BIREME / PAHO / WHO

Latin American and Caribbean Center on Health Sciences Information

Basic concepts of CDS/ISIS databases: an introduction to the use of CISIS

Version 3.x

Sao Paulo - 2007

Basic concepts of CDS/ISIS databases: an introduction to the use of CISIS

Card Catalog

BIREME / PAHO / WHO (Brazil)

Basic concepts of CDS/ISIS databases: an introduction to the use of CISIS. / BIREME (org.). Sao Paulo : BIREME / PAHO / WHO, 2007.

30 p.

1. User manual. 2. Information access. 3. Information systems. 4. Information management. 5. Public health. 6. Public Health services. I. BIREME II. Title

Warning - Any mention in this document to companies, institutions, persons or products are not an endorsement or recommendation given by BIREME / PAHO / WHO, thus it does not mean a preference to a similar one, cited or not.

BIREME / PAHO / WHO

Latin American and Caribbean Center on Health Sciences Information

Rua Botucatu 862 V Clementino

This document was produced with the Documents Conformation Methodology (NorDoc) developed by BIREME.

Table of contents

Abbreviations used

- **ANSI**. American National Standards Institute.
- **ASCII**. American Standard Code for Information Interchange.
- **BIREME**. Latin American and Caribbean Center on Health Sciences Information.
- **CDS**. Computerized Documentation System.
- **CP**. Code Page.
- **FST**. Field Selection Table.
- **FTP**. File Transfer Protocol.
- **IFP**. Inverted File Pointer.
- **ISIS**. Integrated Set of Information Systems.
- **ISO**. International Organization for Standardization.
- **LILACS**. Latin American and Caribbean Health Sciences Literature.

- **PAHO**. Pan American Health Organization.
- **UNESCO**. United Nations Educational, Scientific and Cultural Organization.
- **VHL**. Virtual Health Library.
- **WHO**. World Health Organization.

1 Preface

1.1 About BIREME

Year after year, BIREME has been following its mission of being a center dedicated to scientific and technical health information for the region of Latin America and the Caribbean. Founded in Brazil in 1967, under the name of Regional Medicine Library (which the acronym BIREME comes from), it has always met the growing demand for up-to-date scientific literature from the Brazilian health systems and the communities of healthcare researchers, professionals and students. Then, in 1982, its name changed to Latin-American and Caribbean Center on Health Sciences Information so as to better express its dedication to the strengthening and expansion of the flow of scientific and technical health information across the region, but kept the acronym.

Networking, based on decentralization, on the development of local capacities, on sharing information resources, on developing cooperative products and services, on designing common methodologies, has always been the foundation of BIREME's technical cooperation work. It has been like this that the center established itself as an international model that fosters professional education with managerial and technical information with the adoption of information and communication paradigms that best meet local needs.

The main foundations that gave origin and which support the existence of BIREME are following:

- ✓ access to scientific and technical health information is essential for the development of health;
- ✓ the need to develop the capacity of Latin American and Caribbean countries to operate their sources of scientific-technical health information in a cooperative and efficient manner;
- ✓ the need to foster the use and to respond to the demands for scientific-technical health information from governments, health systems, educational and research institutions.

BIREME, as a specialized center of the Pan-American Health Organization (PAHO)/ World Health Organization (WHO), coordinates and conducts technical cooperation activities on the management of scientific information and knowledge with the aim of strengthening and expanding the flow of scientific health information in Brazil and in other Latin American and Caribbean countries as a key condition for the development of health, including its planning, management, promotion, research, education, and care.

The agreement that supports BIREME is renewed every five years by the members of the National Advisory Committee of the institution (PAHO, Brazilian Ministry of Health, Brazilian Ministry of Education and Culture, Secretary of Health of the State of São Paulo, and Federal University of São Paulo – Unifesp). The latter provides the physical infrastructure necessary for the establishment of the institution.

In 2004 the institution took on the responsibility of becoming a knowledge-based institution.

1.2 The Virtual Health Library (VHL)

With the rise and consolidation of the internet as the prevailing means of access to information and communication, BIREME's technical cooperation model evolved,

as of 1998, to build and develop the Virtual Health Library (VHL) as a common space for the convergence of the cooperative work of producers, intermediaries, and users of information. The VHL promotes the development of a network of sources of scientific and technical information with universal access on the internet. For the first time there has been a real possibility of equal access to health information.

To BIREME, the Virtual Health Library is a model for the management of information and knowledge, which includes the cooperation and convergence between institutions, systems, networks, and initiatives of producers, intermediaries, and users in the operation of networks of local, national, regional and international information sources favoring open and universal access.

Today, every country in Latin America and the Caribbean (Region) participates either directly or indirectly in the cooperative products and services offered by the VHL, which includes over 1,000 institutions in more than 30 countries.

The VHL is simulated in a virtual space of the internet formed by a collection or network of health information sources in the Region. Users of different levels and locations can interact and navigate in the space of one or many information sources, regardless of where they are. Information sources are generated, updated, stored and operated on the internet by producers, integrators, and intermediaries, in a decentralized manner, following common methodologies for their integration in the VHL.

The VHL organizes information in a structure that integrates and interconnects reference databases, specialist directories, events and institutions, a catalogue of the information resources available on the internet, collections of full texts with a highlight for the SciELO (*Scientific Electronic Library Online*) collection of scientific journals, selective information dissemination services, information sources to support education and decision-making, news, discussion lists, and support to virtual communities. The space of the VHL is, therefore, a dynamic and decentralized network of information sources based on which it is possible to retrieve and extract information and knowledge to support health decision-making processes.

The Virtual Health Library can be visualized as a distributed base of scientific and technical health knowledge that is saved, organized and stored in electronic format in the countries of the Region, universally accessible on the internet and compatible with international databases.

2 Introduction

If you have a relevant collection of information, it is important to preserve it for later consultation.

This collection of information or data is a prime candidate to be part of a database.

In more traditional terms, a databank is considered as a collection of databases which, in general, are related to each another.

The basic element of a database is the record, which should be considered as the minimal parcel of information that stands on its own. In turn, this record is composed of elements of data, which occupy fields and represent a smaller unit of information, which has no significance if considered in isolation. Further, the contents of a field can be composed of smaller elements which are located in subfields. Figure-1 illustrates the situation described.

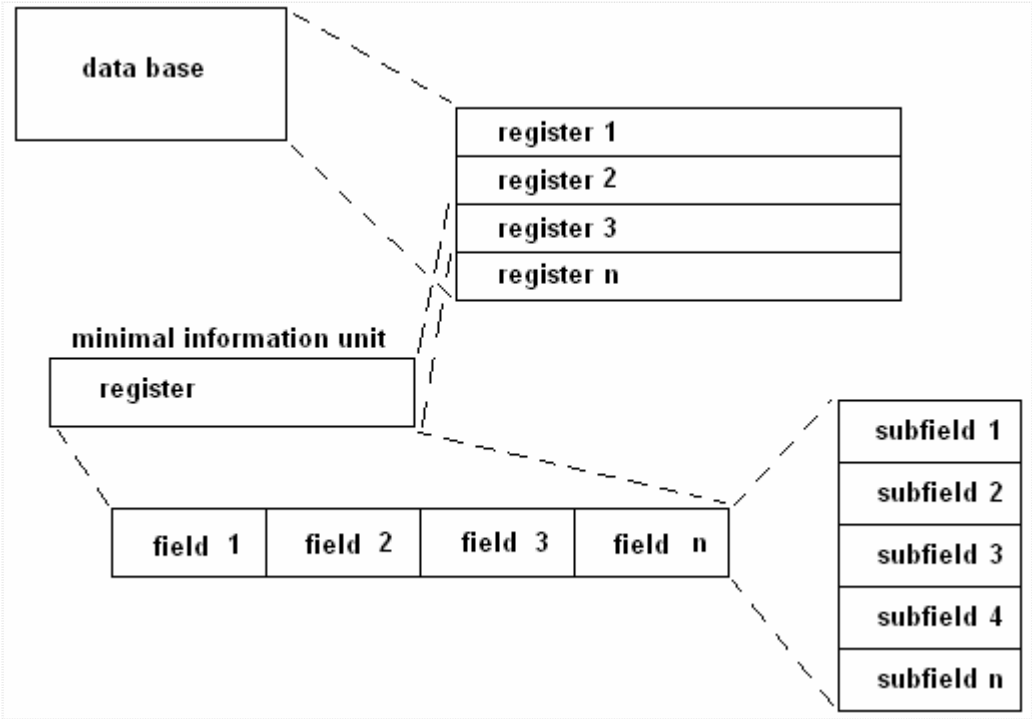


Figure-1: Composition of a database

For the sake of example, each record in Figure-1 represents a catalogued document, each field corresponds to an element of data such as author, title, or theme. Some fields can be composed of subfields, e.g. the author field could be composed of surname, forename, institution to which he/she is affiliated, department, etc.

There are diverse systems for managing databanks and databases; currently the relational system is considered as “state of the art” in the management of databanks. It is based on the idea of a list of tables, between which are established logical relations.

However, in the function of normalization of information, the relational model proves not adequate to the task, since the data are not amenable to normalization, as required by the relational model.

With the type of information and data which represent a bibliographic reference (basically, textual information) it is necessary to adopt a more flexible database model, such as one that accepts records of variable length and admits the existence of fields that repeat in the same record.

3 CDS/ISIS Databases

A CDS/ISIS database is characterized by:

- accepting fields of variable length
- accepting repeatable fields
- permiting the use of subfields
- permiting relations between distinct databases
- employing an inverted file as the index for retrieval

In this format the data elements of each record occupy only the space (in bytes) necessary since they do not need to maintain a fixed "width" in all the records, which avoids wasting storage space.

Therefore is necessary to determine the start and end of each record in the database. Technically it is composed of two files: one contains the data, besides the control information for accessing the data elements, and the other contains pointers which indicate in what place (or position) the data file holds the start of each record. The file with the data and access controls is called the cross-reference file and its name has the extension .XRF.

The so-called "inverted file", is implemented through a data structure called balanced tree (or B-tree) , which is characterized by being extremely efficient in search operations for a random element in its contents. This makes CDS/ISIS highly capable in information retrieval. The technical solution to the inverted file

consists of six physical files, all with the same root name and different extensions: .CNT (control); N01 (nodes of short keys); .L01 (leaves of short keys); .N02 (nodes of long keys); .L02 (leaves of long keys) and .IFP (postings of the inverted file).

The concept of short and long keys, not mentioned until now, is intended to raise further the efficiency and the effectiveness of the database in the process of information retrieval. The pattern of the short keys is of up to 10 characters, while the long keys have from 11 to 30 characters.

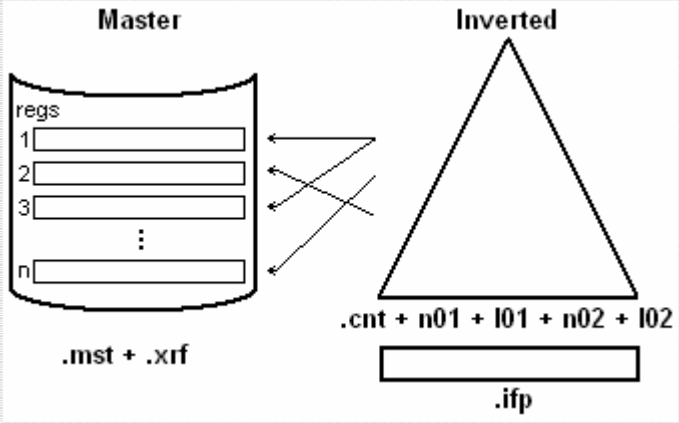


Figure-2: Composition of a CDS/ISIS database as files

From the functional point of view, there are two relevant elements: the master file, which contains all the information of the database; and the inverted file, which contains the list of terms that can be searched, with their locations in the master file.

So a database is made up that can be accessed in an instant to retrieve specific information.

In its applications, BIREME uses special proprietary versions of CISIS, with short keys of up to 16 characters and long keys of 17 to 60 characters, since in the area of health sciences it is common to use long terms.

3.1Maintenance of Information in the database

To be converted into a tool for effective manipulation of information or carrying of knowledge, a database should permit the addition, updating and correction of its

content. CDS/ISIS databases have mechanisms for these basic operations, as described later.

New records are added to the database, as illustrated in Figure-3, which are incorporated in the data file (MST) and receive a reference in the file of pointers to records (XRF).

The files for retrieval from the database receive a marker of "not valid" or are logically deleted, unless that it is effected physically (elimination of data from the database file) in an operation of maintenance of the database, as we will see later.

As regards the occupation of storage space, when a record is edited its size can increase due to the alteration. Following the model of addition of records, the original data remain in position in the electronic file, but are marked as not valid. With that, for each record edited the space ocupied by the database grows, and this is the price of flexibility of the records having fields of variable length.

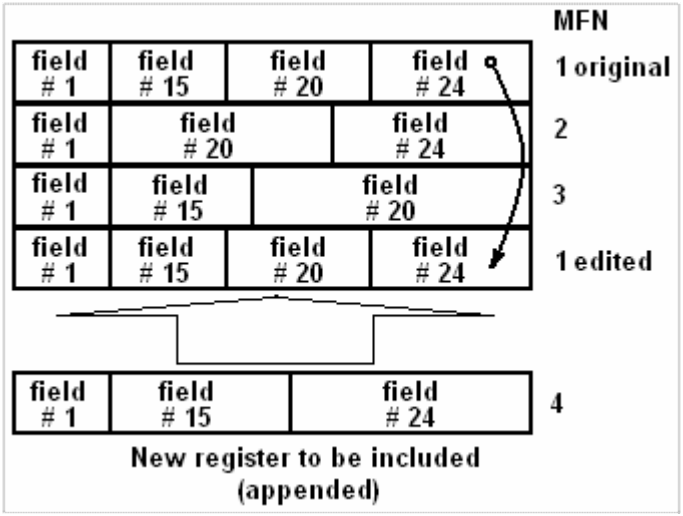


Figure-3: Database during operation

To avoid this undesirable effect, the maintenance operations of the database should include its regular reconstruction, in order to reorganize its contents. Such reconstruction can be carried out in more than one way. The most immediate of all is to read the master file of the database record-by-record, in the sequence of MFNs, and to create a new master file, in which the records are entered in direct order of their numbering. Afterwards the read master file is deleted and the master file created is renamed as it replaces the old one. Another way to effect this

reconstruction, much more common with the users of CDS/ISIS for Windows, is to export the database to a file in ISO 2709 format, and then import over the original database.

In this way, the database represented in Figure-3 passes through the rearrangement of data, as illustrated in Figure-4:

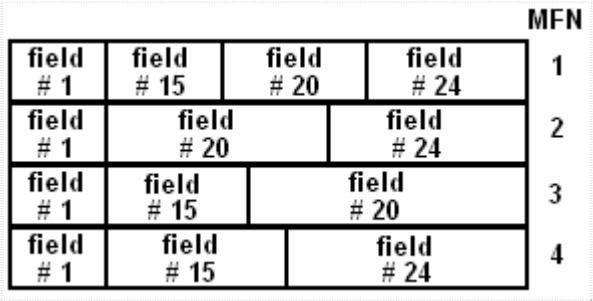


Figure-4: Reorganized database

In general, after the reconstruction you should carry out a complete reindexing of the database, to ensure that, finally, the deleted records are not referred to in old index entries in the inverted file.

3.2Character Strings Used

In the operating environment, a distinct string of codes is used to represent printable characters. For example, when you use a file in MS-DOS (Microsoft Disk Operating System) you use the ASCII codes, and when you edit a file in Microsoft Windows you use the ANSI codes for characters.

For instance, the code 162 (0xA2 in hexadecimal) in ASCII represents the character lower-case “o” with an acute accent (ó), but the same value in ANSI represents the symbol for cents (¢).

In the same way, the localization of the computer system (i.e. nationalization or regional setting), can cause variations in the character codes. In ASCII, the standard is to use what is called CODE PAGE 437, supplying the necessary printable characters for the US market. On the other hand, you can use Code Page 850, which supplies the necessary printable characters for the locations that use Latin languages, especially Spanish and Portuguese.

For example, the value 199 (0xC7 in hexadecimal) of ASCII CP 437 is a graphic character for designing borders, but in ASCII CP 850 it is the character lower-case a with a tilde (ã).

For this reason, a method of converting printable character codes between in character strings is supplied. In the case of CDS/ISIS, this method can be based on the example of special databases called GIZMO. A gizmo database contains two data fields. The input data are matched with the first field and the output is the equivalent data contained in the second field of the gizmo record. The command that makes this conversion feasible will be presented later.

There are four publicly-available gizmo databases: (a) for reading characters in ANSI and outputting characters in ASCII CP 437 (gans437), (b) to read ANSI characters and output characters in ASCII CP 850 (gans650) , (c) to read characters in ASCII CP 437 and output ANSI (g437ans), and finally (d) to read ASCII CP 850 and output ANSI (g850ans).

You need to take account that a Web-based system of data entry uses the ISO-8859-1 character set, which has identical codes and printable symbols to the ANSI set; therefore, if you transfer a database managed in a Web system to a DOS system, you need to convert its characters to the proper set.

3.3Differences between Platforms

Changes in the operating environment can arise from a change of platform, for example, from Windows to Linux, or Unix, etc. Since there are specific characteristics in the recording of files in each of these platforms, the change of platform can produce errors in reading the master and inverted files of CDS/ISIS databases.

To avoid this type of difficulty, the CISIS package provides two applications that perform the necessary alterations in the files for – currently – nine platforms, viz. Linux; HP-UX; Sun; Alpha; Vax; Unisys; MPE; CDC and Windows. One of the applications serves to convert the master file and the other to convert the inverted file.

In this way, before transferring a database between two different systems, you need to verify the need to change the character set and the recording format of the database files (and inverted file, if necessary). The process consists of three phases: conversion of the character set; conversion of the file format; and transfer between the systems, as shown in Figure-5:

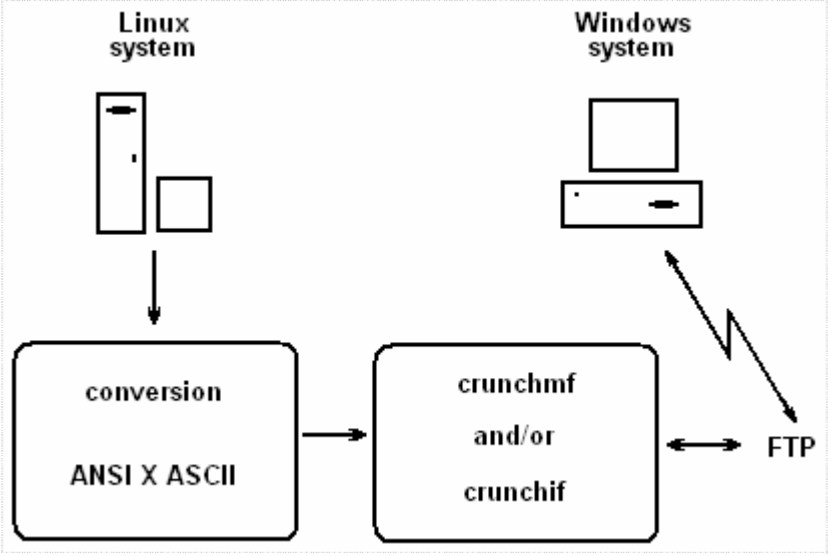


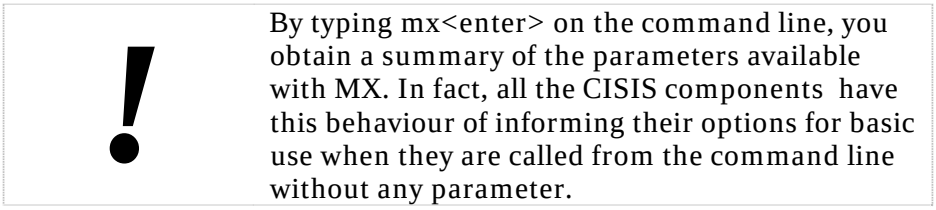
Figure-5: Transfer of a database between systems

4 Introduction to the Use of CISIS

The current version of CISIS is 4.3 (dated 2004). It incorporates various facilities, characteristics and capabilities and it comprises various 'commands', e.g. mx, crunchmf, crunchif, msrt, mxtb, mxcp, mkxrf, and others.

4.1MX - the CISIS Utility

The command of general use in CISIS for reading, writing, retrieval and inversion of databases is MX.



By typing mx<enter> on the command line, you obtain a summary of the parameters available with MX. In fact, all the CISIS components have this behaviour of informing their options for basic use when they are called from the command line without any parameter.

It is sufficient to give the name of the database and mx will proceed to read record-by-record, beginning with the first. Thus, the command **mx lil** returns:

```
mx lil

mfn= 1 [DELETED]
```

```
..

mfn= 2 [DELETED]

..

mfn= 3 [DELETED]

..

mfn= 4 [DELETED]

..

mfn= 5 [DELETED]

..

mfn= 6

1 «BR1.1»

2 «000006»

3 «BR1.1/7.02»

4 «LILACS»

5 «MC»

6 «am»

10 «Banta, David»

12 «The uses of modern technologies: problems and perspectives for
industrialized and developing countries»

14 «29-46»

17 «Conselho Nacional de Desenvolvimento Científico e Tecnológico,
ed»

17 «Organización Panamericana de la Salud, ed»
```

```
18 «Conferência Interamericana sobre a Avaliaçäo Tecnológica em
Saúde»

20 «234»

40 «En»

52 «Brasil. Ministério da Saúde»

52 «Brasil. Ministério da Educaçäo»

52 «Brasil. Ministério da Previdência e Assistência Social»

52 «Instituto Nacional de Assistência Médica e Previdência Social»

52 «Instituto de Pesquisas Econômicas e Sociais»

52 «Organización Panamericana de la Salud»

52 «Conselho Nacional de Desenvolvimento Científico e Tecnológico»

53 «Conferência Interamericana sobre a Avaliaçäo Tecnológica em
Saúde»

54 «14-18 nov. 1983»

55 «19831115»

56 «Brasília»

57 «BR»

62 «Conselho Nacional de Desenvolvimento Científico e Tecnológico»

64 «1985»

65 «19850000»
```

It is necessary to point out that, for each record displayed, the command halts and waits for the operator to request a new record by pressing <*ENTER*>. This condition of waiting (or prompting) is indicated by the presentation of two full stops (..) on the console.

If you wish to verify a certain record, with a known number, the command accepts the following syntax: *mx <base> from=<rec_num>*, for example:

```
D:\Documentos\teste>mx lil from=10000

mfn= 10000

1 «BR1.1»

2 «010000»

4 «IMLA»

4 «LILACS»

5 «S»

6 «as»

10 «Manterola, A»

12 «Coma y alteraciones de conciencia en el nino.»

13 «Coma and consciousness changes in children»

14 «45-58»

30 «Pediatria (Santiago de Chile)»

31 «25»

32 «1/2»

40 «Es»

64 «1982»

65 «19820000»

76 «INFANTE»

76 «NINO»

76 «HUMAN»
```

```
87 «^dCOMA»

87 «^dCONSCIOUSNESS DISORDERS»

90 «b»

..
```

If you do not want to stop after each record, you can include the parameter *now* in the mx command, which causes the sequential reading of all the records of the database to a certain point, if specified, or to the end.

It is possible to specify and combine various parameters in the use of MX, which offers extreme flexibility. To multiply this flexibility you can add the facilities of the native commands of the operating system, for example, counting of lines, sorting with the exclusion of repeated elements, etc.

It is important to underline that, to carry out the change of patterns, the gizmo clause is available, and with it you can use gizmo databases to change the character set of a database, e.g.

```
mx lilasc gizmo=g850ans create=lilans –all now
```

which will take a Lilacs database with ASCII characters and create another with ANSI characters.

4.2 Other commands of CISIS

4.2.1 crunchmf and crunchif

The commands for conversion of the format of a file crunchmf and crunchif are used when you wish to put a database and/or inverted file on another platform than its original, for example, to move a database from Windows to HP-UX.

The general form of use of the two commands is similar and requires that you indicate to the command the file to read (whether master or inverted), the name of

the file to create in the new recording format (whether master or inverted) and the system of destination for the file, among other less frequently used possibilities.

You can see below an example of its use, where a DeCS database is taken, presently on a Linux plataform and prepared for transfer to a Windows system.

```
crunchmf decs win/decs target=pc tell=5000
```

4.2.2 msrt

A rather useful command for making reports is msrt, which organizes the records in a master file, according to a criterion which is based on the content of the fields of the particular database. Thus it is possible to sort the database by date of publication (assuming that this field exists in the database), or by a more complex criterion, such as the alphabetic order of the authors and titles together.

An example of its use is shown below, in which the database Lilacs is sorted by the date of publication (field 65) of the document, with a secondary sort on the title of the document.

```
msrt lilacs 160 "s(v65,v18,v12)"
```

4.2.3 mxtb

Another very useful command is mxtb, whose function is to tabulate data from the database to generate a new database (whose name is specified in the command call). The new database is standarized and contains the frequencies of the data, conforming to the extraction format supplied. The universe of data to tabulate can be selected according to a search expression in the database, which allows restricting the data to tabulate.

An example is shown below, in which the years of publication (first four positions of v65) are tabulated from a specified title of a journal in Lilacs.

```
mxtb lilacs create=rbopd "4:v65.4" "bool=Rev. Bras. Oftalmol"
```

In Table-1 below are shown various types of command, for a list of common situations, to serve as a guide for the execution of routine tasks.

Task	Model Command
Read starting from a record	*mx <base> from=<mfn_initial>*
Read a range of records	*mx <base> from=<mfn_initial> to=<mfn_final>*
Determine the number of records	*mx <base> +control count=-0*
Create a copy of the database	*mx <base_in> create=<base_out> -all now*
Create a copy of the database without deleted records	*del <base_out.mst> (or rm <base_out.mst>)* *del <base_out.xrf> (or rm <base_out.xrf>)* *mx <base_in> append=<base_out> -all now*
Generate an ISO file from the database	*mx <base> iso=<file.iso> -all now*
Generate a database from an ISO file	*mx iso=<file.iso> create=<base> -all now*
Generate a database from an ISO file with mfn given by a field	*mx iso=<file.iso> create=<base> "proc='=’vn" –all now*
Invert a database	*mx <base> "fst=<field technique format>" fullinv=<inverted>*
Invert a database according to an FST file	*mx <base> "fst=<@file.fst>" fullinv=<invertido>*
Invert a database which uses ANSI characters	*mx <base> "fst=<@file.fst>" fullinv/ansi=<invertido>*
Convert the character code of the database	*mx <base> gizmo=<base_gizmo> create=<base_out> -all now*
Convert a database from another plataform	*crunchmf <base> <base_out> target=<plataform>*
Convert an inverted file from another plataform	*crunchif <inverted> <inverted_out> target=<plataform>*
Produce a file legible by Excel	*mx <base> “pft=|”|vn|”|,|,”|vn|”|,|,”|vn|”|/” –all now > <file.csv>*
Import a file from Excel (CSV)	*mx “seq=<file.csv,” create=<base_out> -all now*
Export sorted data from a database	*mx <base> “pft=<format>” –all now |sort –u (linux)*
Produce statistics from fields and characters of the database	*mxf0 <base> create=<base_out> 0 noedit*
Sort the database according to a field	*msrt <base> <#chars_of_the_key> <extraction format>*
Tabulate fields from a database	*mxtb <base> create=<base_out> <length:format> [class=nnnnn]*

Table-1: Common commands in CISIS

5 Practical Aspects

5.1 Equivalence between Linux and Windows commands

A large part of the functions performed in a one operating system match equivalent ones in another. Below, in Table-2, is a list of equivalences between Linux, DOS and Windows.

LINUX	DOS	WINDOWS
clear	*cls*	Not applicable
ls	*dir /w*	Table on the right in Windows Explorer
ls –l	*dir*	Table on the right in Windows Explorer
mkdir	*mkdir*	Create a new folder
rmdir	*rmdir*	Delete a folder (empty)
cd	*cd*	Navigate to another folder in Windows Explorer
rm	*del*	Delete a file (or several)
rm –R	*deltree*	Delete a folder (and its contents)
ftp	*ftp*	FTP application (ws-ftp)

Tabla-2: Equivalent commands between Operating Systems

5.2 Common Commands in FTP

In case you need to transfer files by FTP (File Transfer Protocol) and the application does not have a graphical interface, Table-3 gives a list of the most common commands of FTP.

Comando	Explicación
open	Establishes a connection with a remote system
ls	Presents the contents of the current directory
cd	Changes the working directory
bin	Activates binary mode for transfer
asc	Activates ASCII mode for transfer
put	Sends a file to the remote system
mput	Sends several files to the remote system
get	Receives a file from the remote system
mget	Receives several files from the remote system
prompt	Activates/deactivates the interactive confirmation dialogue

Table-3: Common commands in FTP

5.3 Conversion of Character Strings

To convert the character codes in a database you use the command **mx**, which is one of the components (the most powerful) of CISIS, in conjunction with a gizmo database, as already mentioned.

Supposing that the objective is to change the character codes from ASCII Code Page 850 to ANSI, the following command will perform the operation:

```
mx BASE_IN gizmo=g850ans create=BASE_OUT now –all
```

The original database used in the example is called BASE_IN, but it could be any other name. The gizmo database, in this example, is in the working directory and a second database called BASE_OUT is created (which also could have any other name) with the character codes converted.

In the example below, the database LILACS, which uses the ANSI character codes, is converted by means of the gizmo ANSI ➔ ASCII CP850 in the directory **tabs** (below the working directory) thus creating the database LIL850.

```
mx lilacs gizmo=../tabs/gans850 create=lil850 now –all
```

5.4 Transfering databases between Operating Systems

If you want to transfer a database (just the master file) from one platform to another, with different operating systems, there is as an example the command **crunchmf**, which is one of the components of CISIS.

Supposing that the database is on a Linux server and is to be transferred to a Windows machine, the following command performs the operation:

```
crunchmf BASE_ORI BASE_DEST target=pc
```

The original database used in the example is called BASE_ORI, but you could use any name. It cretaes a destination database called BASE_DEST (again this could be any name) and the conversion of physical format is specified by the parameter **target=** which in this case is **pc**, i.e. a Windows system.

In the example below, the database GANSNA, presently on a PC which runs Windows, is prepared for transfer to a Linux machine, with the result placed in the directory LNX.

```
crunchmf gansna lnx\gansna target=linux
```

A similar situation applies to the inverted file, but the command **crunchif** is used, another component of CISIS.

6 Referencias Bibliográficas

1. UNESCO. *Mini-micro CDS/ISIS:* Reference manual (version 2.3). Organized by Giampaolo Del Bigio. Paris: United Nations Educational, Scientific and Cultural Organization, 1989. 286 p. ISBN 92-3-102-605-5.

2. BUXTON, Andrew, HOPKINSON, Alan. *The CDS/ISIS for Windows Handbook* [online]. Paris: United Nations Educational, Scientific and Cultural Organization, 2001 [cited 30 August 2006]. 164 p. Available from internet: <http://bvsmodelo.bvs.br/download/winisis/winisis-handbook-en.pdf>.

3. SUTER, Tito. “Prehistoria” e historia del MicroISIS [online]. In: *Manual para instructores de Winisis*. Buenos Aires: Centro Atómico Constituyentes (CAC), Comisión Nacional de Energía Atómica (CNEA), 1999 [citado el 30 Agosto 2006]. p. 21-26. Disponible en internet: <http://www.cnea.gov.ar/cac/ci/isis/isidams.htm>.

7 Glossary

- **Backup**. Procedure used to duplicate one or more files and/or directories in another storing device (tape or disc), thus producing a backup copy that may be restored in the event of accidental deletion or physical damage to the original data.
- **Bibliographic Database**. Electronic version of a catalog or bibliographic index.
- **CDS/ISIS – MicroISIS**. Software program developed and maintained by UNESCO to manage bibliographic data.
- **Database**. Collection of data that are structured to be easily accessed and handled. It is formed by units called records whose attributes are represented by fields. For example, in a file called "customer base", each customer is a record, with several fields such as "NAME", "CUSTOMER CODE", "TELEPHONE" etc.
- **Electronic Format**. Any form of storage, retrieval or presentation of information that may be transmitted on-line or recorded in magnetic or optical media.

- **Field**. *See* Database.
- **File**. In computing, a set of data that may be saved into some type of storing device. The data files are created by applications, such as a text processor for example.
- **Indexing**. Procedure to identify and describe the content of a document with terms that reflect the corresponding subject matters to allow the document to be retrieved later.
- **ISO Format (of files)**. Standard established by the ISO to allow the exchange of data between institutions, networks and users.
- **Key**. Record element that allows for storing specific information. *See* Database.
- **LILACS Format**. A bibliographic description format established by BIREME, based on the UNISIST Reference Manual for Machine-readable Bibliographic Descriptions.
- **Record**. Structured data set built to accomodate a specific subject. *See* Database.
- **UNISIST**. Intergovernmental program designed to foster cooperation in the field of scientific and technological knowledge.