of the system **may** be retrieved with a basic English sentence entered by a **computer card** or teletype card image. Various logical and relational search operators **may** be used as constraints in a query to isolate and report only the exact data desired. Also, report operators are available to help manipulate **the** information asked for. The search and report sections of the query **are** based on the narrative file names loaded during the startup. The following table contains the search and report operators and the basic query forms.

TABLE II

<u>Relational Search Operators (so)</u>	<u>Logical Search Operators</u>
IS (equal to)	AND
NE (not equal to)	
GE (greater than or equal to)	
GT (greater than)	
LE (less than or equal to)	
LT (less than)	

TABLE II (Contd.)

Report Operators (ro)
(Data Manipulators)

Logical Report Operator

TOTAL (sum all data available)
AVERAGE (divide TOTAL by number of details
satisfied)
MINIMUM (select the smallest value)
MAXIMUM (select the largest value)
Blank (report all details)

AND

Query Form (Items in lower case are optional and are to be supplied by the
user. Upper case must appear at all times.)

1. IF file name so value THEN LIST ro file name \$
2. IF file name so value AND file name so value AND file name so value
THEN LIST file name AND file name AND file name AND file name \$——

When report operators other than blank are used only one file may be reported
per query.

At present, each English language query is limited to searching and
imposing constraints on up to ten files. Also, the number of files reported
in the same query is limited to ten. Any number of cards or card images may
be used to make a query, but, each query must be terminated by a space and
a "\$". Sample English queries appear in Figure 1.

Check-list version inquiries are submitted on a prepared form that anti-
cipates the types of questions which will be asked and facilitates rapid
entry of the inquiries as input data. Inquiries are entered by keypunching
the checked numbers and letters, with parenthetical entries, in any order
and with any desired spacing,

The form can be set up to search any of the files which are present,
with any desired combination of "and" and "or" logical connectors, and to
report any of the files singly or in combination.

Changes in the layout of the check-list form require corresponding
changes in tables in the subroutine CHECKOF, and recompilation of that sub-
routine.

Programs comprising this system were written in Fortran-V, except for three subprograms written in Sleuth-11, for use with the CSCX operating system or similar Univac systems. The system operates on a Univac-1108 computer with 53K available core, Fastrand drum permanent storage, and FH432 drum temporary storage.

Symbolic programs for the two versions now in use are available from BNW as two files (file 1 for the check-list version, file 2 for the English language version) on working tape U2994, with U9319 as a duplicate record center tape. The queries may be executed with the processed map BDB.

STARTUP INSTRUCTIONS FOR DATA ENTRY

The English language version presently allows the loading of up to 1000 unique items or control words. There can be as many as 500 separate pieces of information or files for each unique control word loaded. Example file numbers and names appear in Table 1.

Data may be entered into these files as numeric items in any standard FORTRAN mode (I, F, E) or as alphanumeric (A). Although only 500 files for each control word are allowed at this time, they need not be numbered continuously starting with 1 nor do they have to end with 500. This 1000 by 500 matrix limit was installed for a specific requirement but may be expanded for other applications.

A complete matrix or set of data need not be known. Any file for any control word may be left blank if there is no data available.

Because this version uses English language for queries the actual naming of the files is very important. Each file must have a unique narrative name. These names are not only used as titles but also as a means of calculating the correct search addresses.

The check-list version of the program is used for a specific application, with data from an existing computer file. Consequently, the entry of data is tailored to the form of this data, and its preparation and entry is mostly written into added subroutines.

The input required for data entry to the English language system must be on cards or on tape in card images. The first card necessary has the total number of control words or unique items to be loaded into the system.

Col. 1-6 (16) Number of Control Words

The next group of cards describes the files to be used for each control word. One card for each file.

Col. 12-13 (A2) Type of data in this file
 "SV" for numeric or fixed length alpha entries
 "DF" for variable length alpha entries
 "ID" for the table of file names

Col. 15-17 (13) File number (any three digit except 998 and 999)

Col. 58 (A1) Mode of entry for this file
 "I" for integer
 "F" for floating point
 "E" for exponential
 "A" for alpha

The next card signals the start for the reading of the data.

Col. 1-4 (A4) "READ"

The first set of data cards are for "SV" entries. One card for each file per control word for which data is known. (Card not required if **there** is no data available.)

Col. 56-58 (13) Control word number

Col. 61-63 (13) File number

Col. 71-80 (any mode) Numeric or fixed length alpha entry

The "SV" file entries are terminated by a single card.

Col. 1-3 (A3) "END"

The second set of data cards are for "DF" entries. One card for each control word per file for which data is known. Each set of "DF" entries starts with a single card describing the file.

Col. 6-9 (A4)	"DATA"
Col. 13-15 (13)	File Number

The form of the "DF" cards are:

Col. 1-54 (9A6)	Alpha entry
Col. 56-58 (13)	Control word number

(If the narrative data is longer than one card will accommodate, subsequent cards may be used for continuation with the same control word number.)

Each completed file is terminated with an "END" card.

Col. 1-3 (A-3)	"END"
----------------	-------

After all data has been entered, the table of file numbers and names must then be loaded. This table is called the IDENT file. First card of the table.

Col. 6-10 (A5)	"IDENT"
Col. 13-15 (13)	File number for this file

Followed by one card for each file.

Col. 3-6 (14)	File Number
Col. 7-60 (9A6)	File Name

The complete input package is then terminated with an "END" card.

Col. 1-3 (A3)	"END"
---------------	-------

Execution of the startup is accomplished with the startup main code TESTOT, A sample output of an accomplished startup, using the check-list version, appears in Figure 4.

FILE CODE	FILE NO.	RPT. NO.	MNEMONICS	FORMAT	DRUM ADD.	TOT. WDS.	WDS. IN FILE
SV	1	101	RECORD STATUS	000200010000	131073	7	1
CW	2	201	PAYROLL NUMRER	0002000050~0	131130	7	1
SV	3	301	FIRST NAME	000200160000	131187	21	3
SV	4	302	FST MIDDLE INITIAL	000200010000	131358	7	1
SV	5	303	SND MIDDLE INITIAL	000200010000	131415	7	1
SV	6	304	LAST NAME	000200220000	131472	21	3
SV	7	401	ORGANIZATION CODE	0002000400~0	131643	7	1
SV	8	501	SEX	000200060000	131700	7	1
SV	9	601	MARITAL STATUS	000200100000	131757	14	2
SV	10	701	SOCIAL SECURITY NUMBER	000200110000	131871	14	2
SV	11	901	GE STARTUP DATE	000200060000	131985	7	1
SV	12	1001	CS STARTUP DATE	000200060000	132042	7	1
SV	13	1101	DUPONT DATE	000200060000	132099	7	1
SV	14	1201	BIRTH DATE	000200060000	132156	7	1
SV	15	1301	BMM DATE	000200060000	132213	7	1
SC	16	1401	JOB CODE	000200040000	132270	7	1
SV	17	1501	PAY STAUS	000200120000	132327	14	2
SV	18	1701	REPORTING LEVEL	000200010000	132441	7	1
SV	19	1801	STREET ADDRESS	000200340000	132498	35	5
SV	20	1802	CITY	000200170000	132783	21	3
SV	21	1803	STATE	000200040000	132954	7	1
SV	22	1901	HOME PHONE NUMRER	000200070000	133011	14	2
SV	23	2001	AREA ADDRESS	000200030000	133125	7	1
SV	24	2002	BUILDING ADDRESS	0002000400~0	133182	7	1
SV	25	2003	PLANT PHONE	000200050~00	133239	7	1
SC	26	2101	COLLEGE CODE	000200040000	133296	7	1
SC	27	2102	MAJOR CODE	000200030000	133353	7	1
SV	28	2103	YEAR OF 1ST DECREE	000200020000	133410	7	1
SC	29	2104	PG COLLEGE CODE	000200040~00	133467	7	1
SC	30	2105	PG MAJOR CODE	000200030000	133524	7	1
SC	31	2106	DEGREE CODE	000200010000	133581	7	1
CB	32	2201	COURSE MNEMONICS	000200060000	135474	39	0
CB	33	2301	HIGHEST DEGREE MNEMONICS	0002000~0000	137549	32	0
DF	34	2401	RESUME	000200060000	135417	7	1
CB	35	2501	MNEMONICS FOR COLLEGES	000200060000	138581	161	0
DF	36	2601	CAPABILITIES	0002000600~0	134597	7	1
CB	37	2701	JOB TITLE MNEMONICS	000200060000	136513	36	0
SV	999	9999	DATE**	000200060000	139742	7	1

FILE \$T0059

TRACKS ASSIGNED = 20
 TRACKS USED = 8
 READ QUALIFIER = 1
 WRITE QUALIFIER = 1

FIGURE 4. Startup Output - Checklist Version

UPDATING AND CORRECTING DATA

The addition of new data or corrections of erroneous data previously loaded into the system is accomplished through the main code UPDATE.

Card 1 signals for an update or copy.

Col. 1-6 (A6)	"UPDATE" or "COPY"
Col. 13 (A1)	"D" or "T" From drum or tape (logical units, drum = A, tape = E)
Col. 14 (A1)	"D" or "T" To drum or tape (logical units, drum = A, tape = B)

Card 2 describes the type of data to be changed, the command word number, and the update terminating symbol.

Col. 1-2 (A2)	"DF" (for DF file) or "VF" for SV file)
Col. 3-6 (A6)	Control word number
Col. 9 (A1)	Termination symbol (\$,*,?,etc.)

Card 3 describes the file name.

Col. 1-2 (A2)	"DF" or "VF"
Col. 3-56 (9A6)	File Name

Card 4 contains the new information to be loaded.

Col. 1-80 (object time format)	New data
--------------------------------	----------

Card 5 terminates a single update.

Col. 1 (A1)	Termination symbol
-------------	--------------------

Cards 2 through 5 can be repeated during any execution for other updates. The last card in the series ends the update.

Col. 1-2 (A2)	"FN"
---------------	------

The program DATEUP adds and corrects data in the check-list version. This routine reads and translates correction cards from an existing computerized personnel file, and enters the new data into the system.

FIGURE 5

SEARCH PROGRAM

FIGURE 5 (cont'd)

INITIAL DATA ENTRY PROGRAM

FIGURE 5 (cont'd)

UPDATE PROGRAM

Figure 5 (Cont'd)

MULTIPLE - USE SUBROUTINES

Fortran-V Utility Routines

A.	BCD2I	(Converts field data to integers)
B.	CBDICT	(Provides CB file numbers)
C.	DICFND	(Provides file index numbers)
D.	DICT	(Provides file information from the dictionary)
E.	DQY1	(Converts data as DDMMYY to YYMMDD)
F.	ERID	(Writes error messages)
G.	ERR	(Writes error messages)
H.	ERRB	(Writes error messages)
I.	ERROR	(Writes error messages)
J.	ERRQRU	(Writes error messages)
K.	FLDIC	(Converts integers to field data)
L.	FMTFND	(Provides format index numbers)
M.	FMTSRT	(Checks order of entries)
N.	FNDOW	(Provides CW file index number)
O.	FNDDIC	(Provides file index number)
P.	MATCH	(Provides ID file numbers)
R.	MATCHU	(Provides ID file numbers)
S.	UNCQDE	(Provides output formats)

Sleuth-I1 Utility Routines

X.	FAKE	(Makes a subroutine look like a main program for walk-back)
Y.	FQURSL	(Compares arrays)
Z.	INRUN	(Provides user identification)

Sleuth-I1 CSCX System Routine Replacement

NTRAN\$

ACKNOWLEDGEMENTS

The PNL Information Retrieval System was designed and programmed by R. L. Engel, S. E. Grady, R. L. Reynolds, R. T. Toyooka, and E. L. Wells.

The data for the steam-electric power plant files was gathered and recorded by Pearl Horton.

REFERENCES

1. R. L. Reynolds, R. L. Engel, R. T. Toyooka, and E. L. Wells. Status of Data Retrieval System as of April 1, 1969, BNWL-1095. Pacific Northwest Laboratory, Richland, Washington, June 1969.
2. Federal Power Commission, Steam-Electric Plant Construction Cost and Annual Production Expenses, Twentieth Annual Supplement, FPC S-192.
3. Federal Power Commission. Statistics of Publicly Owned Electric Utilities in the United States, 1967, FPC S-194.
4. Federal Power Commission. Statistics of Privately Owned Electric Utilities in the United States, 1967, FPC S-190,
5. Rural Electrification Administration. 29th Annual Report of Energy Purchased by REA Borrowers, REA Bulletin 111-2.
6. Rural Electrification Administration. 1966 Annual Statistical Report, REA Bulletin 1-1.
7. National Coal Association, Steam-Electric Plant Factors, 1968 Edition, 18th Edition.

DISTRIBUTION

No. of
Copies

OFFSITE

1	<u>AEC Chicago Patent Group</u> G. H. Lee
2	<u>AEC Division of Technical Information Extension</u> UC-32, Mathematics & Computers
2	<u>AEC Division of Reactor Development and Technology</u> M. J. Whitman S. Strauch
2	<u>AEC Division of Isotopes Development</u> W. K. Eister J. E. Machurek
1	<u>AEC Division of Plans and Reports</u> M. Alexander

ONSITE

1	<u>AEC Chicago Patent Group</u> R. K. Sharp (Richland)
1	<u>AEC Assistant Director for RDT Programs</u> T. A. Nemzek
1	<u>AEC RDT Site Representative - PNL</u> P. G. Holsted
5	<u>AEC Richland Operations Office</u> C. L. Robinson M. R. Schneller Technical Information Files (3)
2	<u>Computer Sciences Corporation - Richland</u> S. E. Grady R. T. Toyooka

No. of
Copies

16 WADCO Corporation
B. W. Bentley
J. F. Fletcher (10)
D. R. Haffner
P. L. Hofmann
L. D. Jacobson
D. P. Shively
K. R. Wise

80 Battelle-Northwest
D. T. Aase
F. W. Albaugh
E. L. Alpen
E. R. Astley
W. L. Bathke
C. H. Bloomster
C. Bromley, Jr.
J. B. Burnham
G. J. Busselman
B. M. Cole (25)
F. G. Dawson
D. R. deHalas
D. E. Deonigi
G. E. Driver
R. L. Engel
S. J. Farmer
J. C. Fox
R. J. Hoch
R. T. Jaske
B. M. Johnson
R. C. Liikala
E. T. Merrill
R. W. McKee
M. A. McLaughlin
W. R. McSpadden
L. T. Pedersen
W. W. Porath
R. G. Rau
W. D. Richmond
W. A. Reardon
R. L. Reynolds (10)

W. Sale
L. C. Schmid
A. M. Schneider
J. M. Seehuus
D. H. Stewart
J. V. Thompson
R. E. Turley
E. E. Voiland
W. W. Waddel
Technical Information
Files (5)
Technical Publications (2)