

FAMULUS: A COMPUTERISED RETRIEVAL PACKAGE

JOHN KIRKLAND

*Education Department
Massey University*

FAMULUS is a set of computer programmes which I brought out from University College, London. These programmes were written in IBM Fortran and have recently been rewritten in Burroughs Fortran by Ross Turner at Massey University's Computer Centre. The package is now running on a Burroughs 6700 machine. A Users Manual is available from the Massey Computer Centre at a cost of approximately \$2.50. A copy of the programmes can be made to users supplying a tape (\geq 60m; 9 track; specify 800 bpi, or 1600 bpi PE).

Although FAMULUS was initially designed as a system for establishing, maintaining, and retrieving bibliographic records it can be used wherever similar data are to be searched, sorted, counted and printed. Thus, these programmes enable users to construct a data base that could contain information pertaining to: personnel files, research literature, departmental library, or other. In short, FAMULUS is a file manipulation package that provides the basic framework for several types of content. For those familiar with other computerised retrieval services (such as Psychological Abstracts; Science Citation Index; ERIC) FAMULUS may be likened to the skeleton for carrying out similar operations on your own data base.

A typical use is in the handling of personally collected references of research literature. In this case each record consists of a single citation. Records may be up to ten free-format fields with four-character names selected by the user. For example in one file the following field names could be assigned: AUTH (for author); TITL (of article or book title); (source journal); SOED (source, edited volume or book, author(s) names if different from AUTH entry, and title if other than TITL entry, as well as place of publication, publishers name, and date); VLPG (volume and page numbers); DATE (of journal); ABST (abstract); AFFL (instructional affiliation of the reprint author). Each record may consist of up to 3,500 characters (around 40 punched cards), enough space for entering most abstracts in full along with the other data. Field sizes are unlimited for a particular record providing the record limit of 3,500 characters is not exceeded. FAMULUS holds information in sequential unformatted data sets on either magnetic tape or disk storage. Files may be of unlimited length.

There are four principal operations available with FAMULUS: Data management; data manipulation; data retrieval; and thesaurus

construction. Each operation consists of several programmes the names of which are printed in upper-case characters in the following outline.

Data Management: The EDIT programme creates, corrects, and updates files. As Burroughs CANDE can be used for entering, correcting, and updating files from a terminal more easily than using conventional punched cards, EDIT is used mainly to convert the work-file into machine readable form for the remaining FAMULUS programmes. OSSIFY does the reverse of EDIT, it creates a card-image back-up file. This makes it possible to exchange data bases with other researchers who might be using different computers.

Data Manipulation: SORT enables a user to sort the entire file. Identical records sorted on the first nominated field are sorted according to the second nominated field, and so on. These fields can be specified in any order.

MERGE will combine separate files into a single one. This is useful where members of a research team each create separate files and a master file is needed.

MULTIPLY will expand a field, such as AUTH, so that each entry is used to create a separate record.

Data Retrieval: GALLEY prints a file in a choice of formats. INDEX prints an alphabetical listing of all words contained in a specified field and provides a reference to the records in which the indexed words appear.

SEARCH enables a user to specify search word combinations (using "and", "or", and "not"). Records matching the search formula (e.g. child and (school or home) and not play) are printed. Any number of fields can be searched simultaneously in this manner.

Thesaurus Construction: VOCAB prints an alphabetical list of all words in a nominated field.

COUNT as for VOCAB, but includes statistics of word occurrence etc.

KEY is an automatic key wording facility which scans selected fields and deposits words into a specific field for each record. Thus a "descriptor" field can be treated automatically to which words can be added to or deleted from later.

KWIC (Key Word in Context) provides, for a selected field, an alphabetical listing of words (a "stop" or "go" list can be used to restrict printing of trivial words) down the centre of the page in heavy type with as much content as can be fitted onto the same line on either side of that word. Each line is numerically referenced to the file.

One particular advantage of these FAMULUS programmes is that they can be used by unsophisticated personnel. An inexperienced person can quickly learn how to set up the control cards and execute the programmes. Clearly, this is an important factor which could encourage an extended use of the service.