A DATA MANAGEMENT SYSTEM FOR PROCESS EVALUATION IN A MENTAL HEALTH SETTING (UN SYSTEME D'ADMINISTRATION DE DONNEES POUR L'EVALUATION D'UN PROCEDE DANS UN CENTRE DE SANTE MENTALE)

H. D. Covvey, R. Cheng, M. Kotkin, G. F. D. Heseltine
University Hospital
University of Western Ontario and
Western Ontario Therapeutic Community Hostel
London, Ontario

ABSTRACT

The Department of Psychiatry at the University of Western Ontario is involved in a process evaluation project with an in-community treatment centre called the Western Ontario Therapeutic Community Hostel (W.O.T.C.H.).

A Database Management Program has been developed on the University of Western Ontario PDP-10 time-sharing facility. The package is written in FORTRAN and operates, at the present time, in batch mode. A simple QUERY system has been developed as well as a basic flexible report generator. A statistics package, partly based on SPSS is available for data analysis.

This package is still being developed and a variety of reports are being created for internal W.O.T.C.H. usage. We have begun development of a sister system for our in-house PDP-11/40. Utilities have been created to transfer the data from the PDP-10 structure to the PDP-11. Input to the PDP-11 will be via an interactive terminal (at W.O.T.C.H.) and an optical mark document reader (centrally located).

Our approach has been to use a hashed ISAM (hierarchical) structure. Records are linked with forward and backward pointers.

The system is proving to be fairly general purpose and a number of other groups are planning to use it for their databases in the near future. (Le département de Psychiatrie à l'Université de Western Ontario s'occupe d'un projet de procédé d'évaluation avec un Centre de Traîtement dans la Communauté du nom de: The Western Ontario Thérapeutic Community Hostel (W.O.T.C.H.).

DATA MANAGEMENT SYSTEM

On a developpé un Programme d'Administration de Données de base sur PDP-10 temps-partageant les facilités de l'Université de Western Ontario. Le paquet est écrit en FORTRAN et on l'emploie en ce moment en genre de tas. On a développé un simple système de QUERY en même temps qu'un rapport générateur simple et flexible. Il y a un paquet de statistiques, fondées en partie sur SPSS que l'on peut consulter pour l'analyse des données.

On est encore au point de développer ce paquet et on a créé une variété de rapports pour leur emploi à l'interieur de W.O.T.C.H. Nous avons commencé à travailler au développement d'un système pareil à celui-là pour notre "in-house" PDP-11/40. On a créé des facilités pour transférer les données de la structure PDP-10 à celles de PDP-11. On mettra le "input" au PDP-11 à l'aide d'une borne interactive (à W.O.T.C.H.) et un "reader" de document marqué opticalement (placé au centre).

Notre manière d'approche fut d'employer une structure hachée ISAM (hiérarchique). Les records sont reliés à des aiguilles qui vont en avant et en arrière.

Le système démontre un plan assez général et un bon nombre d'autres groupes ont l'intention de bientôt l'adopter pour leurs données de base.)

INTRODUCTION

The Western Ontario Therapeutic Community Hostel (W.O.T.C.H.) is a relatively new, in-community treatment centre in London Ontario. It is a paramedical team program offering support to the delivery of psychiatric services by family doctors, psychiatrists, and social agencies. It was initially financed by a Health Resources Development Grant (DM-13) from the Research and Planning Branch of the Ontario Department of Health. The W.O.T.C.H. program handles a large number of clients in a variety of therapeutic programs. Over twenty full-time staff are involved and hundreds of clients are seen on a regular basis.

As part of its program, W.O.T.C.H. has the responsibility for evaluating the cost, and, as much as possible the effectiveness of its programs.

Toward this purpose, the Department of Psychiatry has become involved with W.O.T.C.H. in a process evaluation program which is just now becoming active.

DATA MANAGEMENT SYSTEM

A set of forms has been created to capture data concerning the W.O.T.C.H. programs. These forms are of two types: one relating to the assessment of the client and the system of which he is a member (for instance, his family) and one for the purpose of recording therapeutic contacts with the client.

A batch-oriented database management system has been developed and is currently operational on the PDP-10 computer at the University of Western Ontario and is becoming operational on the PDP-11 computer at University Hospital.

Once the system is fully available on the PDP-11 computer, interactive access to the database will be provided.

A large amount of work is currently underway to create reports of the type required by W.O.T.C.H. in its evaluation project. Reports will either be in tabular format (as are our basic ones) or in graphic format as is currently under development.

Most of the W.O.T.C.H. program is implemented in FORTRAN with a few routines implemented in MACRO.

The data management package is relatively flexible and a number of other groups are currently planning to utilize its facilities. The most interesting of these is the Department of Psychiatry Out-patient Department at University Hospital. This is particularly important as we will be able to acquire data on a program similar to W.O.T.C.H. in that it is for ambulatory patients, but different from W.O.T.C.H. in that it is a hospital service.

To date only process data is being addressed. The program is intended actually to go the full route to outcome evaluation and cost-effectiveness studies. Only the beginnings of this exist, however, in a client-satisfaction assessment form.

GENERAL DESCRIPTION OF THE DATABASE SYSTEM

The W.O.T.C.H. Database System can be sub-divided into six sub-systems:

1. The File Management Sub-system
2. The Command Parsing Sub-system
3. The Query Sub-system
4. The Utility Sub-system
5. The Statistics Sub-system; and
6. The Report Generation Sub-system.

Of the six sub-systems, the first four are implemented and are operational. The fifth one has the most important routines operational meeting the immediate requirements of the user and the sixth is in its design phase. More than 5,000 records consisting of Social Assessment information and Contact Treatment information

are stored in the Database. Based on the stored information, statistics and basic reports are being generated.

Each patient's transaction involves between one and four computer cards. Transactions are of two types: 1. Patient Assessment (PA) record; 2. Treatment Contact (TC) record. The patient assessment information is recorded on four cards and a patient treatment contact produces one card. Record input and file updating are usually done at the beginning of each month. Patient information from the previous month is keypunched and batched into the system. The information on one computer card forms one information record in the Database.

The Database parser scans the first column of each card to determine whether the card contains a query command or a patient's record (data card). The first three characters of each patient information record designates the record type (i.e. assessment or treatment contact).

At the end of each file update, a simply statistical report is generated. It includes the number of new groups (a related collection of members) and new individual members introduced into the system; the number of groups or members discharged from the system; and the number of active groups in the system and their group members.

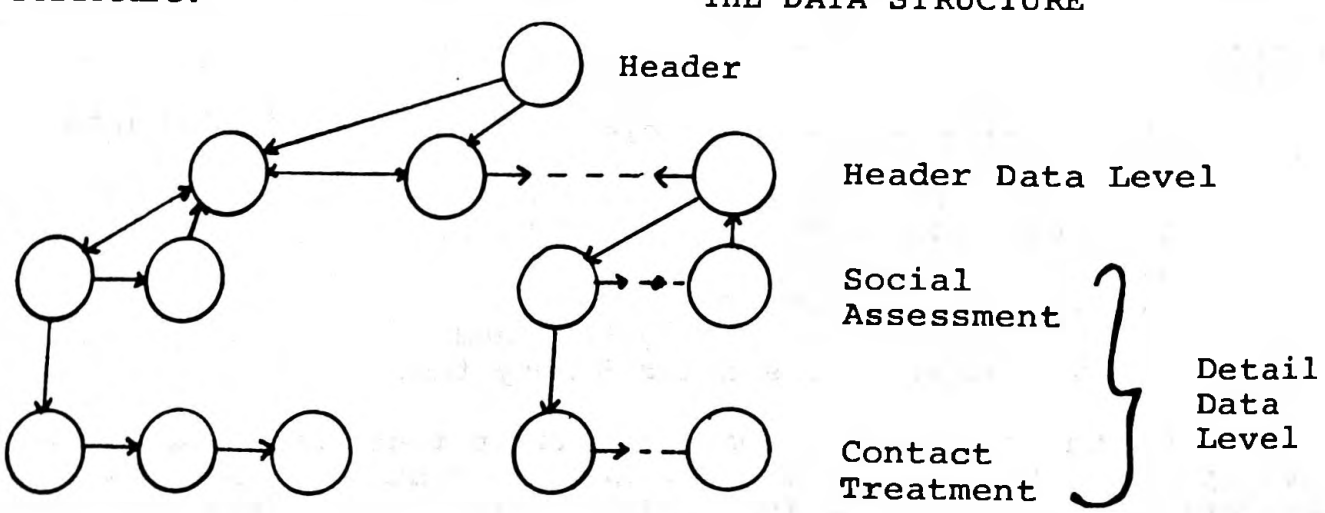
FILE MANAGEMENT SUB-SYSTEM

This sub-system has two parts, the data files and a descriptive table.

A. The Data Files

There are two types of data files. One is called the Header Data File and the other is referred to as the Detail Data File. These two types of data files are linked in a way forming a hierarchical structure (figure 1). Both Forward and Backward pointers are used to link records in the structure.

THE DATA STRUCTURE



DATA MANAGEMENT SYSTEM

All records in the Header Data Level (Social Assessment) are linked by both forward and backward pointers. Each Header Data Record has a pointer linking to its first member record in the Detail Data Level (Contact Treatment).

Each Detail Data Record has a forward pointer linking to the next member record and a backward pointer pointing to its Header Data Record.

The records in the second level (Contact Treatment) of the Detail Data Level only have the forward pointers.

Figure 1

In order to minimize the number of searches for a record in the Header Data level, a hashed ISAM technique is used based on the technique employed by J. K. Mullin (1971).

B. The Descriptive Table

A descriptive table is used to define the fields of each record in the Detail Information Data File (DIDF). It contains:

1. The field names (key words) - which are sorted into ascending order alphabetically.
2. The offset record number within the data block - when more than one physical data record are grouped together, the group is called a data block. (For example, the Social Assessment information takes up 4 physical data records)
3. The starting location of the field in the record.
4. The length of the field; and
5. The type of information in the field (either coding or raw data).

To retrieve a field from a DIDF, the key word (field name) must be supplied to the Database System. Descriptive information associated with the key word is retrieved by a binary search of the descriptive table. Using the descriptive information, data in the field can be retrieved accordingly.

THE COMMAND PARSING SUB-SYSTEM

The ASCII string on each card read into the W.O.T.C.H. system is scanned by the Command Parser. A period (.) in column one indicates that the card contains a command string. For a command string, only the first two characters after the period are parsed. A function code is returned to the system.

Parameters of the command function are parsed according to the function returned. All parameters (if there are any at all) must be enclosed by a set of round brackets. One limitation is that each command string must fit on one card (eighty characters). No continuation is allowed. Any card read into the system which does

not have a period in column one is treated as an information record. The first three columns of a data card designate the type of record.

THE QUERY SUB-SYSTEM

The Query Sub-system consists of all commands which are available in the W.O.T.C.H. Database system for retrieval purposes. A command is of the following format:

.FUNCTION (arguments or expressions)

optional

"FUNCTION" is a command to the system. There are eight different types of functions available presently. By means of these functions, basic reports and statistical output can be generated from the Database.

A FUNCTION may or may not have parameters. The parameters can either be arguments or expressions depending on the type of FUNCTION. An argument is any key word from the descriptive table. An expression can be composed of key words, values, and arithmetic operators or boolean operators (Table 1). Although work is underway to extend this, presently nested boolean operations are not permitted.

Table 1

⟨COMMAND⟩ ::= |. END |.⟨FUNCTION⟩ |

⟨FUNCTION⟩ ::= | ⟨QUERY⟩ (⟨EXPRESSION⟩) | ⟨DISCHARGE⟩ (⟨EXPRESSION⟩) |
⟨FORMAT⟩ (⟨PARAMETER⟩) | ⟨SORT⟩ (⟨PARAMETER⟩) |
⟨CROSS TABULATION⟩ (⟨PARAMETER⟩) | ⟨CODEBOOK⟩ (⟨PARAMETER⟩)
⟨PLOT⟩ (⟨PARAMETER⟩) |

⟨EXPRESSION⟩ ::= | ⟨ARGUMENT⟩ ⟨ARITH OPR⟩ ⟨ARGUMENT⟩ |
⟨EXPRESSION⟩ ⟨BOOLEAN OPR⟩ ⟨EXPRESSION⟩ |

⟨PARAMETER⟩ ::= | ⟨ARGUMENT⟩ |, ⟨PARAMETER⟩ |

⟨ARITH OPR⟩ ::= | = (Equal) | < (Less Than) | > (Greater Than) |

⟨BOOLEAN OPR⟩ ::= | + (Or) | - (Not) | & (And) |

⟨QUERY⟩ ::= [QU]ERY

⟨DISCHARGE⟩ ::= [DI]SCHARGE

⟨FORMAT⟩ ::= [FO]RMAT

⟨SORT⟩ ::= [SO]RT

⟨CROSS TABULATION⟩ ::= [CR]OSSTABULATION

⟨CODEBOOK⟩ ::= [CO]DEBOOK

⟨PLOT⟩ ::= [PL]OT

⟨END⟩::= [EN]D

⟨ARGUMENT⟩::= pre-defined argument list or a constant value

THE UTILITY SUB-SYSTEM

Once a data structure is created it can be maintained on the PDP-10 or converted from a PDP-10 file structure to a PDP-11 file structure by five utility routines presently available:

1. Data Structure Mapping Routine (DSMR) Since W.O.T.C.H. was developed on the PDP-10 and is to be transportable to the PDP-11, the data structure on the PDP-10 has to be converted to the one supported on the PDP-11. The mapping routine is written on the PDP-11.
2. File Merging Routine (FMR) For debugging and file back-up purposes, each monthly transaction can have the choice of being input directly to the system master files or entering a temporary data file which can be merged to the system master later on. The process of file merging is done by FMR.
3. Record Editing Routine(RER) It is impossible to prevent incorrect information from entering the Database. Errors may be caused by bugs, keypunch errors, or incorrect raw data. One of the ways of correcting such a mistake is to delete the whole record or file from the data structure and then reconstruct it again. This is tedious if only single data items are involved. RER enables the user to open any data file in the data structure. It reads which field is to be corrected from the user's terminal. A pointer is used to locate the byte in the data file that is going to be changed and the new user input byte replaces the one in the data structure.
4. ASCII Dumping Routine (ADR) All records in the PDP-11 data structure are unformatted. ADR dumps out a variable number of records from any data file designated by the user to the line printer in ASCII mode.
5. Garbage Collection Routine (GCR) Any record deleted from the system is not physically deleted. Only a flag is set. When the GCR is run it physically deletes the flagged records and re-arranges the records in the Database. Records under the same group number are placed together physically, and the groups are sorted into ascending order. This helps to minimize the disk head movement.

THE STATISTICS SUB-SYSTEM

The Statistics Sub-system of the W.O.T.C.H. system is partially completed. The available routines include Cross-tabulation (CROSSTAB), Frequency Distribution (CODEBOOK), and basic plotting on a terminal (PLOT). The commands of the statistics sub-system belong to the Query Sub-system. Output from this sub-system is similar to that employed in the SPSS package (Nie 1970).

THE REPORT GENERATION SUB-SYSTEM

In the medical environment, reports are required for inter-departmental communication. A sub-system which has the capability of transforming the information from the Database structure to an easily understood English form is required. This will involve string manipulation and sentence concatenation. An interpreter for the report generation sub-language will be required also. We are involved in research in this area currently. NOVEL (Laskin) is being used as a model.

The FORMAT command provides us with basic tabular reports at the present time.

CONCLUSION

A great deal of development is still being done especially on output from the database. Work is also proceeding on the use of mark-sense forms for data entry. Although the system is in test-mode use, it will not be fully operational in terms of all required reports until late Spring of 1976.

REFERENCES

1. Mullin; J. K. An Improved Index Sequential Access Method Using Hashed Overflow, CACM, 15, 5 May 1971, p. 301-307.

2. Nie, N. H., et al. SPSS, McGraw Hill Book Company, 1970.

3. Laskin, E., et al. NOVEL: Multi-State Information System, Rockland State Hospital (Brochure).