

Glossary

16-bit processor microprocessor whose **word length** (q.v.) is 16 bits

6502 an 8-bit microprocessor

68000 a 16-bit microprocessor

access point one of the terms by which a record has been indexed to enable it to be accessed

accession number a number uniquely identifying each item, copy or volume acquired or held by a library

access time the time between requesting data and obtaining it (from the disc)

acronym key see **derived key**

ADABAS a **DBMS** (q.v.) (used by the OPAC MELVYL)

added entry an added entry heading is a name, class mark, title etc. associated with a book in addition to the main author, class mark, title etc., which is a potential additional catalogue access point; cf. **main entry**, see also **heading**

address a number that identifies a storage location (also verb)

address calculation storage see **hashing**

algorithm a finite well-defined procedure for solving a problem

alphanumeric alphabetic or numeric, (sometimes also blank or punctuation)

analytical entry an entry for part of an item for which there is also a comprehensive entry

Apple IIe one version of a widely sold microcomputer made by Apple Computers, USA

archiving making copies of (disc) files, usually on tape, as a security measure

ASCII American Standard Code for Information Interchange, a widely used 7-bit code for representing characters

assembler a program that translates assembly code into machine code; also used, loosely, for **assembly code**

assembly code, language low level mnemonic programming language

authority file file giving the accepted form of headings in a catalogue

B-tree type of **tree structure** (q.v.) well suited for storage of an index on disc

backup copy a copy of a file kept for security purposes

batch a set of data that is to be processed during one run of a program, hence “batch process”, “batch run”; cf. **real time**

baud measure of communication speed, usually a bit per second

beast Okapi jargon coined during discussion of index design: “What sort of beast is it?”. The role(s) played by a term in a particular context (e.g. “wood” could be a surname, a whole personal name, a title word, a whole title, a subject, one word in a corporate name, etc.)

biblish library jargon for library jargon

binary chop, search the fastest method of searching an ordered list: the search key is compared with the midpoint, after each comparison half of the remaining list can be ignored

bit a binary digit, usually represented by 0 or 1

BLAISE British Library Automated Information Service

blank a character code which, when output, causes the cursor or print head to advance one position without displaying or printing

block 512 bytes for Okapi (more generally, the amount of data input or output at one time, usually from/to disc or tape)

BNB the British National Bibliography

Boolean logic or algebra using the operators AND, OR and NOT

BRS/SEARCH an online retrieval service and software package

buffer storage area for input or output data

byte smallest addressable unit of storage, the number of bits needed to represent a character, usually 8 bits

C a high level programming language

CAIRS an IR system developed by Leatherhead Food Research Association

card a circuit board that is plugged into a piece of equipment such as a micro

carriage return special character used to ensure that the next character appears at the beginning of the line; also, often the key a user is required to press to signal the end of some input

CCR Centre for Catalogue Research at Bath University

central processing unit the part of a computer that executes the instructions in the correct sequence

character a letter, digit, punctuation mark, blank or other symbol, or its coded representation

character string a sequence of characters

chip a small package of integrated electronic circuits often made of silicon

chunk 2048 bytes (Okapi jargon)

citation a reference

CITE advanced OPAC developed and used at the National Library of Medicine, USA

class mark a number assigned to a particular subject according to a scheme for classifying knowledge; see also **Dewey**, **DC** and **UDC**

CLR Council for Library Resources (USA)

CLSI a supplier of automated library systems

Cluster/One a small LAN made by Nestar/Zynar (no longer obtainable)

COBOL a widely used commercial programming language

code (noun) (1) low level programming language; see also **assembly code** (2) the actual text of a program; see also **source code** and **object code** (3) a method of representing data; e.g. **ASCII**; (also verb)

COM computer output microfilm/fiche

command language one method of providing user interaction with a system: a language, usually consisting of a few commands in a precise format; cf. **menu**

compiler a program that translates programs as written into programs that can be executed; see also **assembler**, **source code**, **object code**, (also verb: to compile)

computerese computer jargon for computer jargon

content addressable storage that is identified by content rather than address

control number a number uniquely identifying a (MARC) record

core (memory) fast main storage

corporate name a non-personal name such as the name of an institution

CP/M an operating system available on a wide range of microcomputers

CPU see **central processing unit**

CRLF carriage return and linefeed

CRT cathode ray tube, converts electrical signals to visible display

cursor the marker on a VDU screen showing where the next input character will appear

data compression methods of encoding data to save storage space

Datatar a European dial-up bibliographic retrieval system

DBMS database management system. Software to create, maintain and access a database

DC, DDC Dewey Decimal Classification scheme; see also **UDC**

debugging finding and correcting errors in a program or system

DEC-10 a mainframe computer made by Digital Equipment Corporation (a large USA computer firm)

derived key term that is formed from one or more of the original fields, e.g. an acronym made from the first few characters of an author and a title

descriptor term assigned to record for retrieval purposes; see also **keyword**

Dewey classification scheme devised by Melvil Dewey; see also **DC**

DIALOG IR service operated by Lockheed

dictionary ordered list of terms; see also **index**

digram a letter pair (e.g. "re")

directory locates, and may also define, the fields in a record or the records in a file or database

disc, disk fast peripheral magnetic storage device, floppy or rigid, single or multiple, fixed or interchangeable; see also **Winchester**

disc bound state in which the performance of a system is limited by disc speed; cf. **processor bound**

disc drive peripheral device for reading/writing discs

disc operating system operating system (q.v.) for a system with disc drives, which is therefore resident on disc

double-word 4 bytes in Okapi context

dump (noun and verb) copy for backup or debugging purposes

EBCDIC an 8-bit character code, Extended Binary Coded Decimal Interchange Code; cf. **ASCII**

EMMA Extra MARC MAterial

entry (1) a record of an item in a catalogue, (2) an access point

exchange tape magnetic tape of MARC data in format suitable for transferring between systems

false coordination one of the causes of **false drops** (q.v.) in free text Boolean IR systems, e.g. "primary education" might retrieve "The primary purpose of education .."

false drop totally irrelevant record retrieved in a search of an IR system; see also **noise**

field a single piece of information in a record, sometimes divided into subfields; see also **tag**

file a sequence of related records, (also verb)

file server a processor that is dedicated to handling disc I/O in a LAN

firmware program(s) that is/are permanently resident in a computer

FIRS II a microcomputer IR system

fixed field a data field which occupies a fixed number of bytes, and whose location is therefore predefined; cf. **variable field**

form filling a method of entering data in which the screen is laid out with sections to be filled in by the user

free text IR system one in which retrieval is effected partly or only by using the text of the document or reference itself rather than controlled descriptors or index terms

front end see **intelligent front end**

Geac Canadian supplier of library automation systems

go list a list of terms to be included in an index; cf. **stop list**

hardware physical equipment

hash storage non-sequential method of storing records in which, instead of using an index or directory, an algorithm is used to derive the address from the data, hence "hashing"

hashing see **hash storage**

heading a name, title, number, subject etc. that is entered in a catalogue or index, i.e. an access point

hidden Booleans see **implicit Booleans**

hit IR jargon for a record retrieved in a search

IBM PC IBM's personal computer (a 16-bit microcomputer)

IFE see **intelligent front end**

implicit Booleans Boolean operations not explicitly requested by the user

index a reference list of a file, record or document, (also verb); see also **dictionary**

index term a term i.e. name, word, phrase, number in the index

indexing the process of selecting or constructing items to form an index

indicator in a MARC record, one of two digits associated with a tag, to subdivide it, relate it to another tag, or give the number of non-filing characters in the field (e.g. the length of a leading article plus a blank)

input (noun) data going into a system, (verb) to get data into a system, (also used as an adjective)

intelligent front end an intelligent device between a user and a system to make the system more usable

interactive system a system in which the user and computer can interact in a conversational mode in **real time** (q.v.); cf. **batch**

inverted file a list or index that provides access to a file via some or all of its contents and not just the keys by which it is ordered, hence "inversion"

inverted index see **inverted file**

invisible invisible (cf. **transparent**)

I/O input/output

IR information retrieval

IRish IR jargon

ISBN International Standard Book Number, a unique 10-digit number assigned to published books

iteration repetition of a process until a condition is satisfied

K,k 1024 i.e. 2^{10}

KB,kb kilobyte, 1024 bytes

key (1) a term i.e. name, word, phrase, number in the index, or being searched for, (2) a button on a keyboard, (also verb: to key, meaning to press a button on the keyboard)

keyword see **term, descriptor**

kilobyte 1024 bytes

known item search a search for a specific book as opposed to searching for books on a given topic

LAN local area network: a group of computers, usually in one building, connected in a fast network to each other and to shared resources such as disc(s) and printer(s)

LC Library of Congress (USA), and sometimes LC classification

LCC Library of Congress Classification

LCCN Library of Congress Card Number

LCSH Library of Congress Subject Heading

level a number associated with MARC tags to identify elements of a subrecord, e.g. for an analytical entry

LID, line isolation device used to connect stations to the PLAN network

limit commands facilities in an IR system for restricting retrieved items by date, location, material type etc.

linefeed special character used to ensure that the next character appears on a new line

linking process of combining routines in object code to form a complete program

list (verb) to print out program text or data, (noun) types of data structure consisting of items linked to each other in a specific order

local area network see **LAN**

LOCAS (British Library's) LOcal CAtaloguing Service

logging automatic recording of use of a system for later analysis, usually for evaluation purposes

machine code coded instructions executable by the CPU

machine-readable in a form that can be input to a computer, e.g. on magnetic tape

main entry a catalogue record in the preferred form in which it is intended that the work should be cited; loosely, the heading which is used in the main entry (see also **added entry**, **heading**)

mainframe used to mean a large computer (usually serving a whole institution rather than a department or an individual)

MARC MACHine Readable Cataloguing, one particular format and tagging system for bibliographic data

MB, mb see **megabyte**

megabyte 1024 kilobytes i.e. 1,048,576 bytes (about a million)

MELVYL an OPAC developed at the University of California

menu method of implementing user interaction in which the user is offered a selection of options; cf. **command language**

merge combine two or more ordered sequences into a single ordered sequence

MeSH Medical Subject Headings

micro(computer) a desktop computer

microprocessor the CPU of a microcomputer, usually classified by word length i.e. 8-bit, 16-bit etc.

minicomputer used to distinguish a computer that is bigger than a micro and smaller than a mainframe

mouse desktop device whose movements are copied by the cursor

multi-access providing access to more than one user at the same time

Nestar US partner in development of Cluster/One and PLAN; see also **Zynar**

network a system of interconnected computers and shared devices; see also **LAN**

n-gram a group of n letters; see also **digram**, **trigram**, **tetragram**

noise (1) electronic interference, (2) (loosely) spurious information resulting from unpredictable causes

- object code** program code after translation into **machine code** (q.v.)
- OCLC** Online Computer Library Center, Ohio
- Okapi** acronym for our OPAC, chosen because the okapi is a rare, shy, fast-moving animal with a long gestation period: Online Keyword Access to Public Information
- on-line, online, on line** connected to a computer
- OPAC** Online Public Access Catalogue: computerised library catalogue for library users
- operating system** software, usually supplied with a computer, to control the running of user programs
- output** (noun) the results of a computer process, (verb) to transfer these results to disc, or to display or print them, etc. (also used as an adjective)
- PAC** see **OPAC**
- padding** the extending of a field (often with blanks) to a fixed length or to the next multiple of n bytes
- Paperchase** an OPAC developed at the Beth Israel Hospital, Boston, Mass
- pass** a complete scan through a file
- pathname** identification of a disc volume or file in a tree-structured or hierarchical naming system
- PCL** Polytechnic of Central London
- peripheral** a device, usually for I/O, that is peripheral to the CPU
- PLAN 4000** the largest LAN marketed by Nestar/Zynar and the LAN used for the Okapi project
- PL/I, PL/1** a programming language suitable for text processing among other things
- port** socket or other component of computer or peripheral by which connection can be made to another device
- positional operator** a word or symbol in a command language, which is used for term proximity searching (q.v.), e.g. ADJ used to denote that two terms must be adjacent in retrieved records
- post-coordinate** IR system in which descriptors are combined when a search is made; cf. **pre-coordinate**
- posting** in Okapi, the address of a record containing a given term as a particular **beast** (q.v.)
- PRECIS** PREServed Context Index System, a subject indexing system designed for automatic manipulation
- precision** in IR, the ratio of the number of relevant records retrieved to the total number of records retrieved in a search; see also **recall**

pre-coordinate IR system in which descriptors are pre-assigned, and cannot be combined at search time; cf. **post-coordinate**

primary index the upper level Okapi index containing a unique entry derived from the last term in each pair of chunks in the secondary index

procedure loosely, a well-defined part of a program that carries out a specific function; also **routine**

processor see **central processing unit**

processor bound state in which performance is limited by speed of the processor; cf. **disc bound**

program a sequence of instructions in a given programming language to perform a specific job

PROM Programmable **ROM** (q.v.), i.e. ROM that can be programmed by the user; see also **RAM**

prompt request by the system for some input from the user (noun and verb)

protocols (1) rules governing the flow of information in a system, (2) conventions for network communication

proximity see **term proximity searching**

query see **search statement**

query language those parts of a command language (q.v.) which are used for formulating search requests to (e.g.) an IR system

RAM random access memory: memory that can be written to as well as read from; see also **ROM**

read see **input**

real time a system that operates at a speed that is related to the speed of the user or process being served

recall in IR, the ratio of the number of relevant records retrieved to the total number of relevant records in the file being searched; see also **precision**

record a collection of data relating to one entity

relevance feedback using searchers' assessments of the relevance of information retrieved during a search (of an IR system) to improve the (system's or user's) formulation of the search

response time the time taken by the system to reply to the user

return often the key a user is required to press to signal the end of some input; see also **carriage return**

ROM read only memory: used to store permanent information; see also **PROM**, **RAM**

routine (in programming) see **procedure**

run-time (1) the time while the program is running, (2) the elapsed time taken for a program to run

scatter storage see **hash storage**

search state current type of search (during a user session), e.g. author, subject, author/title

search statement expression of a search or query in the language and format of an IR system

secondary index the lower level Okapi index containing actual terms and postings

sector part of a **track** (q.v.) on a disc, the smallest amount that can be read/written in one operation

send often the key a user is required to press to signal the end of some input; see also **return**

session a sequence of searches (of an IR system) by one user at one time

software applications programs, operating systems, compilers, editors, etc.; cf. **hardware**, **firmware**

source code program code as written, before translation into **object code** (q.v.)

source file the main input file to a system

space see **blank**

station an I/O device, often a terminal, in a network

stop list a list of **stop words** (q.v.)

stop word a non-significant word (e.g. “and”, “the”) ignored during index creation, hence “stop list” and the verb “to stop” (cf. **go list**)

subfield in a MARC record, a logical subdivision of a tagged field

subject heading the name of a topic used as an access point in a catalogue

SWALCAP South-Western Academic Libraries Cooperative Automation Project; a provider and supplier of library automation services and systems

tag in a MARC record, a three digit number used to identify each field type (often used loosely to mean the field itself)

term see **index term**

term proximity searching the ability (of an IR system) to allow searches which specify that two terms are to be within a given relative position to each other in the retrieved records

terminal a device for one user to send input to and receive output from a system. Usually a VDT or microcomputer

- tetragram** a group of four letters (e.g. "econ")
- tokens** see **types and tokens**
- touch screen** a VDU which responds to its screen being touched (with a finger or other object)
- tracing** a statement of the headings under which an item appears in a catalogue
- track** circular recording path on a disc surface; see also **sector**
- transparent** self-evident (cf. **invisible**)
- tree** used to represent hierarchical or branching structure
- tree (storage)** see **tree**
- trigram** a group of three letters (e.g. "ing")
- truncation** right truncation is the use of a phrase or word stem, e.g. "econom", in searching, in order to combine terms which begin in the same way; (there is also left truncation)
- types and tokens** type: a natural language word; token: an occurrence of a word. "The cat sat on the mat" contains five types and six tokens
- UDC** Universal Decimal Classification scheme; see also **DC**
- UNIX** an operating system now widely available for minis and micros
- variable field** a field whose length varies with the actual data held, and which usually entails the use of a directory to locate fields and/or field terminator characters (cf. **fixed field**)
- VDU** Video/Visual Display Unit; see **VDT**
- VDT** Video/Visual Display Terminal, a device with a CRT screen and keyboard, or other input device, used for I/O between a user and a computer system
- Winchester disc** widely used type of hard disc
- write** see **output**
- word** 2 bytes for Okapi (more generally, a group of bits or bytes treated as an entity by a given computer)
- word length** the number of bits in a **word** (q.v.)
- WordStar** a word processing system for CP/M machines
- Z80** an 8-bit microprocessor
- Zynar** British partner in development of Cluster/One and PLAN, now called Nestar