

# Information Development

<http://idv.sagepub.com>

---

## **CDS/ISIS Information**

Alan Hopkinson

*Information Development* 1993; 9; 24

DOI: 10.1177/026666699300900207

The online version of this article can be found at:

<http://idv.sagepub.com>

---

Published by:



<http://www.sagepublications.com>

**Additional services and information for *Information Development* can be found at:**

**Email Alerts:** <http://idv.sagepub.com/cgi/alerts>

**Subscriptions:** <http://idv.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.co.uk/journalsPermissions.nav>

# CDS/ISIS INFORMATION

Alan Hopkinson

This column is a regular feature, intended for the exchange of information about the CDS/ISIS software package running under MS-DOS on IBM Personal Computers and compatibles. Readers who have any comments, queries, suggestions, hints or advice on the use of the package are asked to send them direct to:

Alan Hopkinson, Systems Development Officer, Tate Gallery, Millbank, London SW1P 4RG, England. Tel. +44 (71) 821-1313. Fax: +44 (71) 931-7687.

If reporting problems, please indicate which version of CDS/ISIS is being used.

## PROBLEMS, PROBLEMS!

One or two readers of this column have been concerned that CDS/ISIS is all problems!

In reply, it has to be said that all software packages have problems from time to time and CDS/ISIS has no more than any other information retrieval package, which is remarkable considering its power and complexity. The main difference between CDS/ISIS and other packages is that commercial packages have a network of commercial distributors who are supposed to (but do not always) provide helplines. CDS/ISIS relies much more on user groups and columns like this one and the one in the *ASTINFO Newsletter*<sup>1</sup>, and more recently on documents like *ISIS facile*<sup>2</sup> reviewed below. Most problems occur in organizations with only little access to computer expertise or when very large or very complicated databases are being mounted on CDS/ISIS. So, do not be disheartened when you read this column. There are commercial packages where many more problems arise and where many of the solutions that are described in this column are just not available!

## ISIS FACILE

A new book, *ISIS facile* [ISIS made easy]<sup>1</sup> is an introduction to CDS/ISIS in French.

As the author, Jean-Pierre Mitsch says:

The CDS/ISIS software is a paradox. On the one hand it is a package which is outstanding in its powerful functionality; on the other its accompanying manual, despite being well-organized, has put off more than one potential user.

The book is well-illustrated, so that even readers with only a smattering of French will find it useful. The illustrations consist of copies of screens with annotations. Of course, the screens are in French, but then readers could, if they wished, run the French menus to help them work out what points the author was making.

The text is humorously written (it reminds the reviewer of texts written to introduce computers to schoolchildren, which is no bad thing!). For example, the French equivalent of I on the ISISDEF menu (Re-initialize database) is marked with a skull and crossbones! This facility is probably the most dangerous on the menu since it can lead to the wiping out of all the data in the database which is open at the time. Searching is also very effectively covered in just a few pages, and this includes both index and text searching.

Though the book is 'ISIS made easy', it is not just for the beginner. There are four PASCAL programs in chapter 8 which solve problems addressed in other chapters. One takes fields from screen entry and places them in subfields of one field; the next imports into ISIS a Macintosh file; the third converts and exports data in a different form and the fourth allows two databases to be 'open' at once.

## ADVANCED SEARCH TECHNIQUES

Many advanced users began with CDS/ISIS version 1, where search techniques were less complete than they are in versions 2.3 or 3.0. Many such users have not kept abreast of the different search techniques that are now available.

In principle, all the facilities of the print formatting language have been available

since the introduction of free text searching (otherwise known as string or text searching). Free text searches always begin with '?'. They allow you to search for the presence of any pattern of characters in a particular field throughout the database. It is not possible to do a free text search on more than one field in the database at once. The same '?' mechanism can be used to search in numeric fields in what are not truly free text searches.

The function 'val' is very useful. Data in CDS/ISIS fields are stored as strings. As far as the program is concerned '123' is no more different from 'ABC' than from 'XYZ'; all are three characters of text in length. The user may, however, sometimes prefer to treat '123' as a number rather than a string of text. This can now be done with the function 'val'. For example, val(v400) will take the first numeric string it finds in field 400 and convert it to its 'numeric value'. This can then be used in calculations. It can be used also to compare with another number. One can thus make a search:

? val(v440`b) > 1989

to find dates since 1989 in subfield b of field 440 (the CCF field for date of publication).

One CDS/ISIS user wished to find out how many books in a database were of less than eight pages. Field 300`a (the MARC field for collation) in every record held the pagination. A sample of fields follows:

axii, 340p`c25cm  
a145p`c21cm  
axi, 340p,(5)p`c25cm  
a4p`c25cm

The function val used in the expression ? val(v300`a) will read the first numeric value it finds in the designated field or subfield, so in this case it ignored the data in subfield c, it ignored roman numerals, and the only data lost from the above example was the (5)p. The search was achieved by:

? val(v300`a)>5

A number of readers have wanted to find all the records of a database except those with a particular attribute. To look for everything without an index term 'London' can be done as follows. First, search for LONDON (from S on the EXGEN menu). The result will be given as, say, Set 1. Now it is necessary to say 'all the records and not set 1'. This can be done as follows, again from S on the EXGEN menu:

? mfn > 0 and mfn 32000

This becomes search 2, and will take some time as it operates like a string search. Then return to the menu and key S again and enter the search:

#2 ^ #1

This means 'the full range of records' and not 'LONDON'.

If you want to find the records containing 'LONDON' but after record number 3000, you can say:

?#1 mfn 3000

The string search will then look only at the records retrieved in Set 1 and will be much quicker than doing a string search through all the records in the database.

Note that the expressions that can be used in a 'string' search include 'relational' expressions as follows:

- = equals
- <> not equal to
- < less than" less than or equal to
- > greater than
- >= greater than or equal to
- : contains

Note also that, as far as text is concerned, = (equals) and the operators other than : (contains) are sensitive to case, so that Unesco is different from UNESCO. However, : (contains) does not distinguish case, and so is more appropriate for searching for text, particularly since most information retrieval systems do not make any distinction as to case when performing searches and users will not expect to have to bear case in mind.

The following Boolean operators are used:

AND  
OR  
NOT

These are not case sensitive.

## INVERTED FILE SEARCH TECHNIQUES

Remember that in searching the inverted file or index, the operators are:

- \* for AND
- + for OR
- ^ for NOT.

When text searching, the operators must be written out and the use of the symbols \*, + and ^ will create an error.

Section 7 of the CDS/ISIS manual<sup>3</sup>, entitled 'The CDS/ISIS Search Language', contains in subsection B the rules for searching through the inverted file. As well as the usual operators, it shows how other kinds of search may be made, such as searches specifying data in a given field or fields, and proximity searches.

### SEARCHES ON A GIVEN FIELD OR FIELDS

CDS/ISIS has one inverted file or index for any data that is extracted during indexing. Other systems have separate indexes for different fields. Thus title words would be placed in a different index from complete titles and from authors or corporate bodies. Although CDS/ISIS has only one inverted file, it does keep a record of the field from where each term has been extracted and this can be used to narrow a search. This is particularly useful on large databases.

A database may contain a field for type of material: book, serial, film, etc. This field (field 060) may be inverted. It could be inverted with the following line in the field select table:

60 0 "TYPE="v60

In that case, in the inverted file, 'book' would be extracted as "TYPE=BOOK" and would have to be searched as such. If it is extracted in the following way:

60 0 v60

then BOOK would appear in the inverted file. But, book might also occur as a title word, having been extracted as:

200 4 v100 ^ a

where the title is found in field 200 subfield a. If it is necessary to retrieve all books from the database, a search using just the word 'book', would bring out both types of material and title words. Thus, a serial entitled *Book World* would be retrieved when it was not wanted.

The solution lies in searching as follows:

book/(60)

This retrieves only the word 'book' extracted into inverted file 60. If the line in the field select table reads:

1 0 v60

then the search for 'book' from field 60 would be accomplished by:

book/(1)

Note that more than one identifier may be contained in parentheses:

book/(1,200)

would find 'book' in fields 1 and 200.

### PROXIMITY SEARCHES

These work only on fields which have been indexed on a word-by-word basis. They are useful for cutting down on the number of records retrieved when searching large databases. For example, to obtain everything on databases it might be necessary to search for data base and data bases as well as database\$ (which means words longer than 'database' truncated). To ensure that a title like 'The bases of data structuring' would not be retrieved, you can search for data adjacent to bases. The search statement would be:

data . base\$ + database\$

Proximity searching will probably not be used very much, but when it is needed it will prove invaluable and should therefore be considered.

Advanced index searching requires knowledge of the structure of the database since it is useful to be able to search on a particular field.

### PROBLEMS WHEN PRODUCING A PRINTED INDEX

Figure 49 on page 152 of the CDS/ISIS manual<sup>3</sup> and the accompanying text illustrate the production of a sample compact index; the example given is an author index. A reader writes to ask if it is possible to get CDS/ISIS to print the numbers on the same line as the index term, which in this example is the name of the author.

It appears not to be possible, since ISIS

will always print a new line after a heading (which is what the author's name is in this example). The best solution is to remove the carriage return after each heading in a wordprocessor. If you prefixed every heading with a special character like +, it would be easy with most wordprocessors to do a global edit removing every + and its following carriage return.

To do this, the FST for first key in figure 47 on page 150 of the manual should read 1 0 (v70:+:./). When producing the index, make it in one column, and use the wordprocessor to format the columns. Then do a 'global change' in the wordprocessor to replace + CR (carriage return) by one or two spaces.

## REFERENCES

1. *The ASTINFO Newsletter*. Bangkok, Unesco. ISSN 0215-1057. This is the newsletter of the Unesco-supported Regional Network for the Exchange of Information and Experience in Science and Technology in Asia and the Pacific (ASTINFO), and includes a Mini/Micro CDS/ISIS Corner. The newsletter is published by: ASTINFO Secretariat, GPO Box 1425, Bangkok 10501, Thailand. Tel. +66 (2) 390-2564. Fax: +66 (2) 391-0866. Tlx. 20591 TH.
2. Mitsch, Jean-Pierre. *ISIS facile*. Louvain-la-Neuve, Academia, 25/115 Grand'Rue, B1'348 Louvain-la-Neuve, Belgium, 1992. 158p. Price: BEF 580.00.
3. *Mini-micro CDS/ISIS reference manual (version 2.3)*. Paris, Unesco, 1989.

## NEWS OF CDS/ISIS USER GROUPS

### British User Group

A CDS/ISIS User Group, known as ISIS Plus, has been formed in the United Kingdom and is interested in making contact with groups in other countries.

*Further information from:* Chris Addison, Natural Resources Institute, Chatham Maritime, Kent ME4 4TB, England. Tel. +44 (634) 880088. Fax: +44 (634) 880066/77.

### Italian National Users Group

The international CDS/ISIS meeting in Viareggio (see *Information Development*, Vol. 8, no. 4) was followed by an Italian National Users Group meeting, attended by around 1,000 participants. Papers described CD-ROM at the National Library and in the region of Tuscany. Standard formats like CCF and UNIMARC are used extensively in Italy and a large number of ISIS Pascal programs have been written. A cooperative of users has been set up which, as well as acting as a distributor, is producing a newsletter in Italian giving information and news about CDS/ISIS.

*Further information from:* DBA, Largo Fratelli Alinari 11, 50123 Florence, Italy. Fax: +39 (55) 218620.

### Users' Club in Thailand

The CDS/ISIS Users' Club in Thailand, started in November 1989, now has 200 institutional and 197 individual members. The objectives of the Club are:

- to serve as the coordinating centre for CDS/ISIS users in Thailand in sharing and exchanging information, knowledge and experiences
- to distribute, disseminate and repackage information on software development at national and international levels.

The Club publishes a bimonthly newsletter, *ISISTHAI*, organizes conferences, workshops, study tours and other training activities on CDS/ISIS and publishes a directory of CDS/ISIS users in Thailand.

Although CDS/ISIS is very popular in Thailand, one major problem which has still not been totally resolved is that CDS/ISIS, like other software packages, cannot sort Thai records alphabetically due to the different levels of vowels which have to be sorted.

*Source:* Ekvanich, Ploenchan. CDS/ISIS Users' Club in Thailand. *ASTINFO Newsletter*, vol. 7, no. 4, October-December 1992, 5-6.

### CDS/ISIS Bibliography from Malaysian User Group

Aishah Mohammed Nasir has compiled a bibliography on behalf of the CDS/ISIS User Group of Malaysia which includes eighty-five references and is intended to be updated.

*Available from:* Library, Universiti Pertanian Malaysia, 43400 UPM, Serdang, Selangor, Malaysia. (price USD 5.00).

### Jordan User Group

The excellent arabization of CDS/ISIS is the main reason for its popularity in the Arab world in general and in Jordan in particular. Over 120 institutions in Jordan have been licensed to use the package, and more than sixty have used it to establish bibliographic or non-bibliographic databases.

The CDS/ISIS User Group in Jordan was set up in 1989 and has been active in providing training and a forum for the exchange of experience. A monthly meeting for