J Forensic Sci, Apr. 1979, Vol. 24, No. 2

K. A. Oakes,[1] *Ph.D. and Christopher Budnick,*[2] *B.S.*

A Data-Independent Information Retrieval System: A Solution for the Small Laboratory's Computer Needs

At some point in the development and operation of a forensic science laboratory the question of computerization will be considered. Studies on this question done several years ago provided the basis for the development of the Crime Laboratory Information Service (CLIS) program [*1*]. The size and organizational structures of forensic science laboratories were found to vary greatly. Nevertheless, the study revealed a need or desire for computerization to provide analytical support, to provide statistics necessary to establish the uniqueness of a sample, to store and retrieve rifling specifications, to provide bibliographic information and references, and to provide a means of exchanging various types of information with other laboratories. The study recognized two general types of computer applications currently in use: support for instrumentation and management information systems.

The use of computers in either of these two applications will vary depending on the size of the laboratory system. Larger laboratories with heavy case loads may have interfaced many of their analytical instruments to a large computer system. Under these systems, the analyst need only prepare the samples and then press the start button. The computer will then set the parameters, inject the samples, analyze the data, and provide a computer-written report. The same large laboratory may also have a large management information system to handle administrative tasks. Such large, sophisticated computer systems are expensive to develop and to operate. In addition to the high equipment costs, such large systems require a full-time staff of computer operators and programmers to maintain and monitor the system.

Certainly the majority of the small forensic science laboratories can neither afford such a complex system nor justify the need for a large computer system to their supervisors or budget analysts. Nevertheless, as the Search Project studies [*1*] indicated, there are a need and a desire, even in small laboratories, for a computer to aid both the analysts and the administrators in their day-to-day operations. The analyst wishing to identify a chemical compound, for example, would like to compare his unknown spectrum to a standard file of reference spectra until he finds an exact match. Such an identification will be successful if the unknown spectrum is in the standard file and if the analyst can find the spectrum from a large collection of reference spectra. The administrator, on the other hand, wants information on the day-to-day operation of the laboratory. If a computer is available, information such as monthly statistics, budget resources, inventories, case output, and agencies served can be readily compiled.

Since many small laboratories cannot justify a computer system, they must examine

Received for publication 12 May 1978; revised manuscript received 22 Sept. 1978; accepted for publication 6 Oct. 1978.

[1]Head trace analyst, Wisconsin Regional Crime Laboratory, Wisconsin Department of Justice, New Berlin.

[2]Student, Carroll College Computer Science Department, Waukesha, Wis.

their needs and the options available to them more closely. To meet the analysts' needs various instrument manufacturers are now equipping their instruments with small microprocessors and using small computers on their more sophisticated instruments. These more sophisticated instruments are equipped with an appropriate data base that can be searched by the analyst. As an alternative to a data base on its instrument, the small laboratory may subscribe to various commercially available search systems. This approach allows the analyst, for a fee, to search thousands of reference spectra instead of the ones existing in the laboratory's reference collection. The CLIS program is currently being developed to offer this type of searching through the existing National Crime Information Center communications network. Depending on the organizational structure of the laboratory system, the administrator lacking his own computer may be forced to rely on using a larger department's system on a limited basis, or perhaps a remote terminal can be connected to a large time-sharing computer system via a telephone communications system.

None of these alternatives can simultaneously meet the needs of the analyst and the administrator in a small laboratory system. The alternatives, even if available at a reasonable cost, force the laboratory users to conform to an established system not necessarily designed to meet their individual needs. The ideal solution is, of course, to have one's own computer and have programs specifically written to accomplish the desired results. This solution is not very practical for small laboratories because of costs, justification, and the time required for development.

The Wisconsin Regional Crime Laboratory is a relatively small laboratory that faces many of these restrictions and yet has recognized the usefulness of having a computer system. While not being able to justify nor afford an independent "laboratory computer," the laboratory was able to justify and purchase an EDAX energy dispersive X-ray fluorescence spectrometer with a computer system. The EDAX system at the time of purchase was equipped with a Nova 210 computer, Data General's 6030 Floppy Diskette subsystem, and an ASR 33 teletypewriter. By requesting additional memory and a FORTRAN compiler at the time of purchase, the laboratory now has a computer system that can operate independently as well as with the X-ray fluorescence unit. To solve the problem of writing computer programs for both the analyst and the administrator, the laboratory examined the computer system being used at Carroll College, Waukesha, Wis., which is the National Laboratory for Higher Education (NLHE) Information Retrieval System (IRS), written primarily by T. Ray Nanney of Furman University, Greenville, S. C. and E. James Runde of Clark College, Dubuque, Iowa. This system was written for an IBM 1130 computer having 8K core memory, 1 disk drive, 1132 printer, and a 1442 card read/punch. Examination showed the IRS to be ideal for our laboratory's use, providing a data-independent search system that could be used by both the analyst and the administrator. The system, once implemented, can be operated by any of the scientific or administrative staff.

A Carroll College computer science student made the necessary modifications to the NLHE system in less than two months so that it would operate on our small computer system. The program was modified so that the ASR 33 teletypewriter became the input/output device, and the Floppy Diskette system replaced the disk system. These modifications restrict the full capabilities of the NLHE system but were necessary to conform to our existing computer system. By restricting our record size to 72 or 144 columns, a single diskette can accommodate approximately 3000 or 1500 such records. As many diskettes as required (cost, $6.00 apiece) can be used, although each search must be repeated on each new diskette.

The details of the NLHE information system and its logic are available elsewhere [*2,3*]. The system is an example of a set of data-independent programs designed to provide an integrated set of programs for generating, updating, and maintaining disk data files. The IRS sort package was designed to be used with any type of data, the only requirement

being that the data must be initially described by a file description program before the desired sorts can be made. The description program provides a description of the data fields that can be manipulated by the system and stores this information in a "description table." If new data fields are desired or if the original fields are to be redefined, only a new description table is necessary. A single file maintenance program will change, add, or delete records from any usable field.

In addition to the creation and maintenance of data files, the system is designed to allow sorting of the data and the selection of subsets of information. An added feature includes both statistical- and report-writing capabilities. The system, written in FORTRAN, has been designed to be used by persons unfamiliar with computer programming. Simplicity of use was accomplished by making the instructions for generation of lists, statistics, and reports self-explanatory, and the instructions were written with a free format. All instructions for linking the programs and making multiple passes through a given program are generated automatically by the system on the basis of the user-submitted instructions. The system can be used for a variety of applications without any program revisions or the need for a computer programmer or operator.

The use of the NLHE-IRS system at the Wisconsin Regional Crime Laboratory has been extensive and has shown the potential to meet almost all of the computer needs expressed by forensic science laboratories in the survey related to the CLIS program [*1*]. The applications of the system at the Wisconsin Regional Crime Laboratory can be divided into four major areas: (1) analytical, (2) administrative, (3) bookkeeping, and (4) research. A listing of the titles of some of the files created under each of these categories is given in Table 1.

The retrieval system allows the analyst to develop his own in-house data files. The Regional Laboratory has established a file of infrared spectral information for drugs based on our own reference collection. A similar file has been created for corresponding mass spectral data. When an unknown infrared spectrum is obtained or when an unknown mass spectrum is obtained, these files can be searched and the unknown identified if it is in our file. An additional file containing information such as spot test results, ultraviolet data, gas chromatographic data, literature references, or location or source of standards

TABLE 1—*Summary of NLHE information retrieval files being developed at the Wisconsin Regional Crime Laboratory.*

Administrative
1. Purchase orders and supply information
2. Regional laboratory case information for monthly statistics
3. Instrument service information
4. Chemistry-Physics Section information
Analytical
1. Infrared data base search
2. Mass spectral data base search
3. X-ray fluorescence data base search
4. Open bullet file
5. National Bureau of Standards paint information
6. Typewriter classification system
7. Worthless check file
Bookkeeping
1. Evidence location and return file
2. Standard gun inventory file
3. Drug file—type and quantity analyzed
4. Suspect, victim, agency officer, agency case number, laboratory case number files
Research
1. Polygraph case data
2. Gunshot residue information
3. Blood types

is being created to aid the drug analyst. The Chemistry-Physics Section is creating a file that will correlate X-ray fluorescence elemental composition data with the National Bureau of Standards paint collection and manufacturing information publication. The Firearms Section has also set up its own "open bullet" file designed to complement the CLIS program. A suspect weapon can be test-fired and the rifling specifications determined. The open bullet file is then searched to see if the suspect weapon could have been used in any previous unsolved shootings in Wisconsin. If the search is successful, the case in question is reexamined and the firearms evidence compared to the suspect gun. The Documents Section is in the process of creating a worthless check file, which can be used to detect worthless checks throughout the entire state and to associate the style or characteristics found on the checks with a suspected forger. The Documents Section has also found it advantageous to include their typewriter classification card file on the IRS system to aid in identifying the type of typewriter being used.

The administration has also found the IRS system to be quite useful. The monthly statistics relating to agencies served, type of evidence received, section output, and court time can all be extracted from a file containing the relevant case information. Such data may also be readily recalled for quarterly, fiscal, or annual reports. The purchase order file has been created to monitor the operating expenses of the laboratory by vendor, section, month, dollar amounts, and so forth. This information will provide valuable information for planning future budgets. The locations of equipment and supplies are maintained in an inventory file. Our laboratory has also found that a file monitoring the service calls and downtime of the various instruments will provide a service history for each instrument and will prove valuable in determining when an instrument needs replacing. Other types of administrative information are being examined to determine which information would be advantageously stored on diskettes.

Closely related to the administrative applications are the laboratory's bookkeeping tasks. The Regional Laboratory has established an evidence file, which monitors the location and flow of evidence in the laboratory. This file has been established to notify agencies when evidence is ready to be returned, including such information as the agency's case number, submitting officer's name, analyst's name, and the location of the evidence to be returned. A bookkeeping file is also built into the case information file so that the agency's case number, the officer's name, suspect's and victim's names, and the laboratory's case number are cross-referenced. Internally, each section head now has the opportunity to monitor the section's operation, including due dates, case priorities, types of evidence being analyzed, location of evidence, exhibits, examinations, and each analyst's output. Since the laboratory maintains a standard gun collection, an inventory file has been created to include the serial numbers of all guns, their location, who submitted them, date received, or any other relevant information such as destruction date. Statistical data relating to the types of drugs analyzed, their quantities, and the submitting agency can also be obtained from a drug file.

The final category where the IRS system has been implemented is data handling for research projects. Files have been set up in various sections of the laboratory to see if any trends or important information can be obtained from the data compiled over hundreds of cases. For example, the polygraph section is presently coding the data collected from its examinations and storing them on diskettes for future study with the statistical features of the IRS system. The laboratory is also monitoring the results from its gunshot residue examinations. Data are being compiled by type of gun, ammunition, occupation of person swabbed, elapsed time between incident and swabbing, and other classifications. These data can be analyzed by using the statistical aspects of the IRS system. Data collected from blood population studies, as well as actual case work, can be collected and analyzed with this system. Consideration is being given to other types of research projects involving the collection, storage, sorting, and searching of data for important trends and statistical

implications. These types of projects are geared to provide the statistics required to help prove the uniqueness of different types of evidentiary samples.

The IRS system, while meeting the needs of the laboratory in the areas of analytical support, administrative concerns, bookkeeping, and research projects, is a generalized program and does not need to be reprogrammed for use in each of these different areas. Each person (analyst, secretary, director, clerk, or researcher) using the system need spend at most 2 h initially to set up the data file's description table to suit his or her individual needs. Since the input device is essentially a typewriter, no computer skills need to be learned to perform this task. Once the description table is done, the data can be entered, updated, deleted, or modified on a day-by-day or week-by-week basis without any additional file changes. The actual time depends on the amount of data and the person's typing skills. The only problem encountered to date is scheduling each user on the system so as not to interfere with the computer's use with the X-ray fluorescence unit. This problem has been resolved by having the data collected on a coded form over a period of time and then transferred to the computer once a week rather than after each day's accumulation.

To demonstrate the actual usage of the IRS, two simple examples will be given. The first example represents the creation and usage of an administrative file entitled by the user "Chemistry-Physics Section Case Information for 1978." The user's first step is to create the description table for the file (Table 2).

The user defines 19 fields containing the information shown in the field description column. The number of columns allotted for the data in each field and the order of the fields are selected at this time. For example, six columns have been selected for the data entitled "Date of Offense." These columns will be the 16th through 21st columns of the data file. The user may now enter data into the "Chemistry-Physics Section Case Information for 1978" file directly or he may create a coded form to collect data over a time period. The coded form can be any standard computer programming form having at least 72 columns across the page. For this example, the 72 columns would be divided into 19 fields. Each field would consist of the number of columns specified, as shown in Table 2. Each new case will then consist of a new line on the form.

TABLE 2—*Description table for "Chemistry-Physics Section Case Information for 1978" file.*

Sort Field Number	Field Description	Field Abbreviations	Number of Columns per Field
1	chemistry/physics case number	CP	3
2	regional lab case number	RCL	8
3	submitting agency code	AG	3
4	county of submitting agency	C	1
5	date of offense	DATE O	6
6	date case received	DTRC	6
7	agency case number	AG. NO.	8
8	type of evidence code	TPEV	4
9	suspect's name	SUS	6
10	victim's name	VIC	6
11	type of case code	TPC	3
12	number of exhibits	EX	2
13	number of exams	EXM	3
14	date case due	DTDO	6
15	date case completed	DTCT	6
16	analyst	A	1
17	evidence disposition	EVD	3
18	court information	C	1
19	time spent on case	TM	2

Assuming the appropriate data have been entered into the file on a monthly basis, a statistical report can now be compiled. The user on 8 April 1978 might want a listing of all cases completed after 1 Jan. 1978 by each chemistry-physics analyst for the agency with the code number 100 as well as the total number of exhibits and examinations per analyst. The user would type the input control cards shown in Table 3. The computer then prints out the logic expression that will be used during the sort (Table 4).

The output generated by the IRS system is shown in Table 5. This report tells the user that 15 such cases involving 93 exhibits and 436 examinations were completed for Agency 100 after 1 Jan. 1978. Four of these cases (17 exhibits with 217 exams) were worked by Analyst B and eleven cases (76 exhibits with 219 exams) were completed by Analyst O. Other information is also readily available from the printed output. For example, the nature of the evidence is found in Field 8 (TPEV), the nature of the crime in Field 11 (TPC), and the length of time the case was actually in the laboratory can be determined by looking at the date received (Field 6) and the date case completed (Field 15).

The second example is an analytical type of file used to contain information on all bullets that have entered the laboratory for which no suspect gun was submitted. This open bullet file was created by using the description table shown in Table 6.

Tables 7 to 9 show a search of our firearms open bullet file by an analyst. A .38 Special revolver, yielding a test bullet with five lands and grooves, a right twist, land width of 0.088 ± 0.005 in., and groove width of 0.122 ± 0.005 in., was received in the laboratory for examination. Since the gun was seized from a suspect involved in several previous shootings, the open bullet file was searched to see if this gun could have been involved in any cases previously submitted to the laboratory in which no gun was available at that time. A preliminary search (not shown) on all .38 revolvers with a right twist showed the file to contain eleven such open cases. Of these, seven had five lands and grooves, two had six lands and grooves, and two had eight lands and grooves. The file was then searched with the land and groove measurements as two additional selection criteria. The input cards are shown in Table 7 and the logic expression thus generated is shown in Table 8.

TABLE 3—*Listing of input control cards for chemistry-physics file.*

```
DATE 040878 CASES COMPLETED BY ANALYSTS FOR AGENCY 100 AFTER 1/1/78
SELECTION CRITERIA = 2
FIELD = (A, 3) LIMITS = ('100', '100')
FIELD = (B, 15) LIMITS = ('010178', '040878')
A AND B
SELECT ALL AGENCY 100'S CASES COMPLETED AFTER 1/1/78
SORT 16 ASCENDING
TOTAL FIELDS 12 13 ON 16 WITH NM NM
PRINT FIELDS 2 3 6 8 11 12 13 14 15 16
```

TABLE 4—*Selectivity instructions for chemistry-physics file.*

```
SELECT ALL AGENCY 100'S CASES COMPLETED AFTER 1/1/78
BOOLEAN EXPRESSION
A AND B
BOUNDS SET "A" IS SUBMITTING AGENCY CODE
LOWER LIMIT 100
UPPER LIMIT 100
BOUNDS SET "B" IS DATE CASE COMPLETED
LOWER LIMIT 010178
UPPER LIMIT 040878
```

TABLE 5—*Report generated by IRS for the administrative example.*

CASES COMPLETED BY ANALYSTS FOR AGENCY 100 AFTER 1/1/78

THE FOLLOWING ARE THE COLUMN HEADINGS AND AN EXPANDED EXPLANATION.

COLUMN	FIELD NO.	HEADING	EXPLANATION
1	2	RCL	REGIONAL LAB CASE NUMBER
2	3	AG	SUBMITTING AGENCY CODE
3	6	DTRC	DATE CASE RECEIVED
4	8	TPEV	TYPE OF EVIDENCE CODE
5	11	TPC	TYPE OF CASE CODE
6	12	EX	NUMBER OF EXHIBITS
7	13	EXM	NUMBER OF EXAMS
8	14	DTDO	DATE CASE DUE
9	15	DTCT	DATE CASE COMPLETED
10	16	A	ANALYST

THE NUMBER OF RECORDS PRINTED ON THIS REPORT IS 15

DATE 04/08/78 PAGE 1

RCL	AG	DTRC	TPEV	TPC	EX	EXM	DTDO	DTCT	A
R78 0117	100	011078	2	006	01	005	020678	021478	B
R78 0652	100	022878	2	006	01	003	032878	031778	B
R78 5060	100	012878	13	111	09	191	082778	080378	B
R78 0026	100	010378	13	111	06	018	013178	020778	B
A					17	217			
R78 0253	100	012578	0	058	01	002	020878	030378	O
R78 0340	100	020178	5	058	08	028	022878	030378	O
R78 0415	100	020878	5	058	11	039	022278	030378	O
R78 0506	100	031678	30	058	02	004	030278	031578	O
R78 0523	100	021778	5	058	09	030	030378	030378	O
R78 0605	100	022478	5	058	04	017	031578	031578	O
R78 0737	100	030778	5	058	14	033	032778	031578	O
R78 0758	100	030878	5	058	08	028	032278	031578	O
R78 0533	100	022078	17	075	10	020	030678	031678	O
R78 0197	100	011478	0	087	03	003	020878	030878	O
R78 4468	100	092078	4	090	06	018	100878	100378	O
A					76	210			
GRAND TOTAL					93	436			

Of the eleven open cases, only one case successfully meets the conditions imposed by the analyst doing the search. This result as well as the relevant case information is shown in the computer output (Table 9). The analyst can then call the appropriate agency and request that the bullets previously submitted under that agency's case number 75-1321 be resubmitted for comparative analysis with the bullets obtained from the .38 Special in question. If the test firings from the .38 Special and the earlier bullets match, the analyst could provide the agency the missing weapon and a potential suspect on a case three years old. Each of the computer searches took less than 3 min on our diskette system. The same search using a disk system would have been approximately ten times faster.

TABLE 6—*Description table for open bullet file.*

Sort Field Number	Field Description	Column Headings on Report	Number of Columns per Field
1	CALIBER	CAL	3
2	TWIST	T	1
3	LAND AND GROOVES	LG	2
4	LAND WIDTH	L WIDTH	7
5	GROOVE WIDTH	G WIDTH	7
6	FIRING PIN	PIN	3
7	EXTRACTOR	EXTR	4
8	EJECTOR	EJEC	4
9	CHAMBER	C	1
10	BREECH OR BOLT	BOLT	4
11	REGIONAL CASE NUMBER	CASE	7
12	AGENCY CASE NUMBER	AGCASE	8
13	DATE OF OFFENSE	DATE	6
14	AGENCY	AG	2
15	ITEM NUMBER	ITEM	2
16	REMARKS	REMARKS	7

TABLE 7—*Listing of input control cards for open bullet file search.*

```
DATE 040878 SEARCH OF OPEN BULLET FILE
SELECTION CRITERIA = 4
FIELD = (A, 1) LIMITS = ('357', '357')
FIELD = (B, 2) LIMITS = ('R', 'R')
FIELD = (C, 4) LIMITS = ('083    ', '093    ')
FIELD = (D, 5) LIMITS = ('117    ', '127    ')
A AND B AND C AND D
SELECT CASES WITH BULLETS MEETING CRITERIA BELOW
PRINT FIELDS 1  2  3  4  5  11  12  13  14  15  16
END
```

TABLE 8—*Selectivity instructions for open bullet file search.*

```
SELECT CASES WITH BULLETS MEETING CRITERIA BELOW

BOOLEAN EXPRESSION
A AND B AND C AND D
BOUNDS SET "A" IS CALIBER
LOWER LIMIT 357
UPPER LIMIT 357
BOUNDS SET "B" IS TWIST
LOWER LIMIT R
UPPER LIMIT R
BOUNDS SET "C" IS LAND WIDTH
LOWER LIMIT 083
UPPER LIMIT 093
BOUNDS SET "D" IS GROOVE WIDTH
LOWER LIMIT 117
UPPER LIMIT 127
```

TABLE 9—*Report generated for the open bullet file search.*

THE FOLLOWING ARE THE COLUMN HEADINGS AND AN EXPANDED EXPLANATION.

COLUMN	FIELD NO.	HEADING	EXPLANATION
1	1	CAL	CALIBER
2	2	T	TWIST
3	3	LG	LAND & GROOVES
4	4	L WIDTH	LAND WIDTH
5	5	G WIDTH	GROOVE WIDTH
6	11	CASE	REGIONAL CASE NO
7	12	AGCASE	AGENCY CASE NO
8	13	DATE	DATE OF OFFENSE
9	14	AG	AGENCY CODE
10	15	ITEM	ITEM NUMBER
11	16	REMARKS	REMARKS

DATE 04/08/78 PAGE 1

CAL	T	LG	L WIDTH	G WIDTH	CASE	AGCASE	DATE	AG	IT	REMARKS
357	R	05	088	122	R759634	75-1321	122875	100	P	38 SP

Summary

The Wisconsin Regional Crime Laboratory, a relatively small laboratory with a total staff of 27 people, has met its computer needs by modifying the NLHE-IRS to work on a small computer purchased as part of an EDAX X-ray fluorescence energy dispersive data system. The retrieval system, once operational, does not require the services of a professional computer programmer or operator. The operational cost of this type of laboratory computer system is minimal since the computer's operating expenses are also included in the major instrument's operating expenses. This in-house data handling system gives even a small laboratory the opportunity to meet many of the needs expressed in the original CLIS program survey. The in-house capability has the great advantage of being 100% consistent with the laboratory's equipment, reference files, techniques, and even the individual peculiarities of the laboratory's analysts. The system saves much valuable time for both the analyst (searching data files) and the administrator (compiling statistics or budgets). During the creation of a search file the user is forced to adopt a standard format and to use standard coding techniques. Since these formats can be given to other laboratories within the state, data can be compiled from the state's entire laboratory system. Thus the IRS allows each state to create within its laboratory system a miniature CLIS-type program to correlate information and data from participating laboratories. The IRS thus becomes a strong complement to the interstate CLIS program, which is much more extensive and sophisticated. The IRS will function not only on a large computer system but also on a small computer with at least 8K memory, a disk or diskette system, and an ASR 33 teletypewriter. The Wisconsin Regional Crime Laboratory's success with the retrieval system on various pilot programs has encouraged us to extend its use to programs involving surrounding states.

Acknowledgments

The authors wish to thank the Carroll College Computer Science Center for their cooperation and suggestions during the computer program modification aspects of this project. We also would like to thank T. Ray Nanney, Furman University, and James E.

Runde, Clark College, for making the NLHE retrieval system available to us. A special thanks is given to Dr. Michael Fallgatter, Carroll College, for his suggestions during the writing of this paper and to the laboratory staff who were involved in the project and especially Irene Hoernke and Jeri Meils for the typing of this paper.

References

[1] "CLIS: Criminalistics Laboratory Information System," supplement to Technical Report 11, Vol. 1 through 4, Search Group, Inc., Sacramento, Calif., 1975.

[2] *User's Guide* and *Logic Manual*, Educational and Administrative Systems, Greenville, S. C., 1975.

[3] Nanney, T. R., "Data Independent Computer Programs and Information Retrieval in Small Educational Institutions," Furman University, Greenville, S. C., Dec. 1973.

Address requests for reprints or additional information to
Kent A. Oakes, Ph.D.
Wisconsin Regional Crime Laboratory
Wisconsin Department of Justice
New Berlin, Wis. 53151