

BIREME / PAHO / WHO

Latin American and Caribbean Center on Health Sciences Information

**Concetti fondamentali su basi di dati CDS/ISIS:  
introduzione all'uso del CISIS**

Versione 3.x

São Paulo - 2004-2007

Copyright © 2004-2007 - BIREME / PAHO / WHO

Concetti fondamentali su basi di dati CDS/ISIS: introduzione all'uso del CISIS

È garantito il permesso di copiare, distribuire e/o modificare questo documento seguendo i termini della Licenza per Documentazione Libera GNU, Versione 1.2 o ogni versione successiva pubblicata dalla Free Software Foundation; con le Sezioni Non Modificabili ELENCARNE I TITOLI, con i Testi Copertina ELENCO, e con i Testi di Retro Copertina ELENCO. Una copia della licenza è acclusa nella sezione intitolata "GNU Free Documentation License".

### Scheda catalografica

<p>BIREME (Brasile)</p> <p>Concetti fondamentali su basi di dati CDS/ISIS: introduzione all'uso del CISIS. / BIREME (org.). São Paulo : BIREME / PAHO / WHO, 2004-2007.</p> <p>32 p.</p> <p>1. Manuali per l'utente. 2. Basi di dati. 3. Sistemi informativi. 4. Software CDS/ISIS. I. BIREME II. Titolo</p>
--

**Avvertenza** - Che determinate compagnie e/o istituzioni o prodotti vengano menzionati nel testo non implica che essi siano sostenuti o raccomandati da BIREME / PAHO / WHO, e non significa nemmeno che vi sia preferenza loro accordata in rapporto ad altri conismili, citati o meno.

BIREME / PAHO / WHO

Latin American and Caribbean Center on Health Sciences Information

Rua Botucatu 862 V Clementino

*Questo documento è stato prodotto con la Metodologia per la Normalizzazione di Documenti (NorDoc) sviluppata dalla BIREME*

# Indice generale

<b>Abbreviazioni</b> .....	IV
<b>Prefazione</b> .....	1
A proposito di BIREME .....	1
La Biblioteca virtuale della salute (Biblioteca Virtual en Salud -BVS) .....	2
<b>Introduzione</b> .....	5
<b>Basi di dati CDS/ISIS</b> .....	8
Manutenzione delle informazioni nelle basi di dati .....	10
Insiemi di caratteri in uso .....	11
Differenze fra le piattaforme .....	13
<b>Introduzione all'uso del CISIS</b> .....	14
MX il programma di utilità del CISIS .....	14
Altri comandi CISIS .....	17
<i>crunchmf</i> y <i>crunchif</i> .....	17
<i>msrt</i> .....	18
<i>mxtb</i> .....	18
<b>Aspetti pratici</b> .....	20
Equivalenza fra comandi Linux e Windows .....	20
Comandi FTP più comuni .....	20
Conversione di un insieme di caratteri .....	21
Trasferire basi dati sotto un sistema operativo diverso .....	22
<b>Riferimenti bibliografici</b> .....	23
<b>Glossario</b> .....	24

# Abbreviazioni

- **ANSI.** American National Standards Institute [Istituto nazionale americano di normalizzazione].
- **ASCII.** American Standard Code for Information Interchange [Codice americano normalizzato per lo scambio dell'informazione].
- **BIREME.** Centro Latinoamericano y del Caribe de Informacion en Ciencias de la Salud [centro latinoamericano e caraibico per le scienze della salute].
- **BVS.** Biblioteca Virtual en Salud [Biblioteca virtuale della salute].
- **CDS.** Computerized Documentation System [Sistema di documentazione informatizzata].
- **CP.** Code Page [Pagina di codifica caratteri].
- **FST.** Field Selection Table [Tabella di selezione dei campi].
- **FTP.** File Transfer Protocol [Protocollo per il trasferimento di files].
- **IFP.** Inverted File Pointer [Puntatore dell'indice trasposto].

- **ISIS.** Integrated Set of Information Systems [Insieme integrato di sistemi informativi].
- **ISO.** International Organization for Standardization [Organizzazione internazionale per la normalizzazione].
- **LILACS.** Literatura Latinoamericana y del Caribe en Ciencias de la Salud [Letteratura latinoamericana e caraibica in scienze della salute].
- **OMS.** Organización Mundial de la Salud [Organizzazione mondiale della sanità].
- **OPS.** Organización Panamericana de la Salud [Organizzazione panamericana della sanità].
- **UNESCO.** United Nations Educational, Scientific and Cultural Organization [Organizzazione delle Nazioni unite per l'educazione la scienza e la cultura].

# Prefazione

## A proposito di BIREME

Anno dopo anno, BIREME assolve la sua funzione di centro specializzato nell'informazione scientifica e tecnica per la salute nell'area dell'America latina e dei Caraibi. Fondata in Brasile nel 1967, col nome di Biblioteca Regionale di Medicina (Biblioteca Regional de Medicina: da cui la sigla BIREME), sin dall'inizio prestò attenzione alla crescente domanda di letteratura scientifica corrente proveniente dai sistemi sanitari nazionali e dalle comunità di ricercatori, professionisti e studenti. In seguito, nel 1982, prese il nome di Centro Latinoamericano e dei Caraibi per l'informazione in scienze della salute (Centro Latinoamericano y del Caribe de Información en Ciencias de la Salud), allo scopo di esprimere meglio le proprie funzioni, orientate all'irrobustimento e all'ampliamento del flusso dell'informazione scientifica e tecnica relativa alla salute in tutta l'area, tuttavia mantenne la stessa sigla.

L'attività in rete, fondata sul decentramento, volta a sviluppare le capacità locali, condividere le risorse informative, sviluppare prodotti e servizi cooperativi, elaborare metodologie comuni, fu sempre la base del lavoro cooperativo di BIREME. In tale maniera il centro si consolida come un modello internazionale che promuove le capacità dei professionisti dell'informazione a livello gestionale e

tecnico, affinché essi adottino protocolli di informazione e di comunicazione che meglio servano le esigenze locali.

I principi fondamentali che sono alla base e sostengono l'esistenza di BIREME sono:

- l'accesso all'informazione sanitaria scientifica e tecnica è essenziale per lo sviluppo della sanità;
- la necessità di sviluppare le capacità dei paesi dell'America latina e dei Caraibi a sfruttare le fonti di informazione sanitaria scientifica e tecnica in maniera cooperativa ed efficiente;
- la necessità di promuovere l'uso e di rispondere alla richiesta di informazione sanitaria scientifica e tecnica da parte dei governi, dei sistemi sanitari, delle istituzioni educative e di ricerca.

BIREME, in qualità di centro specializzato della Organizzazione panamericana della sanità (Organización Panamericana de la Salud - OPAS)/Organizzazione mondiale della sanità (OMS), coordina e realizza attività di cooperazione tecnica in gestione dell'informazione e della conoscenza scientifiche, coll'intento di rafforzare ed ampliare la circolazione dell'informazione sanitaria scientifica e tecnica in Brasile e negli altri paesi dell'America latina e dei Caraibi, come condizione essenziale per lo sviluppo della sanità, includendovi programmazione, gestione, promozione, ricerca, educazione e servizio.

L'accordo che è alla base di BIREME viene rinnovato ogni cinque anni dai membri del Comité Asesor Nacional de la institución (OPAS, Ministerio de la Salud de Brasil, Ministerio de Educación y Cultura de Brasil, Secretaría de Salud del Estado de São Paulo y Universidad Federal de São Paulo – Unifesp). Quest'ultima mette a disposizione l'infrastruttura fisica necessaria al funzionamento dell'istituzione.

Nel 2004 l'istituzione prese la decisione di mutarsi in un'entità fondata sulla conoscenza.

## La Biblioteca virtuale della salute (Biblioteca Virtual en Salud -BVS)

Con la nascita e il consolidamento di Internet come mezzo prevalente di informazione e di comunicazione, il modello de cooperazione tecnica di BIREME

prese ad evolvere dal 1998 verso la costruzione e lo sviluppo della Biblioteca virtuale della salute (Biblioteca Virtual en Salud BVS) quale spazio comune di confluenza del lavoro cooperativo dei produttori, degli intermediari e degli utenti di informazione. La BVS promuove lo sviluppo di una rete di fonti di informazione scientifica e tecnica con accesso universale in Internet. Per la prima volta si apre la possibilità di un accesso equo all'informazione sanitaria.

BIREME cura la Biblioteca Virtual come modello per la gestione dell'informazione e della conoscenza, il che implica la cooperazione e la convergenza di istituzioni, sistemi, reti e iniziative di produttori, intermediari ed utenti per la messa in opera di reti di fonti di informazione locali, nazionali, regionali e internazionali, privilegiando in tal maniera l'accesso aperto ed universale.

Attualmente, tutti i paesi dell'America latina e della regione dei Caraibi partecipano direttamente o indirettamente nella fornitura dei prodotti e dei servizi cooperativi promossi dalla BVS, il che significa oltre mille istituzioni in oltre 30 paesi.

La BVS si situa in uno spazio virtuale di Internet formato dalla raccolta o rete di fonti di informazione sanitaria di tutta l'area. Utenti di livelli e provenienza diversi possono interagire e navigare nello spazio di una o più fonti d'informazione, a prescindere dall'ubicazione fisica. Le fonti informative vengono generate, aggiornate, conservate e trattate in Internet da produttori, integratori e intermediari, in regime di decentramento, impiegando metodologie condivise per la loro integrazione nella BVS.

La BVS organizza l'informazione in una struttura che integra e interconnette basi di dati referenziali, elenchi di specialisti, eventi ed enti, catalogo di risorse informative disponibili in Internet, collezioni di testi integrali con particolare risalto per la raccolta SciELO (Scientific Electronic Online) di riviste scientifiche, servizi di disseminazione selettiva dell'informazione, fonti di informazione di supporto all'educazione ed alle attività decisionali, notizie, liste di discussione e sostegno alle comunità virtuali. Conseguentemente, lo spazio della BVS costituisce una rete dinamica e decentrata di fonti informative a partire dalla quale si possono recuperare ed estrarre informazione e conoscenza per sostenere i processi decisionali nell'ambito della salute.

**La BVS Biblioteca Virtual en Salud è vista come la base dati distribuita della conoscenza scientifica e tecnica in materia di salute memorizzata, organizzata conservata in formato elettronico nei paesi dell'area, accessibile universalmente in Internet in maniera compatibile con le altre basi dati internazionali.**

# Introduzione

Se si dispone di una mole rilevante di informazioni merita conservarla affinché sia possibile consultarla anche in futuro.

Questa raccolta di informazioni ha buone probabilità di costituire una porzione di una banca dati.

Nell'accezione più consolidata, una banca dati è vista come un insieme di basi di dati di norma correlate le une con le altre.

L'elemento basilare di una base di dati è il record che va visto come l'unità informativa autonoma più piccola. A sua volta il record è composto di dati, come elementi residenti in campi, che costituiscono unità informative ancora più piccole che però potrebbero non avere alcun significato se considerate in maniera isolata. E inoltre il contenuto di un campo può essere articolato in porzioni minori alloggiate in sottocampi. La Figura-1 illustra questo assetto strutturale.

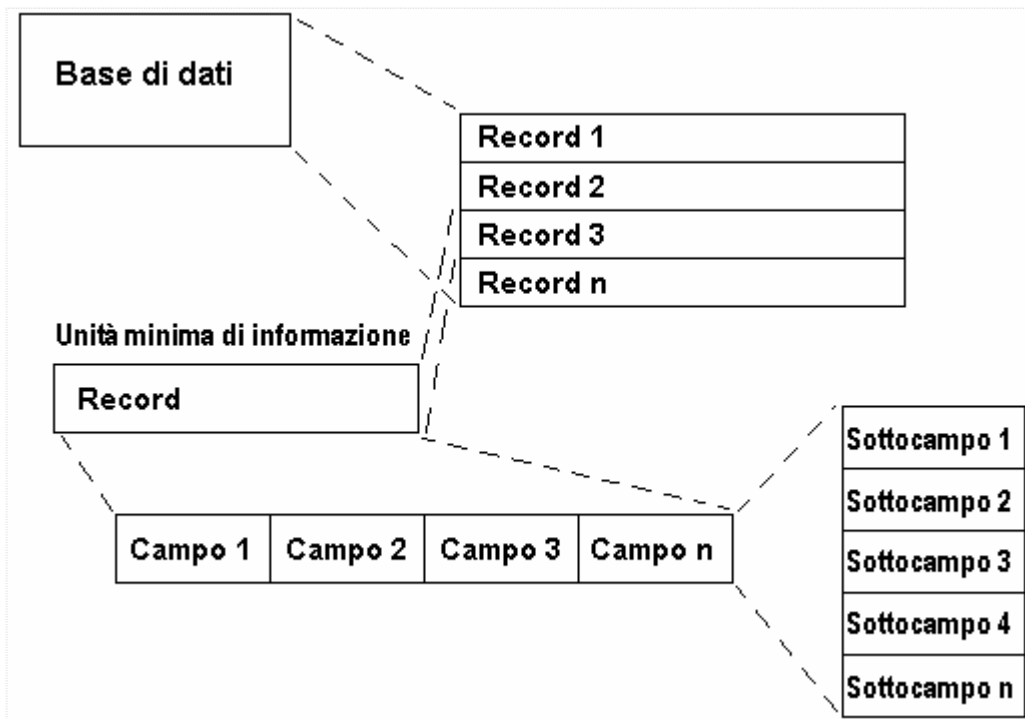


Figura-1: Struttura di una base di dati

Poniamo che ogni record della Figura-1 rappresenti un documento catalogato, e che ogni campo corrisponda ad un elemento di dati, come l'autore, il titolo, gli argomenti. Alcuni campi possono essere composti di sottocampi, come il campo autore, verosimilmente articolato, fra i vari sottocampi possibili, in: nome; cognome; ente a cui l'autore è affiliato; centro di ricerca presso cui egli lavora.

Esistono vari sistemi di gestione di banche dati; al giorno d'oggi il sistema relazionale è riconosciuto come il più alto livello di sviluppo raggiunto nella gestione di banche dati. E' basato su un insieme di tabelle fra le quali vengono stabilite delle relazioni logiche.

Tuttavia, tenendo conto dei diversi tipi di informazione da trattare, il modello relazionale può rivelarsi non adeguato al compito, come è il caso con informazioni che non ammettono quella normalizzazione esigita invece dal modello relazionale..

Per il tipo di informazione e di dati rappresentati nei riferimenti bibliografici (sostanzialmente informazione testuale) è preferibile adottare un modello di banca dati più flessibile come uno che sappia trattare, ad esempio, record a lunghezza

**variabile nonché campi che ricorrono più volte (ripetibili) all'interno del medesimo record.**

# Basi di dati CDS/ISIS

Una base dati CDS/ISIS è caratterizzata da:

- riconoscere campi a dimensioni variabili;
- riconoscere campi dati ripetibili;
- consentire l'uso di sottocampi;
- consentire relazioni fra diverse basi di dati;
- utilizzare un indice trasposto (*inverted file*) come mezzo di recupero dei dati.

In tal modo i dati che costituiscono ciascun record occupano solo lo spazio (misurato in byte) necessario per la loro esatta dimensione, senza pretendere una “lunghezza” fissa per tutti i record, il che evita lo spreco di spazio di memoria di massa.

Poiché in questo modo non è facile indicare l'inizio di un qualsiasi record all'interno della base dati, questa viene strutturata dal punto di vista tecnico in due file: uno che contiene i dati, oltre alle informazioni di controllo per l'accesso ai dati stessi, e un altro che contiene i puntatori che registrano in quale posto –posizione fisica- del file dei dati inizia ogni singolo record. Il file che ospita i dati e i puntatori di controllo per l'accesso è identificato dall'estensione .XRF (cross-reference).

Quello chiamato “*inverted file*” (indice trasposto), o anche solo ‘indice’, è reso operativo attraverso una struttura dei dati definita ad “albero bilanciato”, la cui caratteristica è quella di essere estremamente efficiente nelle operazioni di ricerca

di un contenuto a caso, il che spiega le eccellenti prestazioni di CDC/ISIS in termini di recupero delle informazioni. La soluzione tecnica per realizzare un indice del genere passa per la gestione di sei file, tutti con lo stesso nome e con estensioni diverse: .CNT (controllo); N01 (nodi –*nodes*- delle chiavi brevi); .L01 (foglie – *leaves*- delle chiavi brevi); .N02 (nodi delle chiavi lunghe); .L02 (foglie delle chiavi lunghe) y .IFP (puntatori dell'*inverted file*).

Il concetto di chiavi brevi e lunghe, finora non menzionato, ha come obiettivo quello di aumentare ancor più l'efficienza e l'efficacia della base dati per quanto attiene alle prestazioni di recupero dell'informazione. Le chiavi brevi vanno da 0 a 10 caratteri, mentre le lunghe da 11 a 30.

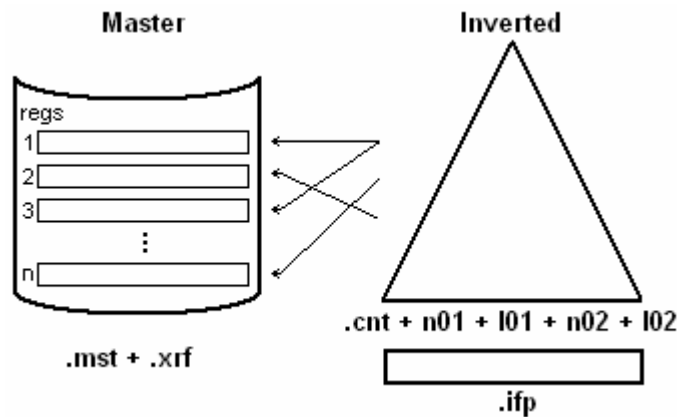


Figura-2: Articolazione in file di una base dati CDS/ISIS

Dal punto di vista funzionale sono presenti due componenti prevalenti: il *master file*, che contiene tutte le informazioni della banca dati e l'*inverted file* che contiene la lista dei termini che possono venire ricercati, assieme con le loro coordinate di localizzazione all'interno del *master file*.

In questo modo si allestisce una base dati alla quale si può avere accesso istantaneo al fine di rintracciare specifici contenuti.



Visto che nel settore delle scienze della salute è comune che vengano adoperati termini più lunghi della media, per le proprie applicazioni BIREME utilizza versioni speciali e proprie del CISIS, con chiavi brevi fino a 16 caratteri, così come chiavi lunghe da 17 a 60 caratteri.

# Manutenzione delle informazioni nelle basi di dati

Per potersi trasformare in uno strumento efficace di manipolazione dell'informazione o di trasferimento della conoscenza, una base dati deve consentire l'aggiunta, eliminazione e correzione del proprio contenuto. Le basi dati CDS/ISIS dispongono di procedure in grado di svolgere queste operazioni fondamentali, così come qui di seguito illustrato.

Possono venire aggiunti nuovi record, come mostrato nella Figura-3, che vanno a fare parte costitutiva del file dei dati della base (MST) e vengono muniti di un riferimento all'interno del file dei puntatori ai record (XRF).

I record da eliminare dalla base dati vengono marcati come non validi e vengono subito cancellati logicamente, il che si trasforma in una cancellazione effettiva (eliminazione fisica dei dati dal file della base dati) durante un'operazione di manutenzione della base dati, come si vedrà in seguito.

In termini di occupazione di spazio di memoria, la registrazione di un record corretto si comporta secondo lo schema dell'aggiunta di record, dato che la sua dimensione può aumentare in ragione della modifica; i dati originali restano nella posizione che avevano dentro al file, però vengono contrassegnati come non validi. In tal modo, in ogni record editato lo spazio occupato dalla base dati aumenta e questo va visto come il costo della flessibilità dei record dovuta alla loro dimensione variabile.

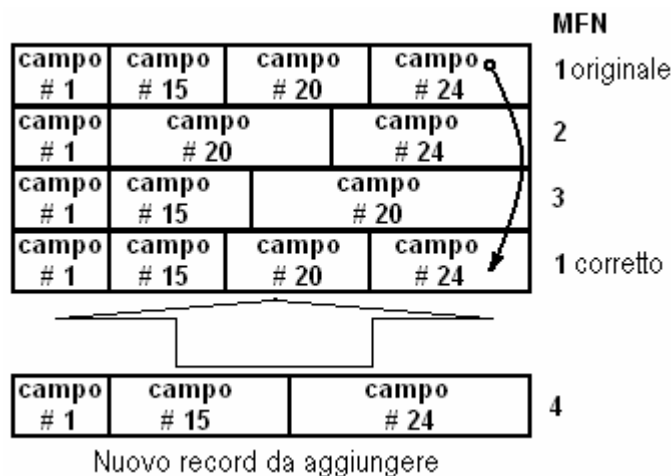


Figura-3: Base dati durante il suo funzionamento

Per porre rimedio agli effetti indesiderati di questo funzionamento, le operazioni di manutenzione della base dati devono comprendere la ricostruzione regolare della base al fine di riorganizzarne il contenuto. Si può portare a termine questa ricostruzione in più di una maniera. La più diretta di tutte è quella che consiste nel leggere il *master file* della base dati record per record, procedendo in sequenza di MFN e creare un nuovo *master file* all'interno del quale i record siano registrati in base al loro stesso numero. Successivamente si elimina il *master file* che si è letto e si rinomina il *master file* ora creato per sostituirlo al precedente. Altra maniera di effettuare la ricostruzione, molto più diffusa presso gli utenti di CDS/ISIS per Windows, consiste nell'esportare i dati come file ISO, e poi reimportarlo nella base originale, così rimpiazzandola.

In questo modo la base dati raffigurata nella Figura-3 subisce una redistribuzione dei dati come illustrato nella Figura-4:

				MFN
campo # 1	campo # 15	campo # 20	campo # 24	1
campo # 1	campo # 20		campo # 24	2
campo # 1	campo # 15	campo # 20		3
campo # 1	campo # 15		campo # 24	4

Figura-4: Base di dati dopo la riorganizzazione

In genere dopo la riorganizzazione si compie anche una reindicizzazione completa della base dati per fare sì che i record cancellati, qualora ce ne fossero, non vengano più indicizzati da voci perente all'interno *dell'inverted file*.

## Insiemi di caratteri in uso

A seconda del sistema operativo nel cui ambito si agisce si usa un particolare insieme *-set-* di codici per rappresentare i caratteri ai fini della stampa. Ad esempio quando si usa un file sotto MS-DOS (Microsoft Disk Operating System), si

utilizza l'insieme ASCII, e quando si tratta un file sotto Windows si utilizza l'insieme ANSI.

Ad esempio, il codice 162 (0xA2 in esadecimale) dell'insieme ASCII, rappresenta una 'o' minuscola con accento acuto (ó), mentre nell'insieme ANSI rappresenta il simbolo dei centesimi come una c minuscola tagliata in verticale (¢).

Contemporaneamente però, in rapporto alla localizzazione linguistico-territoriale del sistema operativo possono vigere variazioni nello stesso insieme dei caratteri. In ASCII vige il modello di declinare l'insieme in varie pagine di codici CODE PAGE numerate, dove la 437, ad esempio, ben si adatta all'ambito statunitense (o anglofono) mentre la Code Page 850 serve meglio le esigenze di lingue latine come spagnolo e portoghese con lettere con diacritici ed accenti.

Ad esempio, il codice 199 (0xC7 in esadecimale) dell'ASCII CP 437 è un carattere grafico che serve a disegnare bordi, mentre nell'ASCII CP 850 è una 'a' minuscola con la tilde.

Dato questo quadro occorre disporre di un metodo di conversione dei codici dei caratteri stampabili fra i vari insiemi, nel caso di CDS/ISIS questo metodo può essere fondato sull'uso di basi dati speciali denominate GIZMO. Una base gizmo è tale che al minimo contempla due campi per i dati in ogni record e pone il dato del primo campo in rapporto coi dati in ingresso (input) mentre il secondo campo contiene il dato di sostituzione che verrà usato per l'uscita (output). Il comando grazie al quale si può realizzare questa conversione verrà illustrato più oltre.

Ci sono quattro basi gizmo di dominio pubblico per leggere caratteri in ANSI e produrre caratteri in ASCII CP 437 (gans437), leggere caratteri ANSI e produrre caratteri ASCII CP 850 (gans650), leggere caratteri ASCII CP 437 e produrre ANSI (g437ans) ed infine leggere ASCII CP 850 e produrre ANSI (g850ans).

Si noti che un sistema di scrittura in ambito Web usa l'insieme di caratteri ISO-8859-1 (Latin-1), con codici e segni stampabili identici a quelli dell'ANSI; conseguentemente una base dati creata in ambiente Web per venire trasportata in ambiente DOS deve venire convertita quanto ad insieme di caratteri usati, e viceversa.

## Differenze fra le piattaforme

Cambiamenti di ambiente operativo possono accompagnarsi a cambio di piattaforma ad esempio da Windows a Linux, o Unix, etc. Visto che ciascuna di queste piattaforme ha le sue tecniche specifiche di scrittura dei file, il cambio di piattaforma può provocare errori nella lettura dei file *master* e *inverted* delle basi dati CDS/ISIS.

Per ovviare al problema il pacchetto di utilità CISIS offre applicativi che realizzano le modifiche necessarie nei file al momento per 9 piattaforme, e cioè: Linux; HP-UX; Sun; Alpha; Vax; Unisys; MPE; CDC e Windows. Un applicativo serve alla conversione del *master file* e l'altra a quella dell'*inverted file*.

Dunque prima di trasferire una base dati da un sistema ad uno diverso si dovrà verificare la necessità o meno di cambiare l'insieme dei caratteri usato e il formato di scrittura dei file (*master* e se necessario *inverted*), e pertanto si disegnerà un processo articolato in tre fasi: conversione dell'insieme dei caratteri; conversione del formato dei file; trasferimento effettivo fra sistemi, si veda a proposito la Figura-5:

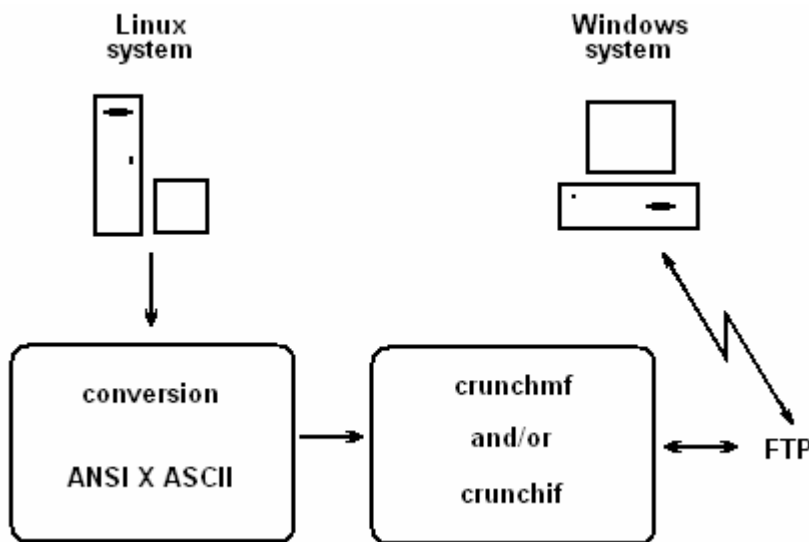


Figura-5: Trasferimento di basi dati fra sistemi diversi

# Introduzione all'uso del CISIS

La versione attuale del CISIS è la 4.3 (del 2004), essa comprende varie funzioni, caratteristiche e potenzialità espresse con diversi comandi come: mx, crunchmf, crunchif, msrt, mxtb, mxcp, mkxrf y otros.

## MX il programma di utilità del CISIS

Il comando di uso più comune del CISIS per realizzare lettura, scrittura, recupero e indicizzazione di basi di dati è MX.



Digitando `mx<invio>` dalla linea di comando, si ottiene in risposta a video un sommario dei parametri di esecuzione di MX. Di norma tutte le componenti del CISIS reagiscono in questa maniera mostrando le proprie opzioni fondamentali d'uso quando vengono invocate dalla linea di comando senza parametri aggiuntivi.

Basta indicare il nome della base dati e mx procederà alla sua lettura record dopo record cominciando dal primo. Perciò un comando come **mx lil** produce:

```
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 Avaliação Tecnológica em Saúde»

20 «234»

40 «En»

52 «Brasil. Ministério da Saúde»

52 «Brasil. Ministério da Educação»

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 Avaliação 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»
```

Si osservi che ad ogni record visualizzato il comando si ferma e aspetta che l'operatore passi alla visualizzazione del record successivo premendo il tasto `<ENTER>` (invio). Questo stato di attesa (prompt) è rappresentato da due punti affiancati sulla linea di comando.

Se si vuole controllare un determinato record, di cui si conosce il numero, il comando accetta la seguente sintassi: `mx <base> from=<reg_num>`, ad esempio:

```
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»
..
```

Se non si vuole che ad ogni record analizzato il sistema si metta in attesa fermandosi, si può includere nel comando mx il parametro *now*, ottenendo così di passare ad una lettura sequenziale di tutti i record della base a partire da un certo record e fino ad un altro, se precisato, oppure fino alla fine.

Si possono combinare vari parametri nell'uso di MX, il che apporta una notevole flessibilità. Questa flessibilità può venire potenziata grazie a comandi che sommino ai comandi MX –evitando le ripetizioni- quelli del sistema operativo atti ad es. a contare linee, ordinare....

E' importante sottolineare che per effettuare il cambiamento dei caratteri è disponibile l'opzione *gizmo*, grazie alla quale vengono usate le basi *gizmo* appunto per convertire insiemi di caratteri, ad esempio:

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

così si prende in input una base Lilacs con caratteri ASCII e se ne crea in output un'altra –di diverso nome- con caratteri ANSI.

## Altri comandi CISIS

### crunchmf y crunchif

I comandi di conversione di formato file *crunchmf* e *crunchif* vengono usati quando si desidera trasportare una base dati e/o il suo indice in un'altra piattaforma, diversa da quella d'origine, ad esempio passare da una base Windows ad una HP-UX.

La sintassi generale per l'uso dei due comandi è quasi la stessa e reclama che nel comando si indichi il file da leggere in entrata (che sia il *master* o l'*inverted*), il nome del file da creare nel nuovo formato di memorizzazione (sia *master* o *inverted*) ed il sistema di arrivo per il file, a parte altre opzioni di uso meno comune.

Qui di seguito si veda un esempio in cui si prende la base DeCS, al momento su piattaforma Linux e si prepara il suo trasferimento in un sistema Windows.

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

## msrt

Comando piuttosto utile per elaborare dati è `msrt`, esso effettua la riorganizzazione di un *master file*, secondo un criterio che si basa sul contenuto dei campi della propria base dati. Così è possibile ordinare i record della base per data di edizione (ammesso che la base disponga di un campo per la registrazione di un tale contenuto), o secondo un criterio anche più complesso, come potrebbe essere quello per ordine alfabetico di autore e titolo insieme.

Un esempio applicativo è mostrato qui sotto, la base *Lilacs* viene ordinata per data di pubblicazione (campo 65) del documento, con ulteriore criterio di discriminazione sul titolo del documento.

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

## mxtb

Altro comando di grande utilità è `mxtb`, grazie a cui si pongono i dati di una base in assetto tabellare e così si riesce a creare una nuova base (il cui nome viene specificato nel comando di esecuzione) normalizzata e che contiene la frequenza dei dati così come estratti dal formato specificato. L'insieme dei dati da porre in tabella può venire selezionato con un'espressione di ricerca da svolgere nella base dati, il che consente di ritagliarsi i dati da tabulare.

Si presenta un esempio riutilizzabile in cui vengono posti in tabella gli anni di pubblicazione (le quattro prime posizioni del campo indicato con `v65`) dei record con un determinato titolo di rivista nella base *Lilacs*.

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

Nella Tabella-1, qui sotto, si mostrano varie forme di esecuzione dei comandi che coprono situazioni d'uso comune, in modo che possano servire da modello e guida per l'esecuzione di procedure ricorrenti.

<b>Funzione</b>	<b>Modello di sintassi del comando</b>
Leggere a partire da un certo record	<i>mx &lt;base&gt; from=&lt;mf_n_inicial&gt;</i>
Leggere un intervallo di record	<i>mx &lt;base&gt; from=&lt;mf_n_inicial&gt; to=&lt;mf_n_final&gt;</i>
Appurare il numero totale di record	<i>mx &lt;base&gt; +control count=-0</i>
Creare una copia della base	<i>mx &lt;base_in&gt; create=&lt;base_out&gt; -all now</i>
Creare copia della base escludendo i record cancellati	<i>del &lt;base_out.mst&gt; (o rm &lt;base_out.mst&gt;) del &lt;base_out.xrf&gt; (o rm &lt;base_out.xrf&gt;) mx &lt;base_in&gt; append=&lt;base_out&gt; -all now</i>
Generare un file ISO a partire da una base dati	<i>mx &lt;base&gt; iso=&lt;archivio.iso&gt; -all now</i>
Generare una base dati a partire da un file ISO	<i>mx iso=&lt;archivio.iso&gt; create=&lt;base&gt; -all now</i>
Generare una base dati a partire da un file ISO con l'mfn specificato da un campo	<i>mx iso=&lt;archivio.iso&gt; create=&lt;base&gt; "proc='vn" -all now</i>
Invertire (indicizzare) una base dati	<i>mx &lt;base&gt; "fst=&lt;campo tecnica formato&gt;" fullinv=&lt;invertido&gt;</i>
Invertire (indicizzare) una base dati in base ad un file FST	<i>mx &lt;base&gt; "fst=&lt;@archivio.fst&gt;" fullinv=&lt;invertido&gt;</i>
Invertire (indicizzare) una base dati che usa caratteri ANSI	<i>mx &lt;base&gt; "fst=&lt;@archivio.fst&gt;" fullinv/ansi=&lt;invertido&gt;</i>
Convertire un insieme di caratteri della base	<i>mx &lt;base&gt; gizmo=&lt;base_gizmo&gt; create=&lt;base_out&gt; -all now</i>
Convertire una base per usarla su una diversa piattaforma	<i>crunchmf &lt;base&gt; &lt;base_out&gt; target=&lt;plataforma&gt;</i>
Convertire l'inverted file per usarlo su una diversa piattaforma	<i>crunchif &lt;invertido&gt; &lt;invertido_out&gt; target=&lt;plataforma&gt;</i>
Formare un file leggibile da Excel	<i>mx &lt;base&gt; "pft= "/vn "/, "/vn "/, "/vn "//" -all now &gt; &lt;archivio.csv&gt;</i>
Importare un file proveniente da Excel (CSV)	<i>mx "seq=&lt;archivio.csv," create=&lt;base_out&gt; -all now</i>
Estrarre dati da una base	<i>mx &lt;base&gt; "pft=&lt;formato&gt;" -all now /sort -u (linux)</i>
Effettuare una statistica di campi e caratteri della BD	<i>mxfo &lt;base&gt; create=&lt;base_out&gt; 0 noedit</i>
Ordinare una base dati secondo un campo	<i>msrt &lt;base&gt; &lt;#carac_de_la_clave&gt; &lt;formato_de_extracción&gt;</i>
Tabulare i campi della base dati	<i>mxtb &lt;base&gt; create=&lt;base_out&gt; &lt;largura:formato&gt; [class=nnnnn]</i>

Tabella-1: Comandi più comuni di CISIS

# Aspetti pratici

## Equivalenza fra comandi Linux e Windows

Gran parte delle funzioni effettuabili in un determinato sistema operativo hanno il loro equivalente in un sistema diverso. Qui sotto, nella Tabella-2, è presentata una lista di equivalenze fra Linux, DOS e Windows.

<b>LINUX</b>	<b>DOS</b>	<b>WINDOWS</b>
<i>clear</i>	<i>cls</i>	Non si applica
<i>ls</i>	<i>dir /w</i>	Riquadro destro dentro l'Explorer di Windows
<i>ls -l</i>	<i>dir</i>	Riquadro destro dentro l'Explorer di Windows
<i>mkdir</i>	<i>mkdir</i>	Creare una nuova cartella (sottodirectory)
<i>rmdir</i>	<i>rmdir</i>	Cancellare una cartella (purché vuota)
<i>cd</i>	<i>cd</i>	Spostarsi navigando verso una nuova cartella nell'Explorer di Windows
<i>rm</i>	<i>del</i>	Cancellare uno o più file
<i>rm -R</i>	<i>deltree</i>	Cancellare una cartella (purché vuota)
<i>ftp</i>	<i>ftp</i>	Usare un applicativo per l'FTP (es. ws-ftp)

Tabla-2: Equivalenze di comandi fra diversi sistemi operativi

## Comandi FTP più comuni

Quando si devono trasferire file tramite FTP (File Transfer Protocol) e l'applicativo funziona con comandi impartiti da tastiera e non con interfaccia grafica, vi sono

parecchi comandi comuni alla maggior parte dei programmi di ftp. Nella Tabella-3 sono elencati alcuni dei comandi ftp più comuni.

<b>Comando</b>	<b>Explicación</b>
<i>open</i>	Stabilisce la connessione con un sistema remoto
<i>ls</i>	Visualizza il contenuto della directory corrente
<i>cd</i>	Cambia directory corrente
<i>bin</i>	Attiva il trasferimento in modalità binaria
<i>asc</i>	Attiva il trasferimento in modalità testo ASCII-ANSI
<i>put</i>	Invia file al sistema remoto
<i>mput</i>	Invia più file al sistema remoto
<i>get</i>	Prende file dal sistema remoto
<i>mget</i>	Prende più file dal sistema remoto
<i>prompt</i>	Attiva/disattiva comunicazione interattiva con conferma

Tabla-3: Comandi FTP più comuni

## Conversione di un insieme di caratteri

Per realizzare la conversione dell'insieme di caratteri della base dati si usa il comando **mx**, che è una delle componenti, in realtà la più potente e quella d'uso più comune, di CISIS, assieme con una base gizmo, secondo quanto già accennato.

Assumendo che l'obiettivo dell'operazione sia la trasformazione di dati di un insieme di caratteri ASCII Code Page 850 in ANSI, il seguente comando la realizza:

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

La base dati d'origine usata nell'esempio ha nome BASE\_IN, ma beninteso può avere qualsiasi nome. La base gizmo, come nell'esempio, dimora nella directory corrente, verrà creata una nuova base dati di nome BASE\_OUT (e anch'essa potrà avere un nome qualsiasi) con l'insieme dei caratteri convertito.

Nell'esempio sottostante la base LILACS che implementa l'insieme di caratteri ANSI, mediante il gizmo ANSI → ASCII CP850 che si trova nella directory **tabs**, vicino a quella corrente, viene convertita creando la base LIL850.

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

## Trasferire basi dati sotto un sistema operativo diverso

Se si vuole trasferire una base dati (solo il *master file*) da una piattaforma ad un'altra, le due avendo sistemi operativi diversi, si deve usare il comando **crunchmf**, che è una delle componenti del CISIS.

Supponendo che la base si trovi in un server (*server*) Linux e che debba venire trasferita su una macchina Windows, il comando seguente realizza la procedura:

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

La base dati d'origine si chiama `BASE_ORI`, e beninteso può avere qualsivoglia nome. Si creerà una base dati d'arrivo dal nome `BASE_DEST` (che del pari può avere qualsiasi nome) e la conversione del formato fisico dei file viene indicata grazie al parametro **target=** che in questo caso è **pc**, cioè un sistema che usa Windows.

Nell'esempio sottostante, la base GANSNA, al momento su PC con Windows, viene preparata per venire trasferita verso una macchina Linux, dentro alla directory LNX.

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

Lo stesso schema si applica al trattamento dell'*inverted file*, impiegando però il comando **crunchif**, altra componente di CISIS.

# Riferimenti bibliografici

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 [visto il 30 August 2006]. 164 p. Accessibile in 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 [visto il 30 Agosto 2006]. p. 21-26. Accessibile in internet: <<http://www.cnea.gov.ar/cac/ci/isis/isidams.htm>>.

# Glossario

- **Backup.** Procedura grazie alla quale uno o più file e/o *directories* vengono duplicati su un altro supporto di memorizzazione (nastro o disco), allo scopo di produrre una copia di sicurezza, che potrà venire ripristinata nel caso in cui dei dati vengano cancellati accidentalmente o si verifichi un danno fisico ai dati originali.
- **Base dati.** Raccolta di dati strutturati allo scopo di rendere possibile avere accesso ad essi e manipolarli facilmente. Composta di unità denominate record i cui diversi attributi sono rappresentati da campi e sottocampi. Per esempio, in un archivio indirizzario dei clienti", ogni cliente rappresenta un record che possiede vari campi, come "NOME", "CODICE DEL CLIENTE", "TELEFONO" etc.
- **Basi dati bibliografica.** Versione elettronica di un catalogo o indice bibliografico.
- **Campo.** Elemento di un record che consente di registrare delle informazioni particolari. Cfr. *Base dati*.

- **CDS/ISIS - MicroISIS.** Software sviluppato e mantenuto dall'UNESCO per il trattamento di dati bibliografici.
- **Chiave.** Espressione che identifica una o più informazioni incluse in una determinata classe o tipo e che può venire usata in fase di ricerca.
- **File.** In informatica, un insieme di dati che può venire scritto su un supporto di memorizzazione. I file di dati sono creati per delle applicazioni, come ad esempio un sistema di elaborazione testi.
- **Formato elettronico.** Qualsiasi forma di memorizzazione, recupero e presentazione delle informazioni suscettibile di venire trasmessa online o per scrittura su supporti magnetici o ottici.
- **Formato ISO (di scambio dei dati).** Modello definito dalla ISO per lo scambio di dati fra istituzioni, reti ed utenti, ci si riferisce alla norma ISO 2709.
- **Formato LILACS.** Formato de Descrizione bibliografica definito da BIREME, basato su UNISIST Reference Manual for Machine-readable Bibliographic Descriptions.
- **Indicizzazione.** Procedimento attraverso cui si identifica e descrive il contenuto di un documento con termini che rapresentano i temi del documento, allo scopo di consentirne dopo il recupero.
- **Record.** Insieme strutturato di dati che permette di registrare una determinata informazione. Cfr. *Basi dati*.
- **Sottocampo.** Elemento che contiene una porzione della informazione di tutto un campo, e il cui significato può risultare non chiaro se non viene analizzato assieme con gli altri elementi correlati.

- **UNISIST**. Programma intergovernativo relativo alla cooperazione nel campo dell'informazione scientifica e tecnologica.