

**ADDIS ABABA UNIVERSITY**  
**SCHOOL OF GRADUATE STUDIES**  
**SCHOOL OF INFORMATION STUDIES FOR AFRICA**

**EXPLORING THE FEATURES AND FACILITIES OF CDS/ISIS FOR  
WINDOWS AS COMPARED TO ITS DOS VERSION AND OTHER  
TEXT RETRIEVAL SOFTWARE.**

**A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE  
REQUIREMENT FOR THE DEGREE OF MASTER OF SCIENCE IN  
INFORMATION SCIENCE**

**BY**  
**SAMSON TIBEBU**

**MAY 1997**

ADDIS ABABA UNIVERSITY  
SCHOOL OF GRADUATE STUDIES  
SCHOOL OF INFORMATION STUDIES FOR AFRICA

EXPLORING THE FEATURES AND FACILITIES CDS/ISIS FOR  
WINDOWS AS COMPARED TO ITS DOS VERSION AND  
OTHER TEXT RETRIEVAL SOFTWARE

By

Samson Tibebu

**Name and Signature of Members of the Examining Board**

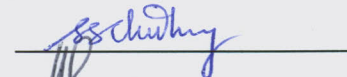
Ato Getachew Birru, Chairman, Examining Board



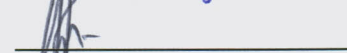
Dr. G.A. Alabi, Advisor




Dr. G.G. Chowdhury, Advisor



Dr. Taye Tadesse, Internal Examiner



Prof. A. Chatterjee, External Examiner



## ACKNOWLEDGMENT

I would like to express my gratitude to my advisors Dr. G.G. Chowdhury and Dr. Gbade A. Alabi for their constructive criticism, and material and intellectual support.

I am highly indebted to the CDS/ISIS Developers Team, and especially to Mr. Davide Storti, member of the team, who was so kind and resourceful during the thesis work, and without his assistance this work would not have been possible.

I am very grateful to all my office colleagues and friends, in particular Ato Sultan Mohammed and Ato Workshut Lamenu, for their moral and material support.

DEDICATION

*To my mother W/o Woubitu Demissie*

## ABSTRACT

Micro-CDS/ISIS, suitable for storage retrieval of structured textual data, has been developed and is distributed by UNESCO free of charge to non-profit making organizations. The first version was released for public use in December 1985. Since then UNESCO has released different versions adding some new facilities and fixing bugs observed and reported by users all over the world. In addition to the different releases, different application programs have been developed by UNESCO and other users to augment and enhance the functionality of the standard releases of the package.

Subsequently, taking into consideration the increasing popularity of Microsoft Windows as an operating system for micros, UNESCO has released a Beta version of CDS/ISIS for Windows operating system, called WINISIS, in July 1996. Since then different test versions have been released.

Taking Micro-CDS/ISIS, which is the widely used text retrieval package in developing countries, and WINISIS, which is most likely to follow the suit, this study compared and contrasted them using features and facilities which are common to most text retrieval packages. The study is based on the May 1997 version, which is the pre-release of WINISIS 1.0, obtained by courtesy of the UNESCO CDS/ISIS Developers Team.

In addition, this study enlisted all the new facilities provided by WINISIS, and recommended areas of application of the software, and identified areas for further improvements.

## TABLE OF CONTENTS

<b>DECLARATION</b> .....	<b>ii</b>
<b>ACKNOWLEDGMENT</b> .....	<b>iv</b>
<b>DEDICATION</b> .....	<b>v</b>
<b>ABSTRACT</b> .....	<b>vi</b>
<b>TABLE OF CONTENTS</b> .....	<b>vii</b>
<b>LIST OF TABLES</b> .....	<b>ix</b>
<b>LIST OF FIGURES</b> .....	<b>ix</b>
<b>LIST OF ABBREVIATIONS</b> .....	<b>x</b>

### CHAPTER I

#### INTRODUCTION

1.1 Background .....	1
1.2 Statement of the Problem and Justification.....	4
1.3 Objective.....	6
1.4 Methodology .....	7
1.5 Scope and Limitation.....	7
1.6 Significance of the Study .....	7
1.7 Organization of the Thesis.....	8

### CHAPTER II

#### TEXT RETRIEVAL SYSTEMS

2.1 Introduction.....	9
2.2 Text Retrieval Systems.....	9
2.3 Text Retrieval Packages .....	13
2.3.1 Desired Features of Text Retrieval Packages.....	14
2.3.2 Additional Features of Text Retrieval Packages .....	22
2.3.3 Applications of Text Retrieval Packages .....	22

### **CHAPTER III**

#### **OVERVIEW OF MICRO-CDS/ISIS**

3.1	Development of Micro-CDS/ISIS .....	26
3.2	Brief Overview of Micro-CDS/ISIS .....	27
3.2.1	Services of Micro-CDS/ISIS .....	27
3.3	Overview of WINISIS.....	35

### **CHAPTER IV**

#### **COMPARATIVE STUDY OF MICRO-CDS/ISIS AND WINISIS**

4.1	Introduction.....	45
4.2	Database Definition .....	45
4.3	Data Entry and Modification .....	58
4.4	Database Indexing .....	66
4.5	Database Browse and Search .....	68
4.5.1	Browsing A Database .....	68
4.5.2	Searching A Database.....	70
4.6	Sorting and Printing.....	82
4.7	Modifying A Database.....	88
4.8	Exporting and Importing Records .....	93
4.9	System Utility .....	97
4.10	Security Measures .....	102
4.11	Advanced Programming Facilities .....	104
4.12	New Features of the Formatting Language of WINISIS .....	108
4.13	Summary .....	117

### **CHAPTER V**

#### **CONCLUSIONS AND RECOMMENDATIONS**

5.1	Conclusions .....	120
5.2	Recommendations .....	124
<b>BIBLIOGRAPHY.....</b>		<b>126</b>

## LIST OF TABLES

1.	Comparison of Micro-CDS/ISIS and WINISIS .....	117
----	--	-----

## LIST OF FIGURES

2.1	Major components of a text retrieval system .....	11
3.1	Main menu of Micro-CDS/ISIS .....	28
3.2	Data entry services.....	29
3.3	Information retrieval services.....	30
3.4	Sorting and printing service .....	30
3.5	Inverted file services .....	31
3.6	Database definition service .....	32
3.7	Master file services .....	33
3.8	System utility services.....	33
3.9	System worksheet editing options .....	34
3.10	System menus editing options.....	34
3.11	The Main window of WINISIS .....	36
3.12	Database window .....	36
3.13	Database menu .....	37
3.14	Browse menu.....	38
3.15	Search menu .....	39
3.16	Edit menu .....	40
3.17	Configure menu .....	41
3.18	Utility menu .....	42
3.19	Windows menu.....	43
3.20	Help menu .....	43
4.1	A sample DBNPAR.PAR.....	50
4.2	A sample FDT .....	51
4.3	A sample data entry/modification definition.....	52
4.4	Worksheet element dialog format.....	52
4.5	A sample display format .....	53

4.6	A sample FST.....	54
4.7	A sample data entry/modification in WINISIS .....	62
4.8	Global field definition dialog box.....	64
4.9	Inverted file maintenance dialog box .....	67
4.10	A sample splitted database window.....	69
4.11	Guided search window .....	77
4.12	Expert search window .....	78
4.13	System print worksheet .....	82
4.14	System sort worksheet.....	84
4.15	General page.....	85
4.16	Presentation page.....	85
4.17	Margin page.....	86
4.18	Layout page.....	87
4.19	System export worksheet .....	95
4.20	System import worksheet .....	97

#### **LIST OF ABBREVIATIONS**

ARCT	African Regional Centre for Technology
ESTC	Ethiopian Science and Technology Commission
IDRC	International Development Research Centre
ILO	International Labour Union
ILRI	International Livestock Research institute
NCIC	National Computer and Information Centre
PADIS	Pan-African Development Information System
PAHO	Pan-American Health Organization
SISA	School of Information Studies for Africa
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNFAO	United Nations Food and Agricultural Organization

# CHAPTER I

## INTRODUCTION

### 1.1 Background

Data processing and information handling have been performed down through ages, first by hand, then by machine-assisted manual and now by electronic machine, such as computers. The capacity of computers have been rapidly expanding due to the development in Information Technology (computer technology, communication technology, microelectronics technology, etc.), the hardware and software. Information Technology (IT) has become an indispensable tool in managing the huge amount of information available in today's society due to information explosion.

Most of the computers in 1960s and 1970s were very large, expensive, and owned by large companies; and they were mostly being used for manipulation of numerical data. In this decade microcomputers have proliferated and are becoming increasingly fast, powerful, small, and cheap. Such development has also resulted in emergence of different categories of computer users, ranging from novice users, who are unfamiliar with both specific syntax or general knowledge in the use of computers, to experts, who develop applications for use by others.

Parallel with the development of microcomputer hardware, the developments in microcomputer software have made it possible the availability of different *data base management systems* (DBMSs) and *text retrieval software*, which are used for handling large amount of textual information.

The advances in microcomputer hardware and software have enabled African countries to use computers in an increasing number to facilitate their activities including *information storage and retrieval* (PADIS 1988,1990). Comparison of surveys conducted by PADIS (Pan-African Development Information System) in 1988 and 1990, reveal an increase in microcomputer use in Africa. In the 1988 survey 58 percent of the responding centres had computer equipment, but in 1990 this percentage had increased to 75. There is also an increase in the number of microcomputer of the different institutions, and the majority of them being used in information services centres such as libraries. There is also an increase in the use of Micro-CDS/ISIS software in these institutions.

CDS/ISIS for DOS, called Micro-CDS/ISIS, is an information storage and retrieval package developed by UNESCO (United Nations Educational, Scientific and Cultural Organisation) and distributed free of cost to non-profit making organisation, especially in developing countries. It is an information retrieval software specifically designed to handle structured non-numerical databases. It provides most of the information retrieval functions available in a number of commercially available software. It enables to create a database, and also to create, delete, and display records in a database. In addition to these facilities, it enables automatic creation of *fast access file* (inverted file), retrieval of records from a database using simple or complex search expressions (using *Boolean*, *adjacency* and other operators), and provides an opportunity to enhance the software's capabilities through programs written in ISIS Pascal language. Furthermore, it can handle unlimited number of databases (with a limit of 500 MBytes each), and with a maximum of 8000 characters per record. It works with *variable length fields* which enables users to save more disk space by releasing spaces not occupied by data. As such it can be said that Micro-CDS/ISIS fulfils most of the functional requirements of a typical text retrieval software.

Since the release of version 1 of Micro-CDS/ISIS in 1985, a number of efforts have been made to improve the user-friendliness of the software. For example, in version 2 of Micro-CDS/ISIS, the six programs, which were functioning separately, were integrated and a main menu was provided to access the different functions (UNESCO 1989). Molla Hunegnaw (1993) has developed an interface called System Interface Search Assistance (SISA) which was intended to improve the search facility of Micro-CDS/ISIS. Another attempt similar to this is that of Heurisko, which is an interface used to perform search and retrieval operations on Micro-CDS/ISIS databases (Pozzana 1993). In the area of data entry, Mohammed Rashid Lugongo (1994) has developed a program to ensure the quality of input data entered in to Micro-CDS/ISIS databases. Sisay Fisseaha (1995) has developed a prototype on-line tutorial facility which enable users to learn the software. In addition to the above efforts, several CDS/ISIS Pascal programs are available to assist in the design and development of databases, data entry, global editing of records, generation of reports, etc.

However, even if some efforts have been made to improve the features and facilities of the DOS version, still in the latest version of Micro-CDS/ISIS (version 3.071) there are problems reported by users. After a careful evaluation of the Micro-CDS/ISIS search interface under various criterion, such as user friendliness, user instruction, output control, and documentation, Perera (1992) has come to the following conclusions:

- Micro-CDS/ISIS is user friendly to experts but hostile to novices;
- does not provide adequate user instruction, and controlling output is not an easy task for novice users; and
- the documentation which is available with the standard package does not provide much help for the novice but they are useful for experts.

Hence, considering these problems and taking into considerations the increasing popularity of Microsoft Windows as an operating system in micros, UNESCO has released the Windows Beta version of CDS/ISIS, called WINISIS, in July 1996; since then different several Beta versions have come out, and the first official version for public use is to come out on July 1997.

## **1.2 Statement of the Problem and Justification**

Micro-CDS/ISIS is increasingly being used by different institutions around the world. International information programs, such as PADIS, are developing in-house databases using Micro-CDS/ISIS package and they organise training in the use of CDS/ISIS in library, documentation and information centres in Africa. Also, ILRI (International Livestock Research Institute, Addis Ababa, Ethiopia - the former ILCA), ARCT (the African Regional Centre for Technology, Dakar, Senegal), and NCIC (National Computer and Information Centre ,Addis Ababa, Ethiopia, which is under Ethiopian Science and Technology Commission, ESTC) offer training in the use of CDS/ISIS. At the regional School of Information Studies for Africa (SISA) of the Addis Ababa University, CDS/ISIS is one of the software introduced to students. International and national databases such as AGRIS of UNFAO (the United Nations Food and Agriculture Organisation), and LILCAS of PAHO (the Pan-American Health Organisation) are available on CD-ROMs as CDS/ISIS databases. In general, some twenty thousand users are reported to be using CDS/ISIS in different parts of Africa and around the world (International Classification 1993; Neelameghan 1992), and in Africa. over 200 centres have the Micro-CDS/ISIS software package (Molla 1993), but currently it is estimated over this figure.

In the last 20 years, in addition to Micro-CDS/ISIS and WINISIS, many information storage and retrieval software have been launched and they show a large variety in what they do and

in how they do it. Indeed, the flexibility and potential of a software package which offers text retrieval facility differ from one package to another.

Citroen (1989) states that: *the big problem facing anyone new in the field of information storage retrieval activities is the enormous number of packages on offer*. There are so many text retrieval packages on the market, and the description of them is so unclear, especially to the uninitiated ones, that, it is very difficult to make a well-founded choice.

Software directories exist which focus on software for information storage and retrieval; these can be very useful, but they are mostly restricted to factual data provided in producers' documentation (Sieverts et al, 1991). Although software reviews may give very valuable information about a particular package, they hardly facilitate a comparison between different packages, as it is a hard task to determine whether the same properties of those programs have been described or tested. Comparative data are given in a few cases only (Sieverts et al, 1991). Other publications just list possible properties and selection criteria without giving actual data.

In addition, since the Windows version of CDS/ISIS is a new one (yet to be released officially), and it is most likely that users are those found in Third World countries, as its DOS version, it is necessary to conduct a study to generate relevant information about WINISIS and to decide on the applicability of it.

Also, to assist those potential or actual users of text retrieval packages, and in particular those of Micro-CDS/ISIS and WINISIS, it is necessary and very helpful to explore the features and facilities of WINISIS as compared to its DOS version and other commercially available text retrieval software. This will help people make their own checklist of their wishes and compare them with the result of this work in order to determine on WINISIS as their future text

retrieval software. The study will also encourage the development and facilitate the use of information storage and retrieval systems in Africa one way or another.

### **1.3 Objectives**

#### **General objective**

The general objective of this study was to explore the different features and facilities of the Windows version of CDS/ISIS, called WINISIS, as compared to its DOS version, called Micro-CDS/ISIS, and other text retrieval software so as to propose areas of activities where the Windows version can be applied, and to recommend measures that should be taken for further improvement of the new version.

#### **Specific Objectives:**

In order to achieve the above general objective, this work had set the following specific objectives:

1. To identify the features and facilities of conventional text retrieval software;
2. To perform a review of the facilities of the Windows and DOS versions of CDS/ISIS;
3. To compare and contrast features and facilities of the Windows version of CDS/ISIS with its DOS counterpart;
4. To suggest areas of activities where the Windows version can be utilised; and
5. To identify shortcomings of the Windows version and propose areas for further improvement.

## **1.4 Methodology**

Since WINISIS has not yet been released officially, a pre-release version (June 1997) was obtained from UNESCO for this study. The information concerning the facilities and capabilities of WINISIS have been collected from available literature review and actual usage of WINISIS and Micro-CDS/ISIS .

## **1.5 Scope and Limitations**

Due, mainly, to the limitation of time, the exploration have been performed irrespective of the type of potential users of WINISIS: novice, intermediate or expert.

In addition, since the official release of WINISIS for public use is yet to come (as checked at the beginning of May 1997), this work utilised the pre-release version (May 1997) of the software. Hence, actual search time and retrieval speed of the software have not been considered as these features vary according to the size and complexity of the database in use, and other hardware and software factors.

## **1.6 Significance of the Study**

CDS/ISIS software is increasingly being used in several institutions in Africa. As explained before, several international and national organizations are using CDS/ISIS (the mini and the micro versions) and some organizations are promoting its usage. In addition, although Micro-CDS/ISIS was initially intended and used for developing bibliographic databases, it is now being used increasingly in developing factual databases and object-oriented databases (Neelameghan 1992). Over 30 institutions are using CDS/ISIS software in Ethiopia (Birru 1994). Therefore, the result of the study can be used by users of Micro-CDS/ISIS in upgrading to the Windows version.

The exploration will also assist those potential users of text retrieval software in identifying available features and facilities of WINISIS, and hence will help them in determining whether to choose it for their need or buying other commercially available text retrieval software. The result can also be used by those people who intend to improve the new version by incorporating additional features and facilities.

In addition, in Ethiopia, the NCIC, which is the assigned local distributor and promoter of CDS/ISIS by UNESCO, and other organizations which promote CDS/ISIS (like ILRI, PADIS etc.) will benefit in getting ready-made information for any wanted advice from users concerning the new version.

### **1.7 Organization of the Thesis**

The thesis is organized into five chapters:

The first chapter deals with background information and justification of the present study.

Review of the features and facilities of software packages which provide text retrieval facilities with possible areas of application of these packages is discussed in Chapter two.

Chapter three gives brief overview of Micro-CDS/ISIS and WINISIS by listing the different functions provided by each package.

Features and capabilities of Micro-CDS/ISIS and WINISIS are compared and contrasted using eleven comparison criteria in Chapter four, including a summary of the comparison made.

The last chapter, Chapter five, list new facilities that are observed in WINISIS, and recommends some areas of application and further improvement for WINISIS.

## CHAPTER II

### TEXT RETRIEVAL SYSTEMS

#### 2.1 Introduction

A variety of different means of organizing information or *knowledge* have been devised and exploited since the earliest of times. In general terms the objective of the organization is to permit the information or the *knowledge* to be found easily again on a later occasion. The organization and later retrieval is known as *Information retrieval* (Rowley 1988:3).

In order to organize information or *knowledge* librarians and information workers create a variety of tools. Traditional tools of information retrieval have been catalogues, bibliographies and printed indexes. But, during the last few decades a variety of approaches to the organization of information have proliferated with the introduction of computer-based methods. Now computer held databases and their indexes are important in the organization of information, and are gradually replacing the traditional tools in a number of applications.

#### 2.2 Text Retrieval Systems

An *information retrieval system* is a device interposed between a potential user of information and the information collection itself (Harter 1986:3). For a given information problem, the purpose of the system is to capture wanted items and to filter out unwanted items. In a town's library, for example, the Card Catalogue or on-line Catalogue is an information retrieval system applied to a local collection of books and other materials.

A functional definition of a *text retrieval system* may be (Chowdhury 1996:61):

*a system that accepts, collates and stores textual information from a variety of sources, retrieves specific information from the text information in response to queries, and process the retrieved text in a user-defined format.*

Text retrieval systems deal with various kinds of documents which are textual in nature. Conventionally, text retrieval systems deal with bibliographic documents/materials characterized by the various keys like author, title, keywords, publication details, abstracts etc. Such systems enable users to search the bibliographic records through any of the keys like author name, title assigned keywords, or by any one or more words occurring in the abstract field. In such a situation, the text retrieval system acts as a bibliographic or reference retrieval system, because the search user have to consult the hard copies of the documents to get the required information.

However, text retrieval systems may also contain full texts of any kind of documents like letters, correspondences, office memos, legal documents, patient records and case histories, complete text of articles and books, and so on. Users of such systems expect to retrieve the complete text or part of it in response to query and can go through it to get the desired information, and such systems are referred as *full text retrieval systems*.

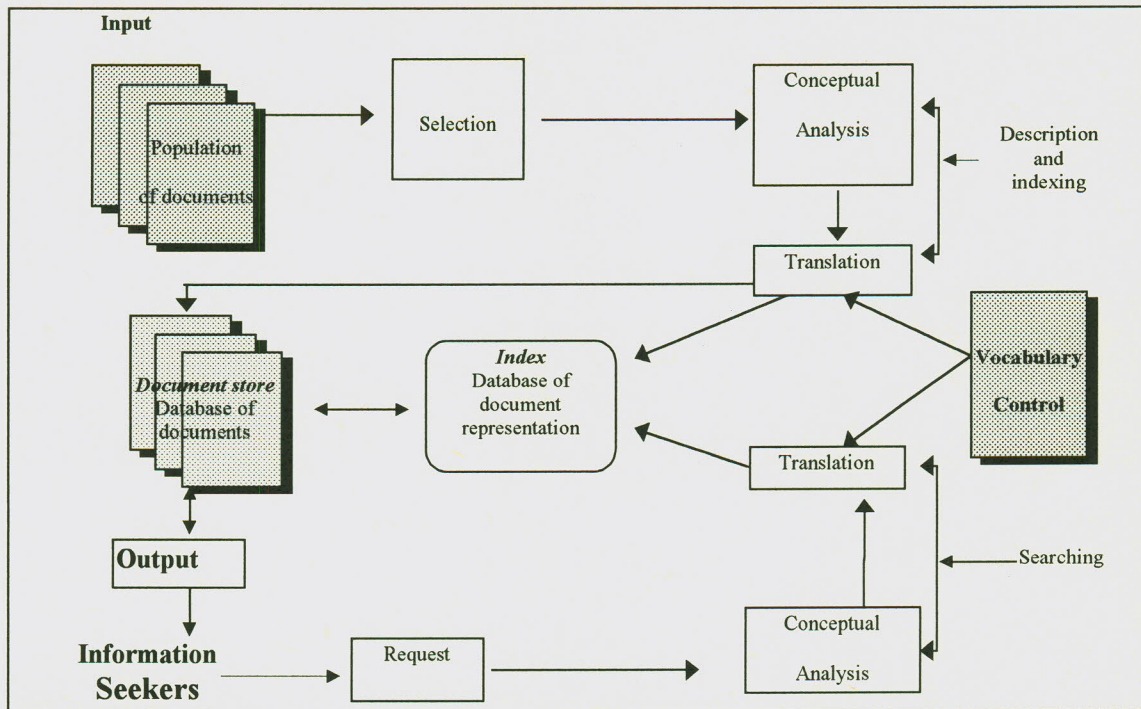


Figure 2.1 : Major components of a text retrieval system

Major components of a typical text retrieval system are presented in Figure 2.1. The *input* consists of documents (printed and other records of all type) acquired by an information center. This implies the existence of selection criteria and policies, which, in turn, implies a detailed and accurate knowledge of the information needs of the community to be served. Once the documents are acquired, they need to be *organized and controlled* so that they can be identified and located in response to the various types of user demand. Organization and control activities include classification, cataloging, subject indexing, and abstracting. Two important elements are the physical description of the document (*descriptive cataloging*) and the choice of access points (e.g., authors, titles) to make the description findable in catalogues and bibliographies.

As shown in Figure 2.1, the subject indexing process involves two quite distinct intellectual steps: the *conceptual analysis* of a document and the a translation of the *conceptual*

*analysis* into a particular vocabulary. For efficient conceptual analysis, the indexer needs both an understanding of the document's subject matter and a good knowledge of the needs of the system's users. The second step in the indexing process is the translation of the conceptual analysis into a particular vocabulary. In most systems, this involves a *controlled vocabulary*, that is, a limited set of terms that must be used to represent the subject matter of documents. Such a vocabulary might be a list of subject headings, a classification scheme, a thesaurus, or simply a list of *approved keywords* or phrases.

Once indexing has been completed, the documents are entered into some form of document store (document database) and the indexing records into a second database, where they are organized in such a way that they can be conveniently searched in response to various types of subject (and other) requests. The database of indexing records, or *document representations*, may be as simple as a card file or an index in printed form. In a modern setting, however, this database is more likely to be a machine-readable file on magnetic tape or disk; it can be regarded as the *index* to the document store. The index shown in figure 2.1 may consist of a bibliographic description of each document in the collection, along with various access points to make this representation retrievable not only by terms describing its subject matter but also by such characteristics as author and title.

The steps involved at the output side of the service are very similar to those involved at input. The user population to be served submit various requests to the information center, and members of the center's staff prepare search strategies for the requests. It is convenient to consider the preparation of search strategies as also involving the two steps of conceptual analysis and translation. The first step involves an analysis of the request to determine what the user is really looking for, and the second involves the translation of the conceptual analysis into the vocabulary of the system. The conceptual analysis of the request,

translated into the language of the system, is the *search strategy*, which can be regarded as a request representation in the same way that an indexing record can be regarded as a document representation.

Once the search strategy has been prepared, it is *matched* in some way against the database of document representations. This can involve a search of card files, printed indexes, microfilm, or magnetic tape or disk. In a contemporary setting, the search might be performed from an on-line terminal at some library whereas the database being searched could be available locally or could be loaded on a computer hundreds or thousands of miles away from the searcher. Document representations that match the search strategy - that is, satisfy the logical requirements of the searcher are retrieved from the database and delivered to the requester, either in printed form or are displayed at a terminal.

### **2.3 Text Retrieval Packages**

A *software* is a program or list of instructions which are necessary to enable a computer system to conduct specific tasks (Rowley 1987:382). A *Software package* is a set of programs to achieve a specific objective to instruct the computer to execute a set of tasks in order to organize information in a pre-assigned manner. *Text Retrieval software* is a software that is specifically designed to support the creation, maintenance, and use of databases of textual data (Kimberly & Rowley 1990:xi). Some such data may also cater for numerical, tabular, graphical data and images.

Kimberly and Rowley (1990:xi) list the following characteristics which typify a text retrieval (TR) package:

- variable-length fields;

- access to records through an inverted file of index keys or text terms which are drawn from the records of the database;
- a range of retrieval facilities which support retrieval based on words in records, where there is a limited control over the form of the search key in the record;
- emphasis on the management of one or more distinct databases, where the ability to draw data from a number of related databases is not critical; and
- fixed applications which require relatively limited programming or system development facilities.

### **2.3.1 Desired Features of Text Retrieval Packages**

Basic features desirable for an ideal text retrieval system may broadly be categorized into two groups (Chowdhury 1996:64-72): *database creation*, and *retrieval provisions*.

#### **A. DATABASE CREATION**

The task of creating a database, using a text retrieval package, is often performed by people with little computing skill, and thus the whole process of database creation should be quite simple. Text retrieval packages have certain basic features which are directly concerned with database creation performance and efficiency of the system. Following are some of the major features:

##### ***1. Variable length fields and records***

In contrast to traditional DBMSs, where a given field occupies more or less the same length in each record, a text database is characterized by the fact that a given field in each record may not always be of the same length. For example, length of an abstract field varies from record to record. Thus, prior allocation of fixed length for each field in a text database

system may not be quite useful. Therefore, text retrieval packages should accommodate variable length fields for best handling of the various retrieval operations and for optimum use of storage space.

### ***2. User-defined record structure and Indexing***

In text retrieval operations, it is always desirable that structure of records in a given database be decided by database designers according to their needs. In other words, users should be able to design the structure of the database to suit the specific needs. Most text retrieval systems provide this facility. In addition, users should also be able to decide which fields are to be indexed and how they will be indexed.

### ***3. Data entry and editing***

Most text retrieval packages provide facilities for on-line data entry and editing, and it is the major feature that should be available in an ideal TR package. However, entering records one by one in the database, particularly in case of retrospective database creation, may be a tedious job. One solution to this could be to download the data from other source(s). This calls for entry of data in a batch mode. However, downloading data involves a number of considerations; the most important ones relate to the compatibility of formats and/or availability of a suitable program for conversion of data from one format to the other. On-line editing of database is required at times, especially when users want to use a database for a specific purpose that was not envisaged to begin with.

## **B. RETRIEVAL PROVISIONS**

Text retrieval software are characterized by the availability of a wide variety of search and retrieval provisions (Kimberly and Rowley 1990; Chowdhury 1996). Basic features relating to retrieval provisions of a text retrieval package are:

### ***1. Boolean query formulation***

A text retrieval system should provide for query formulation by using the Boolean AND, OR, NOT operators, and also should provide *nested Boolean searching*. Boolean search facilities allow a user to use more than one search term in a given search prescription with certain conditions imposed. These conditions specify whether more than one search term should simultaneously be present in the desired records, whether any one of some chosen words should be present, or one or more terms should be present while other term(s) should not be present in the desired records, and so on. Nested Boolean search facility allows to impose more complex conditions along with the search terms. Following is an example of nested Boolean search expression allowed in Micro-CDS/ISIS:

LIBRARY \* (NETWORK + TELECOMMUNICATION)

which will retrieve all those records where the terms LIBRARY and NETWORK occur, and also those records where the terms LIBRARY and TELECOMMUNICATION occur.

It may be noted that text retrieval software uses different conventions for denoting the different Boolean operators.

### ***2. Proximity searching***

In a text retrieval system, there should be a provision for *adjacency/proximity searching*. The purpose of this searching is to refine search statements by permitting a searcher to specify the context in which a term must occur. This search facility allows users to specify whether two search terms should occur adjacent to each other, whether one or more words should occur in between the search terms, whether the search terms should occur in the same paragraph irrespective of the intervening words, and so on.

Text retrieval software differ in using notations for representing various proximity operators. Following are some examples of proximity searches allowed in Micro-CDS/ISIS:

#### MICROCOMPUTERS (G) LIBRARIES

which will retrieve all records where both the search terms occur in the same field; and

#### INFORMATION \$ MANAGEMENT

which will retrieve records where both the search terms occur and in between them at most one word may occur, e.g., Information centre management, Information resources management, *etc.*

### 3. *Search refinements*

The ability to specify a search strategy, and to elicit a response in terms of numbers of postings or records retrieved on the basis of that strategy is fundamental to most text retrieval system. This facility supports the narrowing or broadening of search strategies. Once a search statement is formulated, there should be provisions for refining it in a step-by-step fashion so as to arrive at an optimum level of retrieval.

### 4. *Range searching*

Range searching is most useful with numerical information. It is important in selecting records within certain data ranges. Following options are usually available for range searching:

Greater than (>)

Less than (<)

equal to (=)

Containing

Not equals to ( or <>)

Greater than equals to (>=)

Less than equals to (<=)

Not containing

These operators are used to prescribe a precise condition in a given search statement.

### 5. *Limiting searches*

A database in a text retrieval system comprises of different fields with different items of information. Users in their query formulation should be able to limit the search in one or more fields, *field searching*. Text retrieval software usually provide this facility. For instance, the command:

COMPUTERS \* COMMUNICATIONS /(4)

will retrieve those records where both the search terms occur in a field identified by the field tag 4 in Micro-CDS/ISIS.

### 6. *Display of postings*

A text retrieval system should display the number of *postings* for a given query in a given search session. This helps users approximate whether the number of items retrieved is more, less, or just right. Accordingly, they may just wish to view the retrieved records, print the retrieved records, or refine the search statement in order to arrive at a better result.

### 7. *Truncation*

*Truncation* is a search facility whereby a search can be conducted for all the different forms of a word having the same common root. As an example, the truncated word COMPUT\* will retrieve items on COMPUTER, COMPUTING, COMPUTATION, COMPUTE, *etc.* in Micro-CDS/ISIS.

A number of different options are available for truncation, viz. , *right truncation*, *left truncation*, and *middle truncation*. *Left truncation* retrieves all words having the same characters at the right hand part, e. g. , '\*hyl' will retrieve words like 'methyl', 'ethyl', *etc.* in Micro-CDS/ISIS. Similarly, *middle truncation* retrieves all words having the same

characters at the left and right hand part. For example, a middle truncated search term 'col\*r' will retrieve both the terms 'colour' and 'color'. In the above example, a *wildcard* (\*) is used to allow any letter(s) to appear in a specific location within a word.

### **8. String searching**

*String searching* is the ability to search on character strings within the body of a text in a record, which is usually available for those fields the text of which has not been included in an inverted file, and is, therefore not pre-indexed. In other words, string searching allows one to search those terms which have not been indexed. However, as the process of string searching matches the search term character by character with the stored records, the search process is very slow for a relatively large database. Typically, text string searching may be used on a subset of the database, where that subset having been retrieved by searching on indexed fields.

### **9. Stopword lists**

Use of *stopword lists* helps to avoid inclusion of unsought words while creating inverted file. A text retrieval system should make provisions for using *stopword lists*. This helps in reducing the size of the inverted file thereby making optimum use of the disk space and also making the search faster. In some systems a number of *stopword lists* may be maintained, one for each part of a database.

### **10. Index vocabulary control**

A text retrieval system should make provisions for *thesaurus* support for helping users in choosing appropriate search terms. While searching a database, users can consult an on-line *thesaurus* for choosing appropriate terms, and if they are not satisfied with the output they may choose alternative search terms from the *thesaurus* for further search.

### ***11. Search term widening/narrowing***

Provision of a thesaurus can help users in widening or narrowing a search by using narrower, broader and related terms from the on-line thesaurus. This is particularly useful when the number of retrieved records in a given search session is too low or too high in accordance with a user's expectation.

### ***12. Sorting of search results***

Sorted output of records of a search result often are quite useful to users. Users may want to sort all or a certain number of records in a database in a desired way. For example, users may want to sort the records in a bibliographic database by author, by title, and so on. Provisions for sorting results easily should be available from a text retrieval system.

### ***13. Editing and formatting results***

Users of a text retrieval system may want to print all or a certain number of database records in a chosen format. For example, users of a bibliographic database may want to generate printed library catalogue, a bibliography, an index, and so on. Users may also want an output of a given search in a typical format with necessary editing. Text retrieval systems should also make provision for printing of database records in a chosen format. Users should be able to view the database records in a suitable format, according to AACR2, say.

### ***14. Data security***

Security of database is of particular concern, since most databases offer shared facilities. Hence, in multi-user environment some kind of security is essential. Some of the key security features that are generally available are:

- Password;
- User ID;
- Data encryption; and
- Various levels.

Usually it is necessary to control both databases or parts of a database a user may have access to; and which function a user is permitted to perform in relation to these databases. Passwords are the normal device for controlling access to databases. They may be used alone or in conjunction with other devices such as user ID's or data encryption.

Security can be implemented at different levels. For instance: at system, database level, record level, field level, and command levels by User terminal ID, password or any combination of these. In some systems read/write/delete access can be controlled at the level of individual words or fields; and a user cannot see any facilities which he or she is not permitted to use.

#### ***15. On-line help***

Users often need help in course of searching a database. Same is also required in course of database creation. Text retrieval systems should provide on-line help facilities available from any point of database creation or searching. On-line context-sensitive help make a text retrieval system more friendly to users.

In addition to the basic features of a conventional text retrieval systems, discussed above, since full text retrieval systems handle both bibliographic data as well as full texts of documents, in addition to providing keyword access, these software should discriminate word distances in order to facilitate proximity searching. Full text retrieval software must also recognize and facilitate the interactive nature of full text searching. *Studies have shown that the average number of queries against a full text database prior to display can be ten or more* Ashford and Willet (1986) in Chowdhury (1996). Hence, it is important that

each result be located quickly and new criteria be added to refine the search process. Another useful feature is that which enables user to display retrieved records based on some heuristic determination of relevance to a search. Full text retrieval systems should also highlight the search terms while displaying the retrieved records, and should allow document scrolling, preferably both backward and forward.

In addition to these basic features and facilities, many packages offer a variety of additional facilities that may assist in retrieval or even support specific applications. For example, Micro-CDS/ISIS has advanced programming facilities through which users can develop various application programs that are not readily available in the standard system.

### **2.3.2 Additional Features and Facilities of Text Retrieval Packages**

The flexibility and potential of a software package that offers text retrieval facility will obviously differ from one package to another. In general, a cheaper and smaller micro-computer based packages offer less flexibility and a more limited range of features than the more expensive mainframe or minicomputer based competitors.

In addition to the basic features mentioned above, text retrieval packages may have some additional features and facilities. These have been widely discussed in literature (Rowley & Kimberly 1990; Citroen 1989; Sieverts et al 1991; Sieverts et al 1992; Rowley 1987; Chowdhury 1996).

### **2.3.3 Applications of Text Retrieval Packages**

The serious limitations of the current DBMSs in managing textual information has prevented many organizations from automating document management. Consequently, 90% of all business information is still in paper form: mail messages, policies, procedures,

documentation, contracts, correspondence, and others (Spinello 1994). Given the large productivity and competitive advantage gained from numerically oriented DBMSs, that manage 10% of the information, it is logical to conclude that similar benefits will be acquired through the management of textual information.

Word processing software, email, OCR, on-line databases, and CD-ROM publishing have converged to create vast volume of text that can provide not only operational advantage but also strategic insight for using TR software. Many organizations are recognizing the need for more efficient collection, processing, distribution, and control of their massive volumes of documentation. The current competitive business environment provides a perfect climate for the proliferation of TR systems, and a growing number of organizations have successfully implemented TR systems as an antidote to *information anxiety*.

Many applications can be considered for using TR systems:

- The entire sphere of litigation support is one of the largest application areas for text retrieval technology. With text retrieval software, law offices or corporate legal departments can strongly enhance their approach to automated case management. These software enables them to store and retrieve all legal documentation (e.g., depositions and court transcripts) associated with a particular case. The rapid access to information that these systems provide facilitates the sometimes onerous task of legal research. In addition, automated searches through volumes of data are faster, more accurate and more systematic than manual searches. A menu-driven, natural-language interface can further make case management systems accessible even to novice users.
- Industries, like communications, insurance, and other utilities, can also benefit tremendously from text retrieval systems. Complex regulatory documentation generated by federal and state agencies can be easily incorporated into a full text

retrieval database, especially with the help of optical scanners. Many organizations within these industries pro actively prepare for the possibility of litigation by storing all documents associated with regulatory matters.

- Another area of applications for text retrieval packages is found within certain segments of the consumer-goods industry. Text retrieval software can, for example, be used to maintain key product information. One large brewery has designed a text retrieval system to catalog the results of its taste tests (Spinello 1994). Extensive comments about each beer sample are recorded by the company's many testers. These comments become part of the text database along with related information, such as the tester's name and the date of testing. Product officials can search for all samples that have a certain taste characteristic. This and similar applications help ensure consistently high-quality products.
- Text retrieval system can become a valuable repository for competitive information. Organizations can maintain critical textual information about competitive products, which assists sales and marketing divisions during the sales cycle. Competitive information often comes from many sources, including Internal memos and reports, feedback given to a salesperson, and articles in trade journals or other publications. An integrated software solution can facilitate the process of accumulating information from these diverse sources into a single, comprehensive database of competitive information.
- Publishing represents another potential application area for text retrieval software. Newspapers and other publications are using text retrieval systems to catalog all of their major articles for quick and easy on-line access. Sports columnists wanting to see all the articles published in the past six months mentioning the local quarter back can

initiate a search for all the occurrences of the athlete's name. An archive of articles becomes an invaluable research tool.

Considering the use of TR packages in a library and information related activities, we can find that they can be used for the following applications:

1. On-line information retrieval:

(a) from in-house databases, which may be either catalogue or bibliographic database or factual database relating to persons, institutions, projects and so on.

(b) from external databases, where local packages are used on in-house computer in order to support the interaction with the external database.

2. Printed catalogues and indexes, lists and bibliographies may all be extracted from an in-house database.

3. Printed current awareness bulletins and selective dissemination of information.

4. On-line current awareness services, such as selective dissemination of information or news bulletins.

## CHAPTER III

### OVERVIEW OF MICRO-CDS/ISIS AND WINISIS

#### 3.1 Development of Micro-CDS/ISIS

CDS/ISIS, 'Computerized Documentation System (CDS)/Integrated Set of Information Systems (ISIS)' is a software package developed and distributed by UNESCO (the United Nations Educational, Scientific and Cultural Organization) to non-profit making organizations free of charge. The software is based on the program *ISIS*, which was developed initially in 1965 by ILO (International Labour Organization) and intended for IBM mainframes. By 1987, the Intentional Development Research Centre (IDRC), Canada, has designed MINISIS an ISIS version that can operate on Minicomputers (HP 3000).

By 1986, UNESCO had issued an ISIS version, *Micro-CDS/ISIS*, for use on microcomputers, IBM PC or compatibles with 512 KB memory and a hard disk. There is also a version for the WANG PC and one for the VAX/VMS series. Since then different release of the micro version have been introduced.

Version 1 was released in 1985 (which consists of six independent programs); version 2.0 was released in 1988 (which amalgamated the six independent programs and introduced ISIS Pascal); the next public release was version 2.3 released in March 1989 (which improved indexing speed and the space used by, making the package more resilient, and adding SYSPAR.PAR and DBN.PAR); version 3.0 was released for public in May 1992 (as a 'network sensitive' version by avoiding the network problems of the earlier version); and version 3.07 was released in 1993 (which eliminated some bugs observed in earlier versions and adding some new features like supporting RTF feature). However, at every stage in

the development of Micro-CDS/ISIS, a new Pascal routines and functions are added to the Pascal programming library.

### 3.2 Brief Overview of Micro-CDS/ISIS

Micro-CDS/ISIS is a generalized system designed for computerized management of structured textual databases. Micro-CDS/ISIS enables users to: build, manipulate, maintain, manage, and retrieve records from structured textual databases. Major functions provided by Micro-CDS/ISIS allow users to:

- Define databases containing user defined data elements,
- Enter new records into and modify records from a given database,
- Automatically build and maintain fast access files for each database in order to maximize retrieval speed,
- Retrieve records by their content,
- Display the records or portions thereof to users requirements,
- Sort records in any required sequence,
- Print partial or full catalogues and/or indexes, and
- Develop specialized applications using the CDS/ISIS Pascal programming language.

The above facilities are provided through a set of eight major *services* (see Figure 3.1).

#### 3.2.1 Services of Micro-CDS/ISIS

The various Micro-CDS/ISIS/ISIS services are selected through the main menu, shown in Figure 3.1, which is the first menu displayed each time Micro-CDS/ISIS starts.

Selecting any of the options in this menu, except **C** and **X**, will display another menu or sub menu containing specific options offered by each service. Option **L-Change Dialogue**

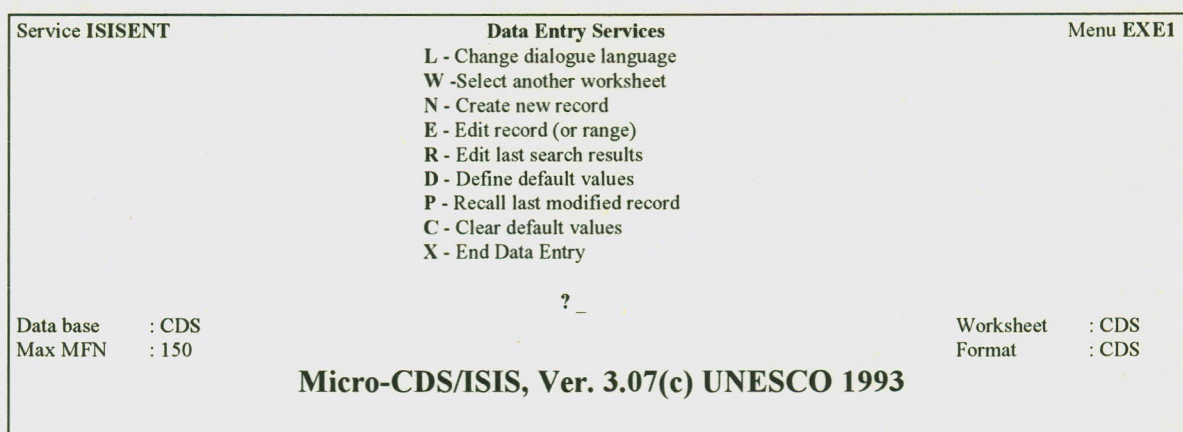
**Language** allows one to select the language in which the system menus, prompts and messages will be displayed (English, French, or Spanish). Initially Micro-CDS/ISIS will start in the language selected as default language during installation(in this case English). However, Micro-CDS/ISIS menus provide an option to change the dialogue language. Option **C-Change Database**- allows one to select the initial database or to switch to another database. But the system will prompt users to provide a database name whenever required, if they don't explicitly select one. Option **X - Exit** terminates the current session and returns control to the operating system.

```
=====
Micro-CDS/ISIS - Version 3.07
=====
C - Change Database
L - Change dialog language
E - ISISENT - Data entry services
S - ISISRET - Information retrieval services
P - ISISPRT - Sorting and printing services
I - ISISINV - Inverted file services
D - ISISDEF - Database definition services
M - ISISXCH - Master file services
U - ISISUTL - System utility services
A - ISISPAS - Advanced programming services
X - Exit
?_
Micro-CDS/ISIS, Ver. 3.07(c) UNESCO 1993
```

*Figure 3.1 : Main menu of Micro-CDS/ISIS*

### **A. - Data Entry Services(ISISENT)**

The ISISENT services provide all functions related to data entry operations. (see Figure 3.2). These functions allow one to add new records to a database (option **N**) and to modify or delete existing records(options **E** and **R**).



*Figure 3.2 : Data entry services*

When creating records, a useful feature allows users to redefine the contents of one or more fields that will then be automatically added to all subsequent new records (option **D**). Another option allows users to select a pre-defined data entry worksheet for the type of record they are creating or modifying (option **W**).

### **B - Information Retrieval Services (ISISRET)**

The ISISRET services provide all functions related to information retrieval (see Figure 3.3). These allow users, in particular, to display the dictionary of search terms and use it in formulating their query (option **T**), submit search expressions (option **S**) and display the corresponding results (option **D**), as well as save search results for subsequent purposes (option **P**).

Other options enables one to display selected sections of the master file (option **B**), re-execute and/or edit a previously submitted search expression, possibly to a different database (option **G**), review the list of submitted search expressions (option **R**) and select a suitable display format (option **F**).

Service ISISRET	<b>Information Retrieval Services</b>	Menu EXGEN
	L - Change dialogue language B - Browse Master file T - Display terms dictionary S - Search formulation D - Display search results F - Change display format R - Recall query formulations G - Execute search results P - Save search results X - Exit	
	? _	
Data base : CDS		Worksheet : CDS
Max MFN : 150		Format : CDS
Micro-CDS/ISIS, Ver. 3.07(c) UNESCO 1993		

Figure 3.3 : Information retrieval services

### C. - Sorting and Printing Services (ISISPRT)

The ISISPRT services allow users to print the output of a given query (which have been previously saved using the **P** option of the ISISRET services) and/or to print a selected range of records. The records may be sorted by virtually any combination of fields and sub fields. The field(s) by which the records are sorted may be used as *headings* in printing.

Each time one uses the ISISPRT services he/she will be asked to provide the specific sorting and page layout parameters he/she require for that particular print run. These are entered through two system worksheets called the *print worksheet* and the *sort worksheet*. Option **P** displays the system print worksheet to accept the print layout parameters. Option **S** enables one to use a pre-defined print worksheet created by ISISUTL services (see Figure 3.4).

Service ISISPRT	<b>Sorting and printing</b>	Menu EXPRT
	L - Change dialogue language S - User print worksheet P - System print worksheet H - Convert Hit file to Master X - Exit	
	? _	
Data base : CDS		Worksheet : CDS
Max MFN : 150		Format : CDS
Micro-CDS/ISIS, Ver. 3.07(c) UNESCO 1993		

Figure 3.4 : Sorting and Printing services

### D. - Inverted File Services(ISISINV)

The ISISINV services perform various operations on an Inverted file (see Figure 3.5). These functions enable users to create (option **F**), update (option **U**), backup (option **B**), restore (option **C**) or print (options **P** and **D**) the Inverted file of an active database.

Service ISISINV	<b>Inverted File Services</b>	Menu EXG1
	L - Change dialog language F - Full inverted file generation U - Update inverted file B - Inverted file backup G - Create unsorted link file S - Sort link file C - Load Inverted file D - Dump inverted file P - Print search term dictionary X - Exit	
Data base : CDS Max MFN : 150	?_	Worksheet : CDS Format : CDS
Micro-CDS/ISIS, Ver. 3.07(c) UNESCO 1993		

*Figure 3.5 : Inverted file services*

Using option **D** one can create a full listing of the inverted file (which includes both the search terms, and for each term the corresponding list of current postings) in a file called **IFLIST.LST**. Options **B**, **S**, and **C** (executed one after another in this order) enable users to create Inverted file step by step, hence have similar group effect as using option **F**, but taking less time especially for large databases.

### E - Database Definition Services(ISISDEF)

The ISISDEF services(see Figure 3.6) enable users of Micro-CDS/ISIS to define new databases (option **C**) or to modify existing database definitions (option **U**). An option is also provided to re-initialize (clear) all records of a database (option **D**).

A Micro-CDS/ISIS database consists in the following components:

**Field Definition Table:** defines fields of the Master record for the database, and is created during database creation.

**Data entry worksheet(s):** one or more screen layouts used to create and/or update the master records of the database.

**Field Select Table(s):** one FST defines fields to be made searchable through the Inverted file, and is created during database creation. Additional optional FSTs are defined to specify sort keys, to be used during sorting and printing, and for reformatting records, to be used during exporting and importing database records.

**Display format(s):** define precise formatting requirements for either on-line display of records and/or printing.

ISISDEF allows users to create and/or modify any one of the above components of a database. Each component is stored as a separate file.

Service ISISDEF	<b>Database Definition Services</b>	Menu	EXDEF
	L - Change dialogue language		
	C - Define a new database		
	U - Modify database definition		
	I - Re-initialize database		
	R - Unlock data base		
	X - Exit		
	?		
Data base : CDS		Worksheet : CDS	
Max MFN : 150		Format : CDS	
Micro-CDS/ISIS, Ver. 3.07(c) UNESCO 1993			

*Figure 3.6 : Database Definition Services*

### **F. Master file services(ISISXCH)**

The ISISXCH services, see Figure 3.7, provide master file backup/restore functions as well as import/export facilities for interchanging data with other systems, including mainframe computers.

Service ISISXCH	<b>Master file services</b> L - Change dialogue language B - Master file backup R - Master file restore C - Reorganize Master file I - Import external file E - Export CDS/ISIS file X - Exit	Menu EXCH
Data base : CDS	? _	Worksheet : CDS
Max MFN : 150		Format : CDS
Micro-CDS/ISIS, Ver. 3.07(c) UNESCO 1993		

*Figure 3.7 : Master file Services*

### F. - System utility services(ISISUTL)

The ISISUTL services provides for miscellaneous system utilities (see Figure 3.8) which enable one to create and edit menus (option **M**) and system worksheets (option **N**), print menus and worksheets (option **P**), define or modify screen attributes (option **A**) and print system message files (option **D**).

Service ISISUTL	<b>System Utility Services</b> L - Change dialogue language N - Create/Edit system worksheets M - Create/Edit system menus P - Print Worksheet/Menu A - Display /modify attributes D - Print System messages H _ Help X - Exit	Menu EXMI
Data base : CDS	? _	Worksheet : CDS
Max MFN : 150		Format : CDS
Micro-CDS/ISIS, Ver. 3.07(c) UNESCO 1993		

*Figure 3.8 : System utility services*

Selecting option **N** from the above menu provides a number of options for creating and editing system worksheets through the menu displayed in Figure 3.9. The following standard system worksheets exist at present:

- xYISI Export worksheet
- xYISO Import worksheet
- xYPRT Print worksheet
- xYSRT Sort worksheet

In addition to these, one may also create *user print worksheet* and/or *user sort worksheets* for specific applications..

Service ISISUTL	<b>System Worksheets Editing options</b>	Menu EXM2
	C - Create new worksheet S - Select a new worksheet R - Restore current worksheet E - Edit worksheet D - Copy worksheet F - Delete worksheet X - Exit	
	<b>Warning! Use ISISDEF to create/update data entry worksheets</b>	
	? _	
Data base : CDS Max MFN : 150	Micro-CDS/ISIS, Ver. 3.07(c) UNESCO 1993	Worksheet : CDS Format : CDS

**Figure 3.9 : System worksheets editing options**

Selecting option **M** from the menu, shown in Figure 3.8, provides a number of options for creating and editing system menus through the menu displayed in Figure 3.10 below.

Service ISISUTL	<b>System Menus Editing options</b>	Menu EXM3
	C - Create new system menu S - Select a new menu R - Restore current menu E - Edit currently selected menu D - Copy menu F - Delete menu X - Exit	
	? _	
Data base : CDS Max MFN : 150	Micro-CDS/ISIS, Ver. 3.07(c) UNESCO 1993	Worksheet : Format : CDS

**Figure 3.10 : System menus editing options**

## **H. - Advanced programming Facilities**

The services discussed so far enable users to perform most of the basic information storage and retrieval tasks. However, users' need change from time to time and they may need to perform specialized tasks which may not be accomplished using the existing functions using the seven services. For this purpose, Micro-CDS/ISIS provides users with a Pascal programming interface which enables them develop applications meeting their specific needs.

The Pascal interface is a library of procedures and functions which enables users to manipulate database records and perform most of the Micro-CDS/ISIS functions using the Pascal library.

### **3.3 Overview of WINISIS**

WINISIS is a generalized system designed for computerized management of structured textual databases(mainly), which uses the Microsoft Windows as an underlying operating system. Like Micro-CDS/ISIS, WINISIS enables users to build, maintain, manage, and retrieve records from structured textual databases.

Unlike Micro-CDS/ISIS - in which case the various utilities/functions provided by Micro-CDS/ISIS are grouped by eight services and each service's function/facility in turn are executed by using the service's menu - in WINISIS all the available facilities/functions are distributed over the various menus of the menu bar of WINISIS.

#### **A. The Main Window**

The main window of WINISIS, shown in Figure 3.11, is displayed each time WINISIS is started. It is normally empty, as shown below, unless a default database is selected through parameter 6 of the SYSPAR.PAR file, in which case the *database window* is also automatically opened within the main window.

The main components of the main window are:

1. The *window title bar*, at the top of the window, displays name of the application;
2. The *menu bar*, which provides access to all the WINISIS functions (some of these functions may also be activated using the various buttons of the *tool bar*);

3. The *tool bar*, located just under the menu bar, provides a quick mouse access to the most frequently used functions of WINISIS, such as opening a database, searching or printing etc. ; and
4. The *status bar*, located at the bottom of the main window, normally used to display help messages.

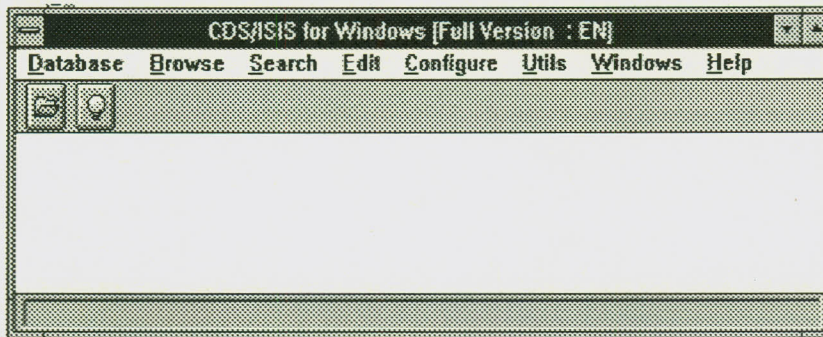


Figure 3.11 : The main window of WINISIS

Each menu contains different options which enables users to perform various functions provided by WINISIS.

### B. The Database Window

The database window is the one which will be automatically displayed whenever a database is opened for use (see Figure 3.12), displaying first record of the opened database, which is ready for browse.

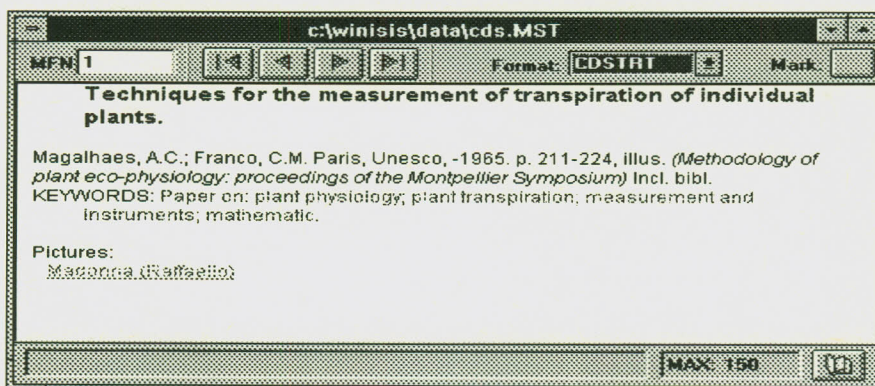


Figure 3.12 : Database window

### C. Database menu

This menu, shown in Figure 3.13, provides the following options:

*Open* - enables one to open different type of file.

There are three ways to open a database:

1. by selecting one of the five most recently opened databases from the list shown at the bottom of this menu;
2. by selecting this option; and
3. by clicking on the open database icon of the toolbar.

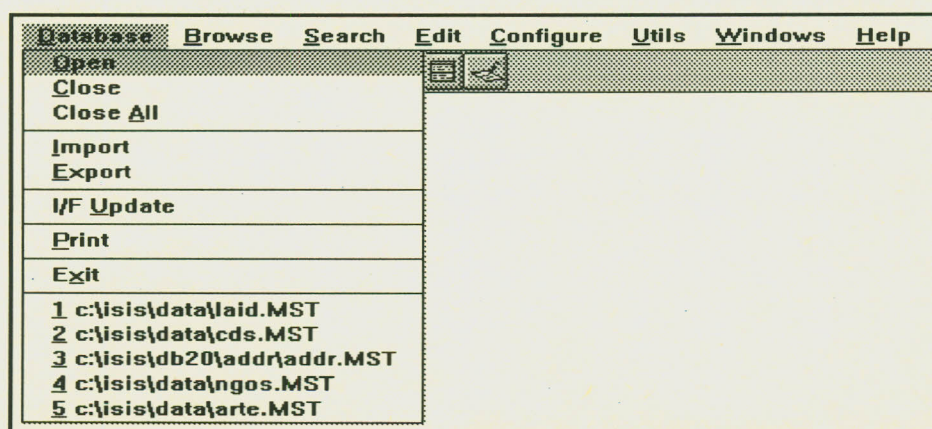


Figure 3.13 : Database menu

*Close* - closes currently selected database. All associated windows, such as a search window, will also be automatically closed.

*Close all* - closes all currently open databases.

*Import* - allows users to import data from external files recorded according to the ISO-2709 standard format for information interchange.

*Export* - allows one to extract a database or a portion thereof normally for transmitting it to other users. Users may also use this option to perform some reformatting of the records of

a database and then use the import function to store the reformatted data into the original or a different database.

**I/F Update** - allows users to perform various maintenance operations on an Inverted file of a database.

**Print** - allow users to print the output of a given query and/or to print a selected range of records. Users may sort the records by virtually any combination of fields and subfields. The field(s) by which the records are sorted may be used as headings in printing.

**Exit** - terminates WINISIS.

### C. Browse menu

This menu, shown in Figure 3.14, contains options that affect the contents of the database window and the data entry window.

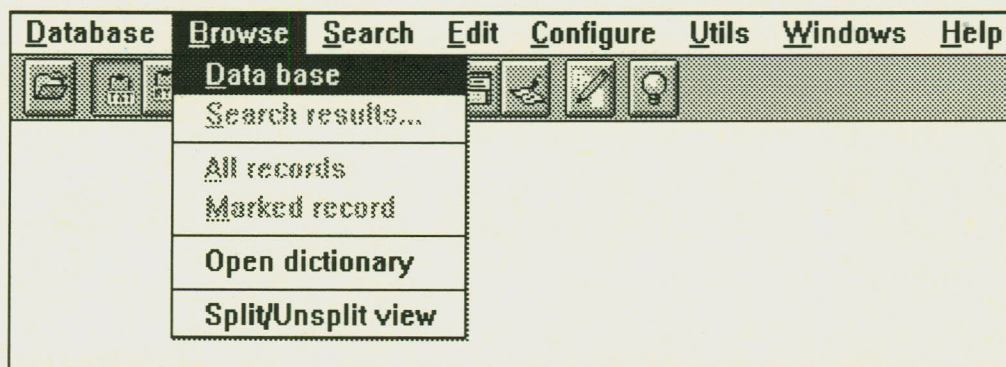


Figure 3.14: Browse menu

**Database** - sets the database in browse mode. In this mode the whole database may be browsed or edited independently from any search performed.

**Search results** - sets the database in display search mode. In this mode only the records retrieved by a selected search are displayed. When users choose this option, WINISIS

displays the list of searches performed so far, from which users may select the one desired.

Users may also use this option to simply view the list of searches submitted so far.

**All records** - is only active when users are displaying or editing search results. It allows users to view all the records retrieved, independently from their *marked* status.

**Marked records** - allows users to view or edit only marked records. This option is only active when users are displaying search results.

**Opens dictionary** - displays dictionary on the database window.

**Split/unsplit view**- split the database window into two and display the current record in the selected format and in ASCII format in each part.

#### D. Search menu

This menu, shown in Figure 3.15, contains the following options:

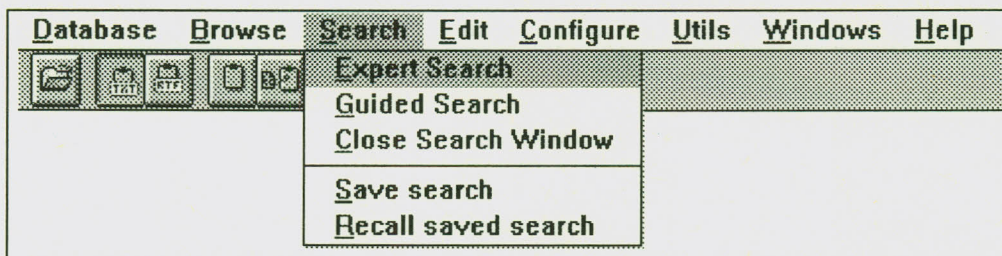


Figure 3.15 : Search menu

**Expert search** - opens the search interface designed for experts.

**Guided Search** - opens the guided search window, which provides a simplified search interface.

**Close Search Window** - closes an opened search window, if any, and activates the database window.

**Save search** - allows users to save results of a particular search, from a list of executed search expressions during a WINISIS session, for later processing.

**Recall saved search** - allows users to select a previously saved search expression. The search corresponding to the recalled save file is then automatically added to the list of current searches and the database window will display the first record retrieved. The recalled search becomes therefore the current search.

### E. Edit menu

This menu, shown in Figure 3.16, contains the following options:

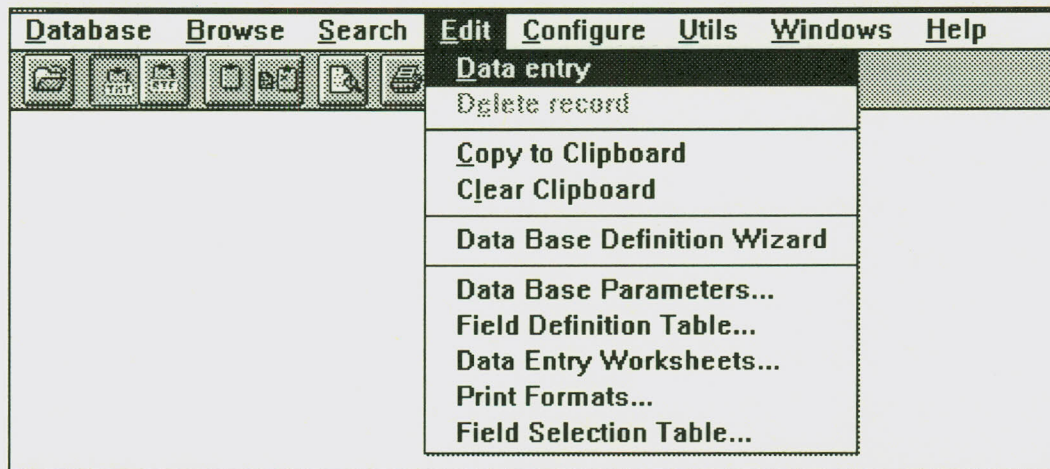


Figure 3.16: Search menu

**Data entry** - opens the *data entry window*, which allows users to add and/or modify database records.

**Delete record** - While in *data entry mode*, this option deletes the current record (i.e., the record currently displayed in the data entry window).

**Copy to clipboard** - copies current record to the Windows Clipboard. It is only operational when the database window is active. Records are copied in either ASCII or RTF format depending on the currently selected option on the tool bar.

**Clear clipboard** - clears the Windows Clipboard.

**Database definition wizard**- allows users to create new databases and/or modify database definition files of the active database.

**Database parameter file** - used to create/modify DBNPAR.PAR for the active database.

**Database definition table**- used to modify the FDT of the active database.

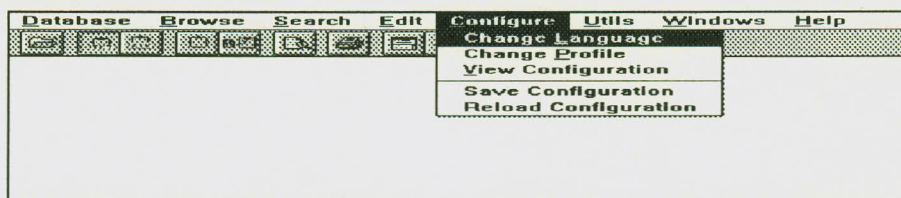
**Data entry worksheets** - used to create/modify data entry worksheets for the active database.

**Print formats** - used to create/modify display and/or print formats for the active database.

**Field select table** - used to modify the indexing FST of the active database, and other FSTs for sorting and reformatting purposes.

## F. Configure menu

This menu, shown in Figure 3.17, contains the following options:



**Figure 3.17 : Configure menu**

**Change Language** - allows users to change the language in which menus, windows, prompts and system messages will be displayed. Initially WINISIS will start in the language selected as default language in the SYSPAR.PAR file, by the new parameter, parameter 101.

**Change Profile** - allows users to dynamically change certain system parameters.

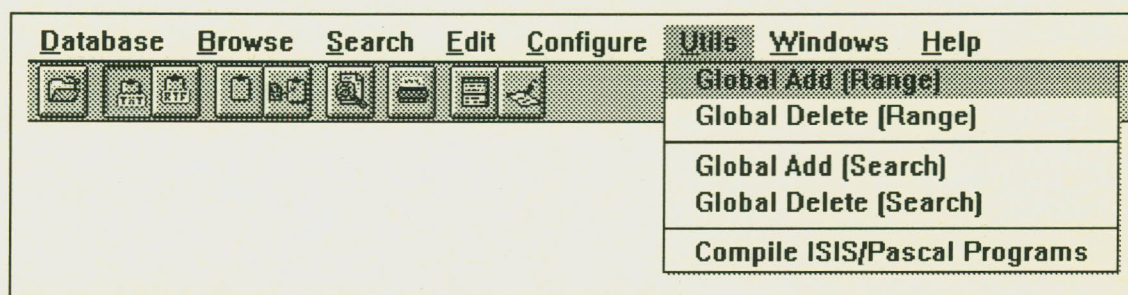
**View Configuration** - displays the current parameter settings.

**Save configuration** - allows users to permanently save the current size of the main window using parameter 122 of the SYSPAR.PAR.

**Reload configuration** - allows users to reload a saved size of the main window.

### G. Utilities menu

This menu, shown in Figure 3.18, contains the following options:



**Figure 3.18 : Utilities menu**

**Global Add (Range)** - allows users to add a field with a specified contents to a range of Master file records.

**Global Delete (Range)** - allows users to delete a field from a range of Master file records using the *Global Delete Dialog Box*.

**Global Add (Search)** - allows users to add a field with a specified contents to all records retrieved by a given search.

**Global Delete (Search)** - allows users to delete a field from all records retrieved by a given search.

**Compile ISIS/Pascal Programs** - allows users to compile a CDS/ISIS Pascal program.

## H. Windows menu

This menu, shown in Figure 3.19, contains the following options:

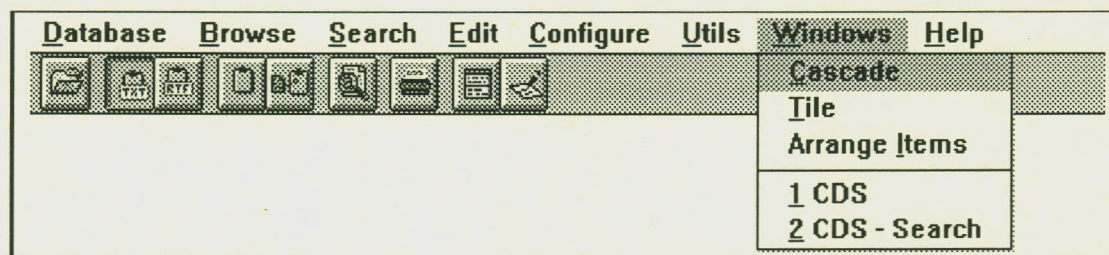


Figure 3.19: Windows menu

**Cascade** - arranges windows so that they overlap, starting in the upper-left corner of the main window.

**Tile** - arranges current windows side by side, so that each window is visible and no windows overlap.

**Arrange items** - arranges all icons in rows.

**1, 2, 3, . . .** - selects a particular window of a currently opened database.

## I. Help menu

This menu, shown in Figure 3.20, allows users to display the WINISIS Help file and contains the following options:

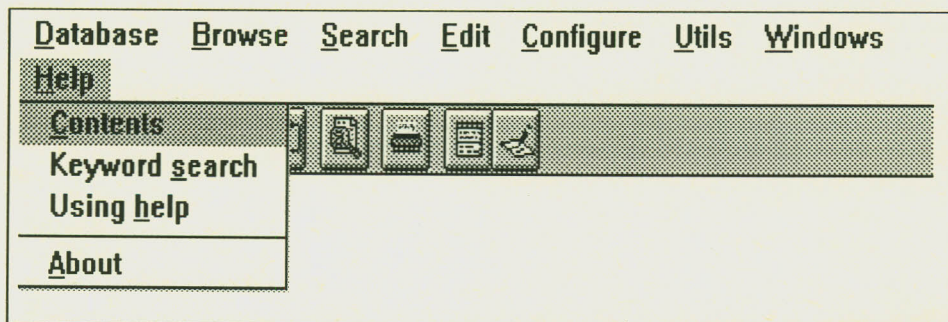


Figure 3.20 : Help menu

**Contents** - displays the WINISIS Help table of contents.

***Keyword search*** - allows users to select a particular help topic.

***Using Help*** - assists users in using the Windows Help facility.

***About*** - displays the WINISIS Copyright notice and the version number.

## CHAPTER IV

### COMPARATIVE STUDY OF MICRO-CDS/ISIS AND WINISIS

#### 4.1 Introduction

The comparative study of Micro-CDS/ISIS and WINISIS has used the following ten elements (different features and facilities generally available in a typical text retrieval packages and those facilities which are found in CDS/ISIS in particular): Database definition, Data entry/modification, Database indexing, Browsing and searching databases, Sorting and Printing facilities, Database modification, Import/Export facilities, additional utilities, Database security, and programming features. In addition, after exploring each package on these elements a discussion is given on each element to compare and contrast the two packages on that element. New facilities of WINISIS, which are totally unavailable in Micro-CDS/ISIS is given in a separate part and discussed by itself. In addition, a comparative summary of Micro-CDS/ISIS and WINISIS based on the different features and facilities of text retrieval software packages is given at the end.

#### 4.2 Database Definition

In both Micro-CDS/ISIS and WINISIS users are allowed to define databases according to their specific requirements. Both packages require that the information which will be stored in a database be structured into fields, and fields into records. This means that the structure of the database has to be designed before users are able to store and retrieve information. Hence, in both packages database definition is the first step in setting up a database system. Database definition specifies the content of the database and the rules that the different modules of Micro-CDS/ISIS and WINISIS may follow in manipulating the data.

## A. Database Definition in Micro-CDS/ISIS

In Micro-CDS/ISIS database definition involves the following four steps, which users must perform one after another in the order specified:

1. Creation of the *field definition table*;
2. Creation of the *data entry worksheet*;
3. Creation of the *display format*; and
4. Creation of the *field select table*.

The *field definition table* (FDT) defines the structure of the proposed database by defining the fields to be included in the database and their characteristics. The *data entry worksheet* is an on-screen form used to create and/or update database records. The *display format* is a set of instructions used for specifying how data elements of a database are to be displayed and/or printed. The *field select table* is used to define criteria for extracting terms from the Master file, which will be inverted and hence will constitute the dictionary of searchable terms.

### 1. Creation of FDT

Micro-CDS/ISIS users define structure of a proposed database using a table called *Field Definition Table* (FDT). The FDT define the characteristics of each proposed field in terms of the following parameters: **Tag**, **Name**, **Type**, **Length**, and optionally including **Repeatability**, and **Delimiter/Patterns**. The **Tag** is a unique numeric value (in the range 1 - 32767) identifying the field being defined; **field name** is descriptive text for the content of the field; **Field Type** defines the type of the field values: *Numeric*, *Alphabetic*, *Alphanumeric*, or *Pattern* type; **field length** is the expected length of field values; **Repeatability** indicates whether a given field can occur more than once in a given record or not; and **Delimiters/Pattern** – define set of subfields allowed in the field being defined, if needed, (and

applicable for all field types except *Pattern* type) or define the character by character description of the field being defined, if needed, (and applicable for *Pattern* type fields only).

## **2. Creation of Default Data Entry Worksheet**

While designing a new database in Micro-CDS/ISIS, once the task of creating the FDT is over, the system automatically leads the designer to the next step: *creating the default data entry worksheet*. For some or all fields defined in the FDT, the user has to create a worksheet which will be used for future data entry and modification purposes.

The Micro-CDS/ISIS *worksheet editor* is used for creating a data entry worksheet, which takes users through a sequence of eight prompts for each proposed field to be included in the worksheet:

1. field tag;
2. starting position of field name;
3. field name (optional but useful);
4. starting position of field value;
5. screen attribute - the way fields will be displayed (normal, underline, bold, etc.);
6. field length;
7. HELP message (optional text not exceeding two lines); and
8. default field value (optional).

When the system prompts for data concerning each proposed field, the system proposes: starting positions for field names and field values (by taking next available positions); field names as field title, and field length of a field as that defined in the FDT. But, users can change any of the proposed values. In Micro-CDS/ISIS a worksheet screen or 'page' can not hold more than 19 fields. Hence, if users have more than 19 fields and/or could not

accommodate their fields in a single 'page' they have to use a second, third etc., 'page'. And, in the process of defining each 'page' it is possible to edit a previously defined 'page' of the worksheet.

### **3. Creation of Display Format**

The display format specifies the way that records of a database will be displayed and/or printed. The way in which records of a database is displayed or printed is specified by users using the CDS/ISIS *formatting language*. Through the *formatting language* users may specify one or more specific data elements in the order they want displayed or printed, and optionally insert *constant text* of their choice, give different screen attributes to these elements as well as specify vertical and horizontal spacing requirements. Using the *formatting language*, users have to write a string of commands (separated by commas or space and which are recognized and executed sequentially by the system from left to right in the order they are listed) to specify their display/print requirements using the line editor which is activated for this purpose. In addition to specify a font and a size that will be initially used for displaying and editing print formats using SYSPAR.PAR file.

### **4. Creation of Field Select Table**

The *field select table* (FST) that users create during database definition is used to define a criteria for extracting terms which will be used to create the inverted file entries, and constitute the dictionary of searchable terms. To define criteria for extracting one or more elements from a Master file record, users have to define three parameters for each groups of terms in the order listed below (using the different lines of the FST for each group):

1. **ID** - the tag or *field identifier* of the field to be indexed;
2. **IT** - the *indexing technique*, which specify what type of terms to be extracted; and

3. **Data extraction format** - specifies how data is extracted from the records.

The three parameters are executed in the reverse order than they are listed above: **data extraction format** tells the system how data is to be extracted from the Master records for the purpose of indexing; **indexing technique** specify what type of elements to be extracted from the result of the data extraction format; and finally the **field identifier** is assigned to each element created during the indexing step (and which will be used as a qualifier during search).

In both Micro-CDS/ISIS and WINISIS there are five basic indexing techniques (numbered 0 - 4) that enable users to extract different type index terms:

0 - index entry will be generated for the whole field;

1 - index entry will be generated for each subfield;

2 - index entry will be generated for each element enclosed in an angular bracket (< >);

3 - index entry will be generated for each element enclosed in slashes (/ . . . /); and

4 - index entry will be generated for each individual word in a field/subfield.

In addition to the above five indexing techniques, one can use four more indexing techniques which allow users to specify a prefix for search terms extracted with indexing technique: 1, 2, 3, and 4 which are numbered 5, 6, 7, and 8 respectively. For instance:

1        8        'TI=',V24

will index each word of field 24 with TI= as a prefix. However, users have to specify the prefix with such words during searching from such prefixed index fields, and these terms will be displayed with their prefix in the terms dictionary.

In Micro-CDS/ISIS and WINISIS it is possible to index a given field more than once using different indexing techniques. When users complete their entries of an indexing FST, they

have the option to go to another *page* (if the current one is full), or to modify existing entries, or to save and exit the FST definition.

## B. Database Definition in WINISIS

In WINISIS database definition involves the following five steps, which are prompted to users to be performed one after another in the order specified:

1. Creation of DBNPAR.PAR definition
2. Creation of FDT;
3. Creation of data entry worksheet;
4. Creation of display format; and
5. Creation of FST.

In WINISIS also the purpose/reason for which the DBNPAR.PAR, FDT, worksheet, display format, and FST are defined are identical to that of Micro-CDS/ISIS.

### 1. Creation of DBNPAR.PAR

WINISIS prompts users to define database parameter file for the database being defined by providing the default path for each associated file of the database being defined (according to parameter 5 of SYSPAR.PAR, if any). Hence users can modify each path or accept the default. A sample DBNPAR.PAR definition is given in Figure 4.1.

**Data Base Definition - DBNPAR Settings**

Database Name: CDS.MST (NEW DATABASE)

Advanced Settings: [File Pathnames]

Multi-user environment

MST path:	c:\winisis\data\	L01 path:	c:\winisis\data\
XRF path:	c:\winisis\data\	L02 path:	c:\winisis\data\
CNT path:	c:\winisis\data\	IFP path:	c:\winisis\data\
NB1 path:	c:\winisis\data\	ANF path:	c:\winisis\data\
NB2 path:	c:\winisis\data\	Others:	c:\winisis\data\

Data Base Definition.  
Sets your component file pathnames or leave them empty for a default setting.

Cancel Help Next page

Figure 4.1 : A sample DBNPAR.PAR

## 2. Creation of FDT

In WINISIS, the FDT defines the characteristics of each proposed field in terms of the following: **Tag**, **Name**, **Type**, **Rep**, and/or **Pattern/Subfields**. Using the **Field Table** dialog window, see Figure 4.2, users specify the **tag** in the **tag** text box, **name** of the field in the **name** text box, select the field type among the three: ALPHANUM, ALPHA, NUMERIC from the **type** drop-down-list box, and optionally make the field repeatable by checking the **Rep** check box and/or specify the pattern or subfield identifier of the field in the **Pattern/Subfields** text box.

Tag	Name	Type	Rep	Pattern/Subfields
12	Conference main entry	ALPHANM	-	npdz
24	Title	ALPHANM	-	z
25	Edition	ALPHANM	-	
26	Imprint	ALPHANM	-	abc
30	Collation	ALPHANM	-	abc
44	Series	ALPHANM	R	vz
50	Notes	ALPHANM	-	
69	Keywords	ALPHANM	-	
70	Personal Authors	ALPHANM	R	
71	Corporate Bodies	ALPHANM	R	
72	Meetings	ALPHANM	R	mdz

**Step 1. Field Definition**  
Define the structure of your data base by entering tag fields, descriptors, types and patterns.

Figure 4.2 : A sample FDT

## 3. Creation of Data Entry/Modification Worksheet

In WINISIS data entry worksheet is defined by selecting the fields that users want to include in the worksheet from the **Fields** list box (see Figure 4.3). Fields will be ordered (and displayed during data entry and modification) according to their selection order. But, users can change the order of any field after selection. For each field included in the worksheet users can change the description/field title, give default value, associate HELP

messages, and/or field validation criteria using the **Worksheet Element** dialog window shown in Figure 4.4.

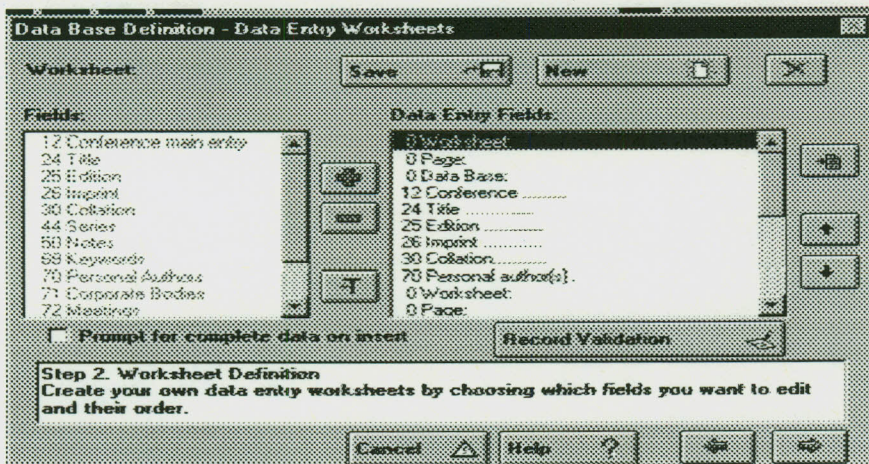


Figure 4.3 : A Sample worksheet definition

In addition to field validation criteria for any field, users can also give record validation criteria which is used during data entry and modification sessions to validate/check the content of the record when modifying existing records and/or adding new records. Each time a field is created and/or modified, WINISIS will execute the validation format. If the format produces any output, then it is displayed as an error message and users are forced to correct the field. Similarly, record validation format is executed before updating a record. If the validation format produces any output, then it will be displayed as an error message and users are forced to correct the record. A sample validation is given below:

**if size(v20)>7 then 'Field 20 may not be more than 7 characters long' fi**

**if p(v20) and p(v10) then 'Field 10 and 20 are mutually exclusive' fi**

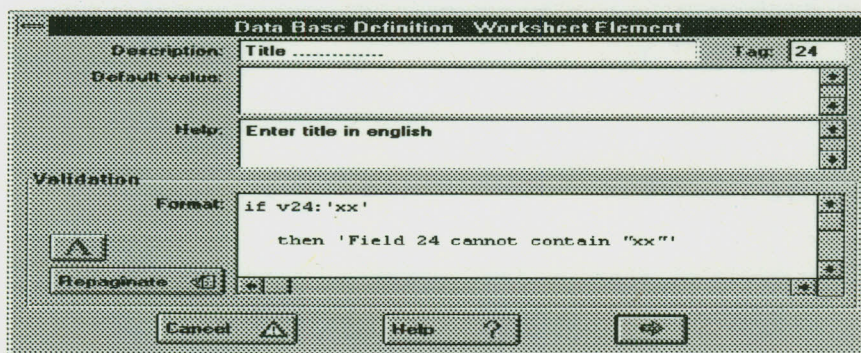


Figure 4.4 : Worksheet element dialog window

The above validation criteria validates fields 20 and the whole record. For field 20 an error message is displayed if the field is longer than 7 bytes. Finally, an error message is displayed if both fields 10 and 20 are present in the record which is about to be updated

#### 4. Creation of Display Format

Like Micro-CDS/ISIS, in WINISIS also users specify the display/print format by using the formatting language commands separated by commas. In addition, in WINISIS users can also specify font attributes (italic, bold, underline) for any element in the display/print format directly by highlighting the element and selecting any of the font attributes buttons (see Figure 4.5). It is also possible to specify hyper-text link.

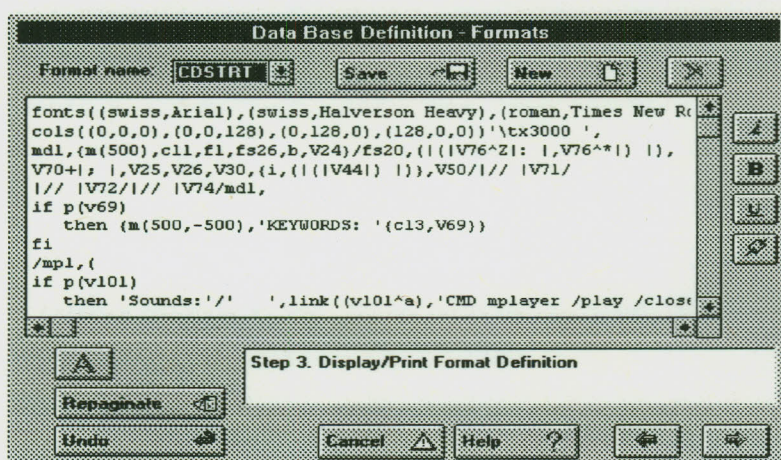


Figure 4.5 : A sample display format

#### 5. Creation of FST

In WINISIS users define their indexing FST by using *Extraction* table shown in Figure 4.6. To specify the FST users select the field tag from the *tag/name* drop-down-list box; select the indexing technique from the *technique* drop-down-list box; and specify the data extraction format in the *format* box. Each entry selected for indexing is displayed in the *entries* text box, and hence available for modification.

In designing these five files users can go back and forth to modify what they have defined (which is not possible in Micro-CDS/ISIS).

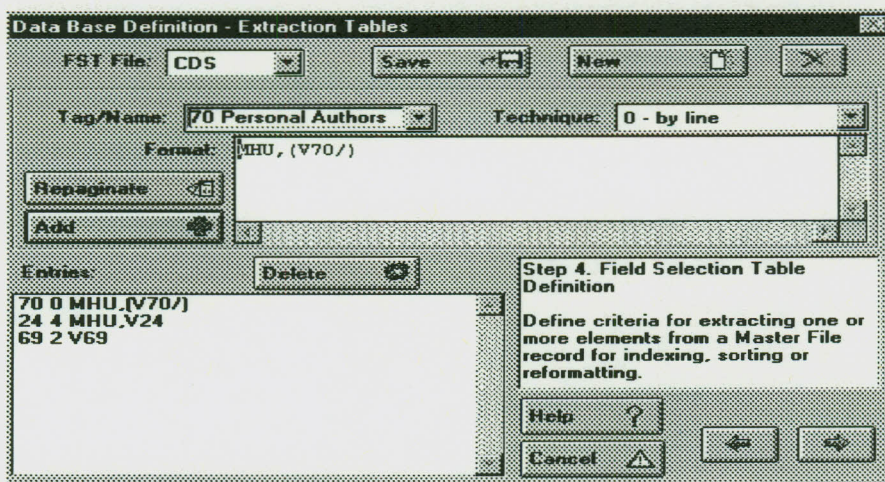


Figure 4.6 : A sample FST

## Discussion on Database Definition

### **A. DBNPAR.PAR**

In Micro-CDS/ISIS when a database is created the system places it automatically in the DATA subdirectory. If users have more than database there could be inconveniences. For instance, users would not be able to give the same name for two print formats of different databases. The way to avoid this problem is to place each database in its own directory. In Micro-CDS/ISIS users do this by creating a new directory and move all the files belonging to the database in to the new directory. And, in order to tell the system where users have put these database files, they must create a database parameter file in the DATA subdirectory outside Micro-CDS/ISIS using a text editor.

However, since WINISIS included the creation of DBNPAR.PAR as part of database definition procedure all these burdens of creating directories, moving files, and creating database parameter file itself are simplified.

In WINISIS, print worksheets (.WPR) will be saved in parameter 10 directory of DBNPAR.PAR. In addition, parameter 6 of SYSPAR.PAR can accept more than one database.

### **B. Creation of FDT**

The main difference in Micro-CDS/ISIS and WINISIS concerning FDT, and of course data entry worksheet is that, there is no field width definition. The system itself adjusts the width according to its content.

In WINISIS, it is possible to create new fields by modifying existing ones. Also, field types are not specified using a single letter representative, by clicking on a more descriptive terms.

In WINISIS there is no PATTERN field type, hence, no usage of fixed length field. In addition, for any field with specified field pattern can be repeatable.

In Micro-CDS/ISIS FDT, when a column can contain only a single letter (as with **field type** and **repeatability**) the cursor will jump straight on to the next column. Otherwise users have to press the ENTER key to move to the next column. Hence, there is inconsistency in ending table entries, and this can be disconcerting to beginners, especially.

### **C. Worksheet definition**

In Micro-CDS/ISIS for each field to be included in a data entry worksheet, at minimum one has to give the following data or accept what is suggested by the system:

tag, field name, starting positions for field name, and field value.

However, in WINISIS all the fields which are defined in the FDT will be displayed in a box, and what is expected from users is selecting wanted ones simply by clicking. Hence, what users do can be seen as a final touch to what the system already prepares. Creation of data entry worksheet is thus much more simplified.

In addition, unlike Micro-CDS/ISIS, there is no need of dividing a worksheet into *pages* when users have more than 19 fields or when they could not accommodate what they have in a single *page*. All fields will be found in a single data entry worksheet/window, and all what is expected from users, during data entry/modification, is just scrolling or pressing **ENTER** key to bring in to view the remaining fields.

During data entry worksheet preparation, it is possible to define field validation criteria for any field and/or one record validation criteria, which will be used to validate contents of field(s) and/or the whole record respectively during data entry/modification session. Also, in addition to validating data by field type, unlike Micro-CDS/ISIS, WINISIS does not allow entering of more than occurrences of a non-repeatable field, and it checks the validity of subfield delimiters.

Many worksheets can be created in a single session of creating a worksheet by copying and modifying the one just prepared, i.e., there is no need to go back to menus to make a copy, like Micro-CDS/ISIS.

#### **D. Display format**

In creating a display format in WINISIS, in addition to the formatting language commands of Micro-CDS/ISIS users can use new formatting commands and functions which enable them display their record elements in a much better way. For instance:

- users can access individual occurrences of a repeatable fields;
- it is possible access an alternate database which is different from the currently opened;
- apply character attributes (bold, underline, and italic) directly from the dialogue window;
- apply default or user defined non-proportional, proportional, and/or the mixture of the two.
- use different colors;
- apply new spacing commands for proportional fonts;
- include other formats saved by themselves using the *at* (@) sign; and/or
- write interactive formats where users will be able to perform some operations such as displaying images and/or playing music while displaying records, give a second simultaneous view of a record, open another database in a second window etc.

In addition, like data entry worksheet, it is possible to create so many display/print formats in a single session.

#### **E. FST Creation.**

In WINISIS, when defining an FST all the fields defined in the FDT of the current database are listed in a drop-down-list box with their tag. Similarly, all indexing techniques (numbers) with their description are listed in another drop-down-list box.

Hence, while designing an FST users simply select the field tag and the indexing techniques from the two boxes, they don't have to memorize the tag and the meaning of indexing techniques. Data extraction format is specified in a third box, and when all this three elements are specified users instruct the system to add the new entry. And the new entry will be listed in a fourth box for further modification

Finally, concerning database creation as a whole, in WINISIS it is possible to go back and forth and make any modifications at each files created in each step database definition, which is not possible in Micro-CDS/ISIS. In addition, there is direct HELP in the same windows used to create the files, and there is a broader context help.

### **4.3 Data Entry and Modification**

In both Micro-CDS/ISIS and WINISIS records may be created through direct keyboarding of the data using data entry worksheets, or importing/downloading from another database. Records may also be modified using user-written Pascal program that allows specialized kind of record modification.

#### **A. In Micro-CDS/ISIS**

##### **1. Creating New Records.**

In Micro-CDS/ISIS one way in which data entry is done is using a previously defined data entry worksheets. A Micro-CDS/ISIS worksheet is similar to a printed form with spaces to be filled in, which is displayed on the screen. Users use the keyboard to fill it in as they would in a paper form.

Each field of the worksheet is entered or modified by means of the field editor. But, in addition, Micro-CDS/ISIS provides screen editing facilities, such as moving from field to field in a *page*, or switch back and forth between *pages*.

In Micro-CDS/ISIS users enter new records field by field; if users need to enter several values in repeatable fields individual values must be separated by the field separator with no space before and after; and if a field is to be divided into fields, users must enter the data with no space before and after the subfield markers.

If a field type is alphabetic, numeric or pattern then its content will be checked against its type and will be rejected if it does not match the type.

During data entry there are some special keys users can use:

**F1** - to display a pre-defined HELP of the current field (and if there is none, Micro-CDS/ISIS displays the field type of the current field);

**F2** - to delete what has been written in a field and go to the next field;

**F3** - to mark a block of text in a field;

**F4** - to cut a marked block of text;

**F5** - to paste a block of text in another field or record; and

**F6** - To delete data from the current cursor position till the end of the field.

During data entry, it is possible to use the following movement keys:

**Arrow keys** - to go character by character within a field;

**Home key** - to go to the beginning of a field;

**End key** - to go to end of a field;

**Ctrl + Arrow keys**- to go word by word to the right or left in a field;

**ENTER key** - to go to a next field;

**TAB key** - to go to a previous field;

**ESC key** - to cancel any changes users have made to the current field and move to the next field; and

**PgDn key** - to move straight to the bottom of the page, so that users can choose from the options there.

In Micro-CDS/ISIS, always, to create new records users have to press **N** after finishing one.

## 2. Record Editing

In Micro-CDS/ISIS users can either edit a record as soon as they have completed entry of each worksheet page of a new record or they can come back for editing some time later and make changes. To edit previously entered records users have two options: select option **E-Edit record/range** -(to edit records by specifying a range of MFN) or select option **R-Edit last search result** (to edit records which are result of a last search) from **Data Entry Services** menu.

The first record of the specified range or the search result will be displayed on a selected worksheet for edition. Editing record(s) is then just like editing a record during data entry, and the same function keys explained above and keyboard movement keys can be used. During data entry, users can also define default value for a field.

In Micro-CDS/ISIS one way to make changes to all records in a database is through the export and import functions. By exporting using *export FST* and then importing using *import FST* to the same database one can make such type of global change such as reorganizing subfields of a field, for instance. Using this technique users can move data automatically from one field to another, eliminate a field content, add new occurrences to a field in all records, change field tag of a field etc. The other way of making such type of global changes is to use the UNESCO's GLOB program which can be used as a format exit from the **Data Entry Services** menu.

## 3. Deleting records - logically and physically

In Micro-CDS/ISIS a record is deleted (logically) by calling it up for editing and selecting the **Delete** option. A deleted record will then not be retrieved when users search a database and its MFN will not be reused when new records are created. A logically deleted record

can be undeleted by calling it for editing (unless the Master file is reorganized after logically deleting the record). But, calling a physically deleted record has the same effect pressing **N** - to create new records.

In Micro-CDS/ISIS, if users want to get rid of the record altogether and 'close up' the gap in master file numbers, the only way is to export the whole database (which will exclude deleted records) and then import it again. (Exporting and importing are discussed in section 4.8.)

## **B. In WINISIS**

In WINISIS new data entry and modification of existing records is performed first by selecting option **Data Entry** from the **Edit** menu. When initially opened, Data Entry window/worksheet displays the current record (i.e., the one which is currently displayed in the corresponding *Database window*, which can also be result of a search, or a marked record). Hence, before records can be modified users can perform a search and, further *mark* selected records for editing.

Like Micro-CDS/ISIS fields of the record are displayed in the lower part of the window according to a selected (or default) data entry worksheet. But, it is possible to change data entry worksheets during data entry/modification. Moving from one field to another is possible by pressing the **ENTER** and **SHIFT+ENTER** keys or the **PgUp** and **PgDn** keyboard keys. It is also possible to move to any field with a single mouse click and use the scroll bar to display hidden fields.

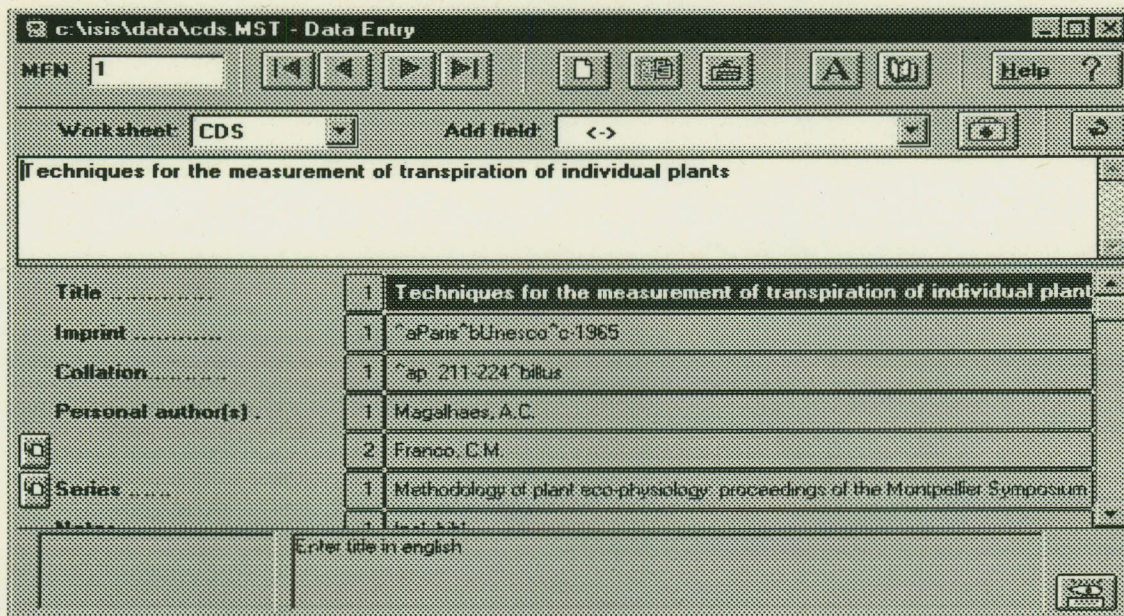


Figure 4.7 : A sample data entry/modification in WINISIS

### 1. Creating new occurrences

New records can be created from scratch or by copying existing records and modifying there after. In WINISIS adding occurrences is performed by choosing a field from the **Add field** combo box. Users can also add an occurrence to a repeatable field by pushing the *repeatable button* that is shown just near each repeatable field label in the window(see figure 4.7 ). In WINISIS each occurrences of a repeatable field are displayed and modified as a separate line of data, so there is no need of entering occurrence separators.

WINISIS provides a facility of linking file(s) and field(s). Using the Windows File Manager and applying the drag-and-drop technique or by directly specifying the name(s) of the file(s) in the *field edit* box users may easily establish links between fields and external files such as pictures, movies, and sounds.

### 2. Editing a field

Records to be edited can be selected using the **Browse** toolbar : by specifying its MFN and/or using the *Arrow* buttons. To edit a particular field users must first click on the

corresponding *field* button (the long box where field contents are displayed). Its contents will then be brought up in the *field edit* box. In editing a field users can use all standard Windows edit function keys (such as CTRL+X for cut, CTRL+C for copy, CTRL+V for paste, etc.) and, in addition, the three following special keys:

- **ENTER** to update the field and select the next field.
- **F2** to delete field
- **ESC** to ignore all changes (field will redisplay as it was before the edit started).

In WINISIS users can change the size of *field edit* box to increase or decrease the amount of information displayed from a field. It is also possible to change the font used to display a field in the *field edit* box.

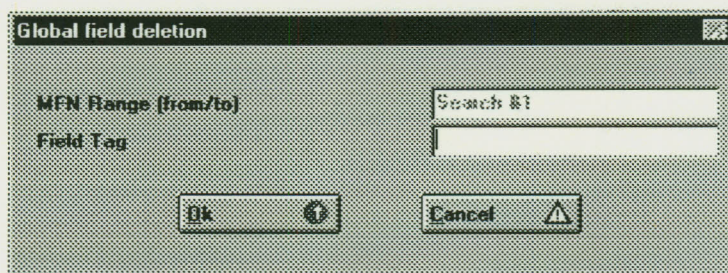
During data entry and modification it is possible to add a term to a field from terms dictionary by double-clicking the term or using Windows's *drag-and-drop* technique, and users are allowed to save a new or modified records to a disk. If a record is not saved to a disk, users can cancel all modifications they made to the current record and restoring the original content from the disk.

### **3. Deleting and Undeleting records**

Being in the data entry window, selecting the Delete Record option from the Edit menu enables users to logically delete the current record. The status bar will display that the record has become deleted. A logically deleted record (i.e., a record marked for physical deletion) can be undelete using the **RED CROSS** button on the toolbar, which immediately stores the record on the disk without asking for any confirmation.

#### 4. Global Editions

In WINISIS, in addition to the possibility of adding new occurrence(s) to a field or modifying existing occurrence(s) of a field record by a record, there is a facility of adding new occurrence(s) to or deleting existing occurrence(s) from a field in a range of records or in all those records retrieved from a previously submitted search. Using the **Global Add (Range)**, **Global Delete (Range)**, **Global Add (Search)**, and **Global Delete (Range)** options of the **Utilities** menu users can (see Figure 4.8):



*Figure 4.8 : Global field deletion dialog box*

- add a field with a specified contents to a range of Master file records,
- delete a field from a range of Master file records,
- add a field with a specified contents to all the records retrieved by a given search, and
- delete a field from all the records retrieved by a given search.

#### Discussion on Data Entry and Modification

##### Record Creation

In WINISIS, to create new records users can copy a record AND create a new one by modification or create it from scratch. The data entry window is not divided into *pages*, all the selected fields of a data entry window are observable by scrolling, if they are more than a screen. When entering new occurrences of a repeatable field, each occurrence will be listed in different lines, and there is no need of entering the repeatable marker. Fields can be linked to external files such as pictures, movies, sound etc., for usage in display/print format

## **Record Editing**

In WINISIS, in addition to editing records by range of MFN, it is possible to edit fields which are results of a search or those records marked from the result. But, it is possible to change this default setting using the new parameter (parameter 123) of SYSPAR.PAR. Also, it is possible to select other worksheets during data entry and modification; users can also change the font by which fields will appear for edition in the field edit box.

For editing purpose, data entry worksheets/windows, displays only those records which have entries. Also, it is possible to resize the data entry window as a whole, and/or the size of the *field edit* box. Hence, the display area can be optimized for edition. During data entry/modification users can add terms from the dictionary, and they can access the print facility of WINISIS. While editing fields, what is being edited is seen in the *field edit* box, hence it will not replace existing field content until users press the ENTER key or click on another field, i.e., editing a field does not mean replacing existing contents of a field as Micro-CDS/ISIS. Hence, this facility enables users to make comparisons of new and old versions of their fields before confirming changes.

## **Global Editing**

In WINISIS it is possible to add new values to all fields in a range of records or to all records retrieved by a previous search. And, it is possible to delete in the same way. However, the global addition facility adds only new occurrences. It can not add new values/occurrences between existing ones. It is not possible to remove an occurrence or part of a field selectively. So, in this respect the GLOB exit program is much better.

#### 4.4 Database Indexing

In both Micro-CDS/ISIS and WINISIS after data has been entered into the database Inverted files should be generated so that users can take advantage of the powerful search facility. Inverting the whole database may take longer period of time for relatively large databases. But, once an inverted file is created users can use the 'Update inverted file' facility to invert only the newly created, or modified records. In both Micro-CDS/ISIS and WINISIS inverted file generation process consists of three steps:

1. Creation of the *Link files*;
2. Sorting the *Link files*; and
3. Loading the *Sorted Link files* into the Inverted file.

The *link files* contain all the information necessary to generate the Inverted file and are created according to the **Inverted file FST** defined for the database. Link files contain one entry for each couple search term/posting to be stored on the Inverted file. The second step sorts the *Link files* into alphabetical order. The purpose of this sort is to collect together all postings related to a given search term, in preparation of the last step which loads the sorted Link files into the Inverted file.

These three steps may also be executed one at a time by using the options **Create link files**, **Sort link files** and **Load link files** in this order. Hence, if users have large databases, where the full Inverted file generation may require a relatively long time and considerable disk space, they can perform the full inverted file generation step by step as described above.

The **Inverted File Maintenance window** of WINISIS is shown in Figure 4.9.

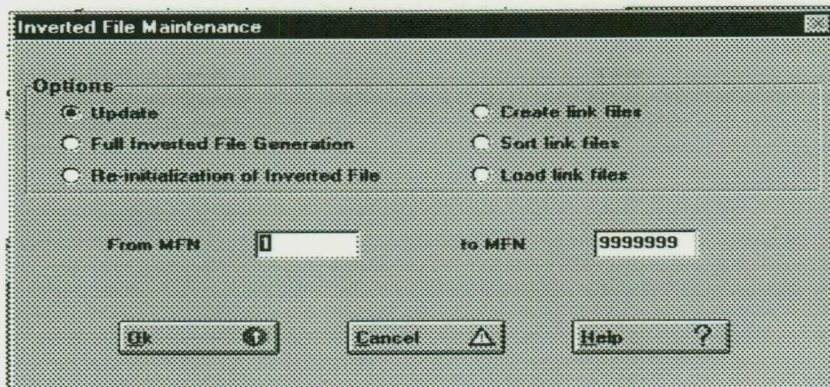


Figure 4.9 : Inverted file maintenance dialog box

## Discussion on Inverted File

When new records are created and existing records are modified. Micro-CDS/ISIS informs users to update the Inverted file, whether or the modified/new fields are indexed. But, in WINISIS users can make updating the Inverted file automatically using the new parameter, parameter 130. Otherwise WINISIS does not inform users to update the Inverted file whenever they make change. The Backup option of Micro-CDS/ISIS converts the inverted file into a link file, by excluding term with no postings, for subsequent reload. Hence, when restored, the inverted file will be reorganized, compact and reclaim any unused space which might have been produced by updates since last updates.

Similarly, the *Re-initialize Inverted file* feature of WINISIS erases the inverted file content and reallocates all associated files to a minimum size. When the inverted file is regenerated (reloaded from the sorted link files) it will be compact and reclaims any unused space.

However, *Inverted file backup*, *Dump inverted file*, and *Print search terms dictionary* facilities are not found in WINISIS.

### **Stopword lists**

In both Micro-CDS/ISIS and WINISIS if users are indexing a field by separate words (indexing technique 4) they can prevent common 'no information' words such as *and* or *the*

from being indexed by creating an ASCII file called *Stopword file* that contains these type of words. In both Micro-CDS/ISIS and WINISIS there can only be one Stopword file/list for a given database, not different lists for different fields or part of a database. The Stopword file needs to be setup outside these packages using text editors or word processors. It must have the same name as the database with extension STW. It must reside in the same directory as FDT file. The Stopword file must contain one Stopword on each line with no preceding spaces, and the words must be in capital letters. Whenever both systems generate an index file they look for such file in the FDT subdirectory and if present they will utilize it.

## **4.5 Database Browse and Search**

### **4.5.1 Browsing A Database**

#### **A. In Micro-CDS/ISIS**

In Micro-CDS/ISIS, users can display records of the Master file using a pre-defined display format by choosing option **B-Browse master file** from **Information Retrieval Services** menu. Users can start browsing from any record.

In standard Micro-CDS/ISIS, as many records will be displayed as will fit on the screen and pressing the **ENTER** key displays more of the records. Any record which is not complete at the bottom of the screen will be displayed at the top of the next one. Users can terminate browsing before they reach the end of the database.

#### **B. In WINISIS**

In WINISIS opening a database is always accompanied by opening a window (called database window) which displays the first record of the database in it. Hence, users can directly use the navigation toolbar's *goto* and/or *direction buttons* (see Figure 4.10) to

browse records. In WINISIS users also have the facility of browsing records of a previously executed search, and browsing only marked records among them.

During or before browsing users have the options of splitting the database windows and display records in two different formats. The first window is used to display the records, but the other window can be used to display records in ASCII format or to display the print format itself (PFT). Hence, users can modify their PFTs and see the effect immediately in the other part of the database window (see Figure 4.10).

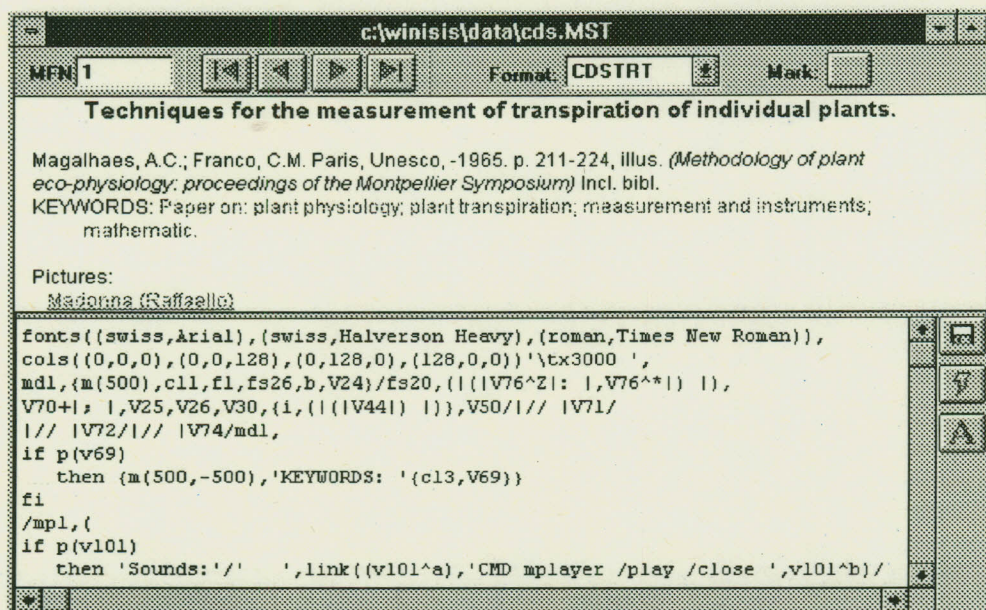


Figure 4.10 : A sample splitted database window

Unlike Micro-CDS/ISIS, in which case printing of records is done only through a single remote outlet, in WINISIS users can access the print facility during browse, data entry, and search. It is also possible to open more than one database and browse them side by side, of course one at time.

### Discussion on Browse

In Micro-CDS/ISIS, users can browse only Master file records starting from a certain MFN.

In WINISIS, first of all, whenever a database is opened for use, its first record is displayed

in database window, hence ready for browse. It is also possible to browse results of a search or only marked records of a search result. In WINISIS it is also possible to browse records using two formats at the same screen by splitting the database window into two. In addition, while browsing, it is possible to copy records in to the Clipboard either in ASCII or RTF format so as to paste them in other applications for further processing.

It is also possible to change display format while browsing (which is possible in Micro-CDS/ISIS only before browse) and hence view different records in different formats.

#### **4.5.2 Searching a Database**

##### **The Search Language Micro-CDS/ISIS and WINISIS**

The search language of Micro-CDS/ISIS and WINISIS is based on *Boolean algebra*, which provides a convenient way of expressing logical operations between *classes*. Each search term associated with a given record can be viewed as representing the *class* of all those records associated with that term. Thus by expressing logical operations between search terms users can define precisely the class of records to be retrieved in response to their needs.

The most efficient way to carry out a search is through the Inverted file. Users may only do this, however, if their search requirements are within the scope of the dictionary of search terms defined for a given database. However, there will still be cases where users are not able to formulate a search based on the Inverted file. For these cases, both Micro-CDS/ISIS and WINISIS provides an alternate search method, called *free text searching*, which allows users to satisfy *any search requirement* independently from the contents of the Inverted file. Because this second method is much less efficient, it is normally used in

conjunction with an Inverted file search to refine search strategy, but may, if necessary be used independently, even though it may take a longer time.

### **Search Expressions**

In both Micro-CDS/ISIS and WINISIS a search is performed by formulating a *search expression*. Eventhough the way a search expression is formulated is different in Micro-CDS/ISIS and WINISIS, it is a sentence composed of search terms, search operators, previously executed search results, and/or pairs of parentheses.

The terms used in formulating a search expression can be *precise terms* (any searchable elements defined for a given database), *right truncated terms* (root of terms), and ANY term (a collective term standing for a pre-defined set of search terms).

In both Micro-CDS/ISIS and WINISIS users can combine two or more search terms by using search operators, which indicate the intended relationship between the search terms. The three basic *operators* in both Micro-CDS/ISIS and WINISIS are: the logical **AND**, **OR**, and **NOT**.

In addition to the above three basic operators in both Micro-CDS/ISIS and WINISIS there are additional *field level and proximity operators*, which are more restrictive type of the logical **AND** operator, and are particularly useful for natural language searching:

1 **G** - to specify that the terms must be in the same field. All occurrences of a repeatable field are treated as a single entity;

2 **F** - to specify that terms must come from the same field or individual occurrences of a repeatable field;

- 3□. - to specify that terms are no more 'n' words apart in the same field or individual occurrences of a repeatable field, where 'n' is the number of periods plus one; and
- 4□\$ - to specify that the terms must occur exactly 'n' words apart in the same field or individual occurrences of a repeatable field, where 'n' is the number of dollar sign(s) plus one.

In response to a search expression both packages assign a unique number called *set number* to the search expression users have just entered (provided it contains no syntactical errors), and display the number of *postings* for each term in the expression and the number of records retrieved. If the search expression contains an ANY term and/or a right-truncated search term all the postings of the individual search terms as well as the aggregate posting value for the aggregate ANY term and/or the truncated term will be displayed. If users use a term which is not a valid search term, both Micro-CDS/ISIS and WINISIS will set its posting value to zero and flag the term with the message:

**\*\* NOT FOUND \*\***

In addition to assigning a unique number to search expression both Micro-CDS/ISIS and WINISIS also save the records matching each search expression. In later search expressions users may refer to previously entered ones by simply using the set number assigned preceded by a number sign (#). This facility allows users to develop their search strategy step by step. Breaking the query into elements also allows users not only to check at each step the number of records retrieved, but also to display the retrieved records themselves, so that they may verify at any time the logical validity of the search formulation in terms of the relevance of the records retrieved. A reference to a previously

defined search expression is called a *backward reference*, and it is possible to apply a qualifier to a backward reference. For example,

#1/(64)

which selects, from the records retrieved by search expression 1, only those in which the search expression was verified on field 64.

Users may form complex search expressions by combining two or more search terms with the search operators described above. As in normal algebra, users may use parentheses to alter the order of evaluation.

In forming search expressions the following rules hold both in Micro-CDS/ISIS and WINISIS.

- no two logical operators may be adjacent to each other, except for repeats of " . " and "\$ " (which may not, however, be mixed together);
- the parentheses used must be balanced, i.e. the number of open parentheses must be equal to the number of closed ones and each open parenthesis must have a matching closed one; and
- if a search term contains parentheses or any one of the search operators (\* + (G) (F) . \$ ^) or begins with a number sign (#) users must enclose it in double quotation marks (") in order to avoid a possible ambiguity.

### **Free-Text Search**

**Free text search** allows users to specify search requirements on fields which have not been inverted and/or to specify conditions which users not otherwise be able to indicate through the search expressions described above, such as comparison of fields or comparison numerical values of fields.

In Micro-CDS/ISIS and WINISIS *free text search expression* is declared by beginning a search expression with a question mark (?). Free text enables users of Micro-CDS/ISIS and WINISIS to perform the following operations (which are not possible using the normal search method):

1. select only those records containing a particular string of characters in a given field.

Example:-

**?V200: 'Cambridge University Press'**

2. select only those records satisfying certain arithmetic conditions using arithmetic operators (<, >, >=, <=, and =). Example:

**?val (V260^C)>=1990**

3. select only those records satisfying certain date conditions. Example:

**?v10='1994-02-01'**

4. select those records in which a given field is absent or present. Example:

**?a (V94)**

**?P (V94)**

During free-text search and before displaying search results, Micro-CDS/ISIS displays the following information:

- the total number of records available in the database;
- the total number of records searched;
- the hits; and
- the percentage of records scanned.

Similarly in WINISIS it is possible to display each record satisfying a free-text search expression as the search proceeds, and finally the hit. But, this facility is optional and is specified using parameter 103 of the SYSPAR.PAR.

### **A. In Micro-CDS/ISIS**

In Micro-CDS/ISIS a search expression is formulated by selecting option **S - Search formulation** or using option **D - Display dictionary terms** from **Information Retrieval Services** menu.

After formulating their search expression (whether it is a free-text search or search through the Inverted file) and executing it, users have the option of displaying the result immediately or selecting any valid options of the **Information Retrieval Services** menu.

During search formulation if users do not know the exact terms indexed in their database, they can request for terms dictionary to select one or more terms. It is possible to see further down the file by pressing the *PgDn* key. Users can go to a term which they want to select (using **Arrow Keys**, **Home**, or **End** keys) and then pressing **S**. To select other terms users should take the cursor/pointer to the next wanted term and press any valid operator to combine it with the rest. Finally pressing the **ENTER** key execute the search.

After executing a search expression, and getting some hits, users can display the records immediately or they can return to the **Information Retrieval Services** menu and choose the display option of it. When displaying search results, Micro-CDS/ISIS will display as many records as will fit in one screen using any selected format. Then, users can display more records, and any record which is not complete at the bottom of the screen will be displayed again at the top of the next one. The order of records will be the same as in the database, i.e., it is not possible to sort records just before they are displayed on the screen.

But, by exporting using *export FST* and then importing in to the same database records can be sorted. In addition, it is possible to display records as they are stored in the system, preceded by the numeric tag for each field (for checking records and investigating problems, for instance).

Results of a search can be saved for other purposes, for instance, for printing or perhaps exporting to another user. What is saved is actually a list of the master file numbers of the records retrieved. If users delete one of these records it will not be accessible through the save file.

## **B. In WINISIS**

In WINISIS, searching a database can be performed in two ways: **Guided search** and **Experts search**, each designed for different ways of formulating search expressions.

### **Guided Search**

In guided search a search, expression is formulated using a simplified search interface (see Figure 4.11) which is designed for inexperienced users of WINISIS. Four **text boxes** (search elements) are provided where users may type a search term or select any one from the terms dictionary in each box. Each search element may be associated with a particular field by selecting the appropriate one from the corresponding **Searchable fields** list. Search elements may be linked with one of the three search operators **AND**, **OR**, or **NOT**, by selecting the appropriate one from the corresponding **Operators** drop-down-list box.

When one or more previous searches have been made, users may select one from the **Previous search** list and connect it with a new set of search elements. When a search expression is executed the actual text of the search statement generated by WINISIS will

then be displayed in the **search expression** box. When the search is completed its number will be shown in the **Search Number** box and the total number of records retrieved in the **Number of Hits** box. Users may display the result of the search or may return to the database window without executing the search or clear the current search expression and specify a new one.

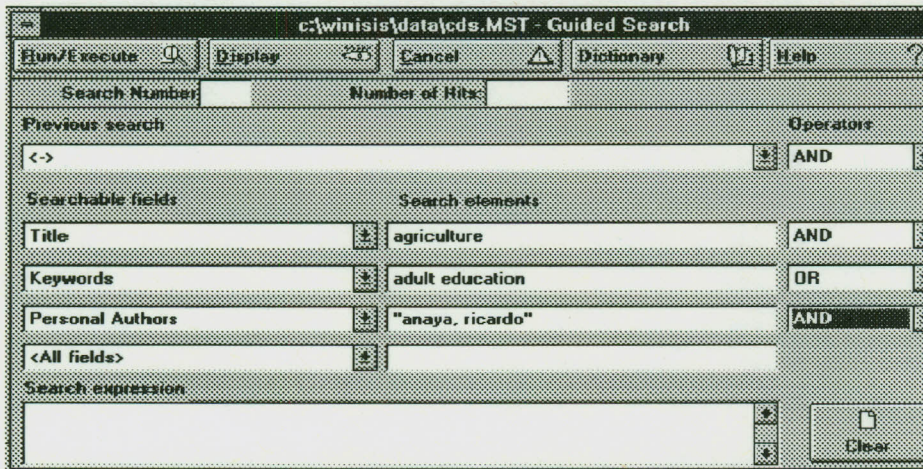


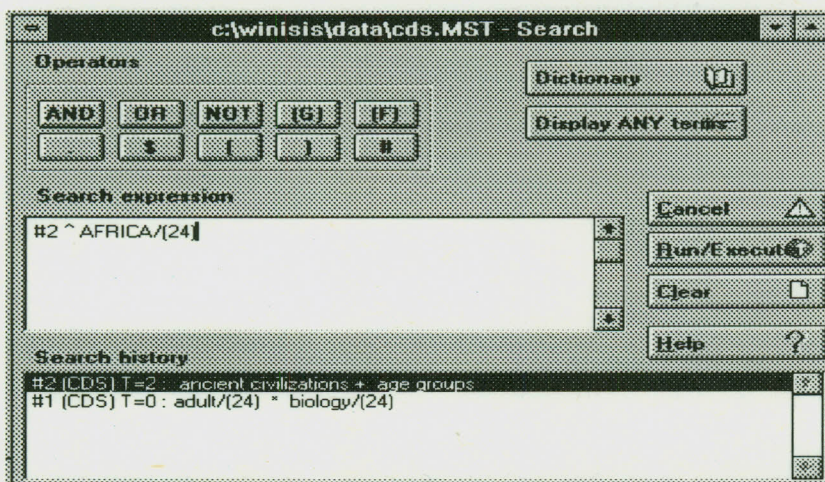
Figure 4.11 : Guided search window

## Expert Search

In WINISIS **Expert Search** is performed using interface (see Figure 4.12) that is designed for more experienced users. In the **Search expression** dialog box an experienced user may formulate his/her search expression by typing term(s) and/or selecting them from terms dictionary. Search operators, parentheses, set number, and/or pre-defined ANY terms can also be selected directly from the dialog window.

The **Search history** window contains a list of previously executed search expressions. For each expression it gives the set number, the database name, the number of hits, and the search expression. A new element is added to this list each time a search expression is executed. Double clicking on an element of this list will display the corresponding results in the database window. A single click will normally insert the set number in the **Search**

**expression** box. However if the selected element corresponds to a search done on a different database the search expression itself is copied. Users may also copy all or part a search expression (rather than the set number) to the **Search expression** box by dragging the element from the **Search history** list and dropping it in the Search expression box.



*Figure 4.12 : Expert search window*

During search, the **Searchable fields** combo box (at the top of the **Dictionary** window) allows users to select the field to be searched when they are in expert search mode and to select search terms from an indexed field when they are in Guided Search mode. Typing one or more characters in the **Term** box will position the dictionary window accordingly (e.g. typing F will show the dictionary terms starting with F). The up and down buttons will page forward and backward. Note, however, that backward paging will go no further than the starting term.

In addition, using the new parameter of the SYSPAR.PAR (parameter 105), users can instruct to display or not to display results of a search expression automatically.

## Discussion on Search

The search interface of Micro-CDS/ISIS is the same for both novice users or expert users. In Micro-CDS/ISIS, users have to formulate their search expressions by themselves without almost any assistance from the system. Of course, users have access to the terms dictionary and they can reuse previously executed search results while they are formulating their search expressions.

But, even if the terms dictionary can be displayed there is no guidelines or HELP on how to select terms and how to go back and forth between the dictionary and the menu/their search expression.

In addition, the screen/interface used to formulate a search expression is simply a blank screen, without any relevant information on how to formulate a search expression and what keys to use. For instance, there is no guidelines or help which assist users:

- how to select search terms from the dictionary;
- type of operators which can be used and/or their keyboard key characters;
- how to use a previously executed search expressions;
- how to clean a search expression and/or formulate others; and/or
- how to execute search and how to display results.

Even after executing a search and getting some hits the system simply shows an arrow sign:

---->

which is neither indicative of next expected key to be pressed nor intuitive for trial and error. Also, to edit and submit a previously executed search expression users have to remember the set number or should see it before editing it.

These problems however have been largely solved by the search interface HEURISKO which has been developed by UNESCO and now comes with Micro-CDS/ISIS, where both novice and expert mode of search are available.

In WINISIS, first of all, the search interface is divided into two: for novice users and expert users. In guided search window, which provides limited search facilities, users formulate their search expressions using four boxes for specifying search term. From that window users can select a field from which to search a term and select an operator (among the three basic ones) to connect the term with the next term, if any. That is, in guided search a search expression is formulated by specifying parts of it step by step.

Since previously executed search expressions are available in a drop-down-list box, it is possible to connect any one of them with what is formulated or re-execute it. But, once a search is executed it is possible to modify it and re-execute, for instance, by adding operators, search term, and/or search expressions.

In expert search window, which provides more search facilities than guided search, a search is formulated by writing and/or selecting search terms, operators, and/or search expressions in a text box. But, unlike guided search, here all operators, type of terms, and all list of search expressions executed during a session of WINISIS are available for selection. During formulating a search expression, it is also possible to copy part of a previously search expression and/or include it in what is being formulated.

In WINISIS, re-execution of a previously search expression is just selecting it from a list, whereas in Micro-CDS/ISIS it is by specifying its set number.

The display option of Micro-CDS/ISIS displays records of a last search, hence to re-execute and display records of other search it has to be current first of all. In WINISIS users can re-execute any search and redisplay the records, even from a saved one, which in done is Micro-CDS/ISIS only for further processing, like printing, for instance. But, in WINISIS it is possible to recall any saved search and re-execute it and/or modify and save it again.

In addition, in Micro-CDS/ISIS users are allowed to save only a last search, otherwise it has to be made current. In WINISIS saving a search result only needs selecting it from a list of executed search expressions during a WINISIS session. In WINISIS, when displaying result of a search, it is possible to select any pre-defined display format or select the ALL format to see all fields of a record.

In WINISIS, users have the option of controlling the display of search progress and also automatic display of results of a search using parameters 103 and 105 respectively.

Also, in WINISIS, it is possible to affect how the dictionary terms are displayed in the dictionary window when a field is selected from the dictionary window using parameter 104 of SYSPAR.PAR: it is possible to display only those terms from the selected field or including other terms, that are indexed from other fields, but being dimmed and in lower case. In either case, WINISIS will display the number of occurrences of each term in the selected field. Finally, concerning search, in WINISIS, it is possible to put the search window and the database window side by side, perform searches and see the result on the same screen.

## 4.6 Sorting and Printing

Both Micro-CDS/ISIS and WINISIS have incorporated facilities for generating outputs either in the form of printed output or to a file in a user defined format. This facility enables users to obtain printed outputs, which are important for generating notices, reports, bibliographic indexes, etc. Printout of previously saved search results may also be obtained. Users may also sort the records to be printed in almost any combination of fields.

### A. In Micro-CDS/ISIS

Since sorting and printing requires a number of parameters to be defined, Micro-CDS/ISIS, uses a system worksheet to collect these parameters from users. However, users may also define and use their own print and/or sort worksheets containing the most frequently used parameters.

#### Using the Print Worksheet

In Micro-CDS/ISIS, sorting and printing facility is accessed by choosing option **P - System print worksheet** from the **Sorting and Printing Services** menu. The system print worksheet, displayed in Figure 4.13, is used to collect fifteen items of information regarding users' requirements of sorting and printing. Users are expected to submit some or all of the data concerning their print requirement through the system print worksheet.

Database Name _____	MFN Limits _____	Save file name _____
First Title _____		
Second Title _____		
Third Title _____		
Print format _____		
Line Width     70	Line Width     1	Column Width     70
Lines/page    60	First page number   1	EOC tolerance     3
Data indentation 0	Sort ?            N	Sort worksheet name _____
		Column Width     _____
EDIT: Replace		Last Page

*Figure 4.13 : System print worksheet*

## Sorting Output

In Micro-CDS/ISIS sorting of a print output can be specified by using the **system sort worksheet** or using a pre-defined users sort worksheet. The **system sort worksheet** (see Figure 4.14) will be presented after the system print worksheet if users specified that the output be sorted but did not submit name of their own sort worksheet file in the **system print worksheet** or when they are preparing their own print worksheet.

In Micro-CDS/ISIS records may be sorted in virtually any desired combination of fields and subfields. The field(s) by which the records are sorted may be used as headings, and users can define the number of levels of headings they want, for instance, two levels of headings may sort data by authors (first level), and then by titles (second level).

To sort their printouts users have to specify some of all of the following data:

1. the number of levels of heading to be used (the maximum is four);
2. name of a Stopword file - to avoid unwanted words from being used for sorting when sort keys are generated from a field indexed using indexing technique 4;
3. heading format - to specify how headings will be printed instead of using the default heading format provided by Micro-CDS/ISIS; and/or
4. for each sort key users: define the content of the sort key, specify the way headings are generated, and size of the sort key.

During printing/sorting users can not instruct Micro-CDS/ISIS to save the worksheet they just filled in for future use. They have to use the **System Utility Services** for creating their own print and/or sort worksheets.

Number of Headings _____ <b>1</b>	Stopword file name _____
Heading format _____	
_____	
Length of first sort key ____	Heading processing indicator <b>0</b>
FST for the first sort key _____	_____
_____	
Length of second sort key ____	Heading processing indicator <b>0</b>
FST for the second sort key _____	_____
_____	
Length of third sort key ____	Heading processing indicator <b>0</b>
FST for the third sort key _____	_____
_____	
Length of fourth sort key ____	Heading processing indicator <b>0</b>
FST for the fourth sort key _____	_____
_____	
- Next page   M - Modify   C - Cancel   X - Exit	

**Figure 4.14 : System sort worksheet**

## **B. In WINISIS**

In WINISIS sorting and printing facility is accessible by users using the Print option of the File menu. WINISIS implements the Print dialog box to collect sorting and/or printing requirements of users. The Print dialog box consists of five *pages* (or *tabs*) which are designed to collect groups of information regarding printing and sorting requirements of users:

### **1. General Page**

Using the *General Page* of the **Print** Dialog box (shown in Figure 4.15) users should specify what and where to print, and the print format they want to use. Users should specify:

- records they want to print, which can be: records retrieved by a last search or records marked from them, or range of records specified by MFN,
- an output device, which can be: the current Windows default printer or an ASCII file, and

- name of the print format used for printing the selected records.

Users can modify a print format shortly before printing and saving it for future use; also it is possible to save the entire print setting for future use. Users can also see how the records are printed- a sample print preview.

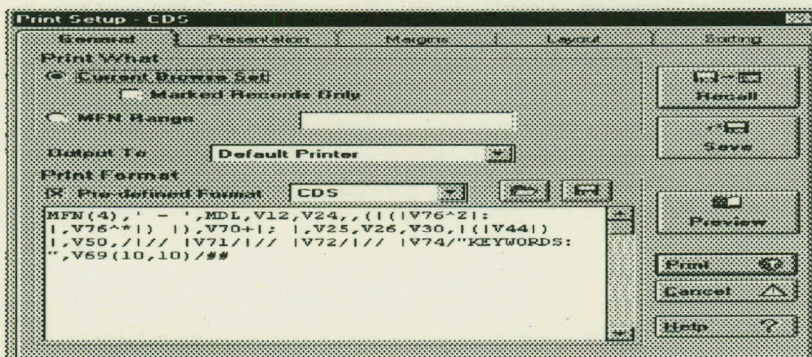


Figure 4.15 : General page

## 2. The Presentation Page

The Presentation page of the print dialog box (shown in Figure 4.16) is used to define and edit any of the following: up to three titles; a header; and/or a footer. In addition, for each title users can choose: to keep the title on all pages; the title's text alignment; the title's font, font size, and/or font style; and/or to have frame around a title text.

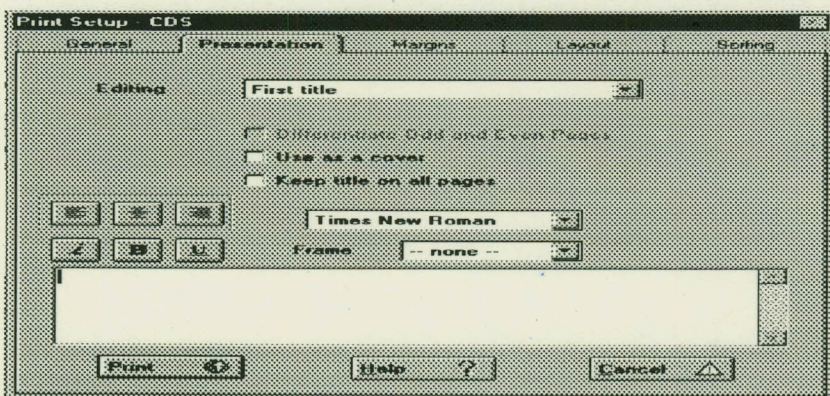


Figure 4.16 : Presentation page

### 3. The Margins page

The margin page of the Print dialog box enables users to define the page printable area. \users can define or change the default: margins - top, bottom, inside, and outside; binding space; and paper size. Using the Margins page users can also change the measurement unit used for the various margins: Millimeters, Centimeters, Inches, Twips; data indentation and end of column tolerance (see Figure 4.17).

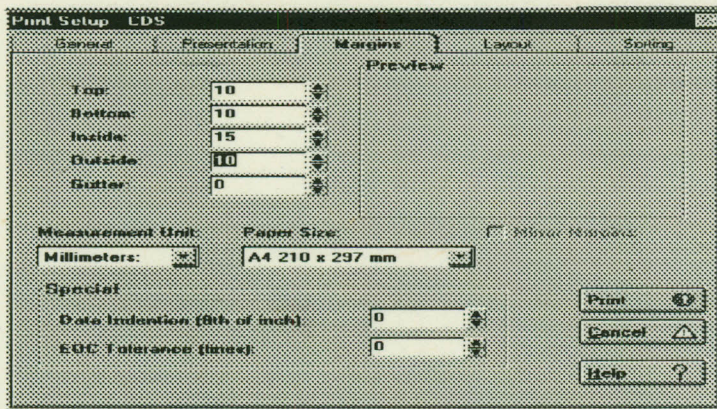


Figure 4.17 : Margin page

### 4. The Layout Page

In the Layout page of the print dialog window (see Figure 4.18) the layout of a print document, such as page numbers and number of columns can be defined or modified. In particular users can:

- define number of columns per page (up to six) - to print in columns;
- specify space between columns - for columnar printing;
- specify whether or not they want lines between and/or boarder around each column;
- specify whether or not they want page numbers, and if they want whether or not they want the page number appear on the first page, the starting number of the page number, its position, and its alignment; and
- specify whether or not they want sorted output.

In the Layout page users can also specify the scale by which the records will be printed on the selected paper size, which enables users to fit more (or less) records on a single page. In addition, even if titles, header & footers, and/or sorting have been defined, users have the option of ignoring any of them just before printing.

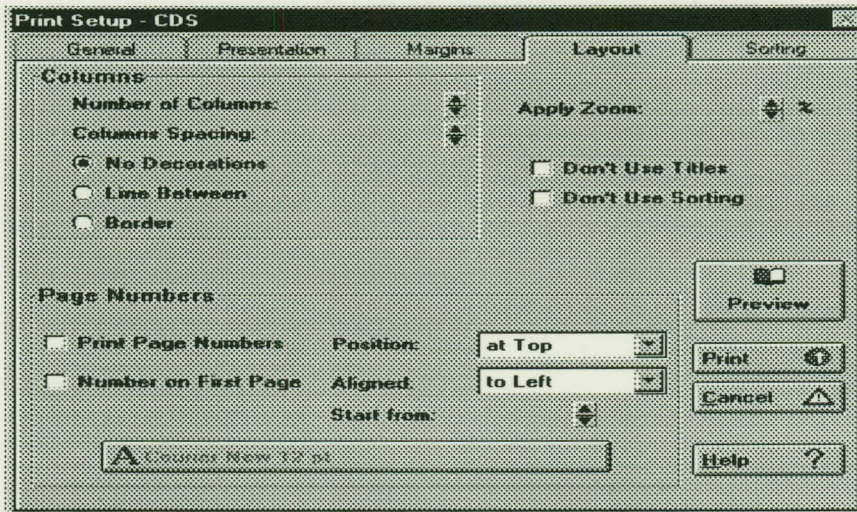


Figure 4.18: Layout page

## 5. The Sorting Page

The Sorting page of the print dialog box is exactly equivalent to the system sort worksheet of Micro-CDS/ISIS except for a check box beside each FST Key fields which tells WINISIS to consider (if selected) or ignore (if not) the related sort key.

### Discussion on Sorting and Printing

The parameters used by Micro-CDS/ISIS and WINISIS concerning users sort requirement are identical. No single difference. However, in the case of WINISIS, first of all, the parameters used to collect information concerning printouts are divided into four groups:

**General, Presentation, Margins, and Layout.**

Some features of WINISIS, which are not found in Micro-CDS/ISIS, enable one to have more control on what and how records are printed. In particular:

- marked records can be printed;
- it is possible to modify an existing PFT and use it for printing, and even save it for future use;
- print titles can be formatted almost as in any word processors;
- users have more control in defining page printable areas, and change measurements used for that purpose;
- position of page numbers and their fonts can be specified;
- it is possible to increase or decrease the scale by which records will be printed, hence enable to print more or less fields and records in a page;
- print preview is possible for the first page (as a sample);
- it is possible to print in ASCII; and
- It is possible to save a print set up (i.e., all individual print and sort parameters) during a print session for future use. Hence, there is no need of having other facilities, like System Utility of Micro-CDS/ISIS, where users create and save their own print and/or sort worksheets, which contain most frequently used users' parameters.

#### **4.7 Modifying A Database**

In addition to defining a new database, in both Micro-CDS/ISIS and WINISIS users can modify definition of an already existing database. Some of the fields may be deleted from the field definition table, or additional fields may be added to it. In both packages, the field definition table, however, can neither be deleted nor is there an option for making a copy of it. Whenever a change is made to the FDT necessary modification may have to be made to the worksheet, display format, and/or the FST. Otherwise, inconsistencies among the database files may exist which in turn may result in the corruption of the database files.

Users can create additional worksheets, display formats, and field select tables by copying, modifying and saving existing worksheets, display formats, field select tables, or can define any of them from scratch. Modification to existing worksheets, display formats, and field select tables can be made, and those which are no longer needed can also be deleted. However,

the default worksheet, display format and field select table created during database definition cannot be deleted.

### **1. Modifying The FDT**

In both Micro-CDS/ISIS and WINISIS, users have to modify their FDT if they want to perform any of the following:

- add new fields,
- delete existing field(s),
- change field length (for Micro-CDS/ISIS only),
- change field type,
- change field pattern,
- add subfields to fields,
- change field names, and/or
- change field tags

However, each time users change the FDT both packages perform a consistency check of the worksheets defined for the database and will, if necessary, automatically modify them under certain conditions. Additional changes, however, may be required not only to the worksheets but also to the FST and the display formats defined for the database. In case of major modification of the FDT, changes may be required in the database file itself, users may be able to perform them automatically using the import/export facility or they may have to do them manually. For instance, adding a new field may arise a need of adding it to a worksheet, a display format, and/or possibly to the indexing FST of the corresponding database; and if users delete a field from the FDT, both packages will automatically prevent further data entry into the field and the field will be changed automatically to a title field. But, if there is previously entered data users should change the display format(s) to stop the field from being displayed, delete the field from the worksheet(s), remove the data from the master file by exporting and then importing or manually record by record in Micro-CDS/ISIS and including the Global Field Deletion facility in WINISIS. If the field was inverted, users should delete its entry from the indexing FST and re-invert the database.

major modification of the FDT, changes may be required in the database file itself, users may be able to perform them automatically using the import/export facility or they may have to do them manually. For instance, adding a new field may arise a need of adding it to a worksheet, a display format, and/or possibly to the indexing FST of the corresponding database; and if users delete a field from the FDT, both packages will automatically prevent further data entry into the field and the field will be changed automatically to a title field. But, if there is previously entered data users should change the display format(s) to stop the field from being displayed, delete the field from the worksheet(s), remove the data from the master file by exporting and then importing or manually record by record in Micro-CDS/ISIS and including the Global Field Deletion facility in WINISIS. If the field was inverted, users should delete its entry from the indexing FST and re-invert the database.

## **2. Modifying Worksheets or Creating a new one**

For existing worksheets in both packages it is possible to make any of the following changes:

- Add new field(s);
- Delete existing field(s);
- Change alignment of existing fields (only for Micro-CDS/ISIS);
- Modify one or more parameters of a field or fields;
- Insert or delete one or more blank lines(only in Micro-CDS/ISIS); and/or
- Reposition existing field(s).

If users want to have more than one worksheet for a database (for example, one for books, one for serials, and one for conference proceedings in a bibliographic database), instead of

creating each from scratch they can create by copying an existing one and modify it to suit their needs.

### **3. Modifying A Field Select Table**

In both packages users should modify the FST of a database if they want any of the following changes concerning the index file (and hence subsequently the search):

1. change the indexing technique(s) used for a field;
2. add new index terms from a field; and/or
3. avoid index terms of a field.

The above changes may also arise due to any of the following changes in the FDT on fields which are used or intended to be used for searching:

- change of field tag of a field;
- field(s) has been deleted;
- new field(s) has been added; and/or
- subfield structure of a field has been changed.

But changing the FST changes the rules for the indexing of records subsequently added to the database. It does not alter the entries of records which have been added or modified since the last update. Both Micro-CDS/ISIS and WINISIS do not consider updates needed for old records. Therefore, for the modified FST to take effect on existing records, users will have to fully invert the database. But, whenever new records are added and existing ones are modified, Micro-CDS/ISIS reminds users to update the inverted files, so that only added records or modified records will be considered in the index. But, in WINISIS users can use parameter 130 of the SYSPAR.PAR to automatically update the Inverted file

whenever users modify records, otherwise WINISIS will not remind users to update the Inverted file.

#### **4. Modifying/Creating New Display Formats**

In Micro-CDS/ISIS, there are two ways in which users can modify a display format: using the **Database Definition Services** or using **Information Retrieval Services**. **Information Retrieval Services** allow users to make a temporary change, try it out immediately, and save it when they are happy with the result. Similarly, in WINISIS users can modify display/print formats using the Edit menu or by splitting the database window into two, one for records and the other for PFTs and modify them while seeing the effect. In either case of both packages, users can modify the format by deleting, changing, or addition of individual formatting commands and/or additional texts so that new way of visualizing and/or printing of records can be attained. Users can create different display formats for different purposes either from scratch or by copying and modifying existing formats.

#### **Discussion on Database Modification**

In both Micro-CDS/ISIS and WINISIS the reasons for modifying files of a database are almost the same. But, the way modification is done is somewhat different, and probably easier in WINISIS. For instance:

- in WINISIS to insert a new field or modify an existing one in an FDT there is no need of notifying the system before hand; and hence the first field of the FDT in Micro-CDS/ISIS is not necessary in WINISIS;
- in Micro-CDS/ISIS when users remove a field, they have to delete existing entries manually one by one or using the Export/Import facility. But, in WINISIS it is possible to use the more simpler method: using the Global Field Deletion facility.

- in WINISIS, whenever a new field is added in an FDT it will be appended, but it is possible to sort the entries by their tag;
- in Micro-CDS/ISIS adding new fields among existing fields needs creating a new room before hand (both in the FDT and data entry worksheet)
- in Micro-CDS/ISIS, removing a field in a data entry worksheet or decreasing its width does not re-adjust the location of remaining ones;
- in Micro-CDS/ISIS order of a field of a worksheet is changed by removing the field, create a room for it, and re-creating it;
- in the data entry worksheet of Micro-CDS/ISIS, modifying a parameter of a field needs answering unnecessary other seven prompts of the editor;
- in Micro-CDS/ISIS the presence of immediately following fields prevents significant increase of field length (unless it is made scrolling or a new room is prepared); and
- in WINISIS it is possible to modify PFTs by looking at the effect of modification on the same screen with the records.

But, in Micro-CDS/ISIS there is a direct option to delete unwanted worksheets, display formats, and FSTs, which is not available in WINISIS even if these operations can be done by other means.

#### **4.8 Exporting and Importing Records**

Both CDS/ISIS and WINISIS provide import and export facilities for interchanging data with other systems that support ISO 2709 standard. Micro-CDS/ISIS uses import/export system worksheets and WINISIS uses the import/export dialog windows which are almost identical to those of Micro-CDS/ISIS.

During export or import operations, in both Micro-CDS/ISIS and WINISIS, users can use an FST to limit part of a record which will be transferred, to move data from one field to another on transfer and/or to transfer the content of a single field into many fields. Thus, another use of the export/export facility is to reformat a database by exporting and then re-importing the same records.

The record structure which is used by Micro-CDS/ISIS and WINISIS for exporting and importing purpose is known as ISO 2709. This is used by some other bibliographic software, such as ISIS, MINISIS and any other software that allows the import or export of MARC records. But, it is not used by all bibliographic software. However, a program called *Fangorn*, developed by UNESCO, is available which can convert data from on-line databases and other sources into ISO 2709 format, so that they can be imported into Micro-CDS/ISIS or WINISIS.

#### **Exporting in ISO 2709 Format**

It is possible to export records on to hard disk, a floppy disk or via a network. If users want to export using floppies they have to make sure that they have sufficient formatted discs ready.

Normally the system worksheet shown in figure 4.19 is used for collecting the necessary information. But, it is possible to use a pre-defined users Import and/or Export worksheet which has been prepared using the **System Utility Services** (in Micro CDS/ISIS). Users should provide some or all the data requested by the fields of the system export worksheet, given in Figure 4.19, concerning the records to be exported.

When used as an export reformatting file, the FST is interpreted in the following manner:

- each line of the FST represents an output field;

- each output field is assigned an ISO tag equal to the field identifier defined in the corresponding FST line; and
- The data extraction format given in the export FST defines the contents of the field. In this format users must use the CDS/ISIS tag of the fields as defined for the database. Each line produced by the format (or each element, if the FST specifies indexing techniques 2, 3 or 4) will generate a new occurrence of the output field.

<b>Data interchange services - Import parameters</b>			
Database name _____			
<b>ISO file parameters</b>			
Output ISO file name	MST.ISO	Field separator	#
		Record separator	#
<b>Selection parameter</b>			
MFN Limits	_____	Save file name	_____
		Hit file (Y/N)?	N
<b>Reformatting parameters</b>			
Reformatting FST	_____	Gizmo conversion file	_____
Renummer records from	_____		

*Figure 4.19 : System export worksheet*

If the file to be exported is transmitted over a telecommunication line, the characters used as field separators and record separators may be deleted by the communication software. Therefore, Micro-CDS/ISIS allows users to redefine the field separator and the record separator as normal ASCII characters which will pass through. However, the characters selected must be such that they are never used as data characters in the database to be transmitted.

In WINISIS to allow users specify any ASCII characters as field separator and record separator, the following notation is introduced:

**\xxx**

where, xxx represents the corresponding ASCII number of the character selected by users as a field or record separator. For example: to specify as field separator the ASCII character number 255, users must enter: \255 in the field separator box.

### **Importing in ISO 2709 Format**

The importing function of Micro-CDS/ISIS is reached by selecting option **I-Import external file** from **Master File Services**. Normally a worksheet (shown in Figure 4.20) is used to collect the necessary information about the source ISO 2709 file and the destination database. Users of Micro-CDS/ISIS are expected to provide some or all of the data requested by the fields of system import worksheet, given in Figure 4.20.

The FST, when used as an import reformatting file, is interpreted in the following manner:

- each line of the FST represents an output field;
- each output field is assigned a tag equal to the field identifier defined in the corresponding FST line;
- the data extraction format given in the FST defines the contents of the field. In this format users must use the ISO tag of the fields as defined in the input file. Each line produced by the format (or each element, if the FST specifies indexing techniques 2, 3 or 4) will generate a new occurrence of the output field. The Stopword file of the database, if present, will be used in processing indexing technique 4

Data interchange services - Import parameters			
Database name _____			
<b>ISO file parameters</b>			
Input ISO file name	MST.ISO	Field separator	#
		Record separator	#
<b>Load option</b>			
First MFN to be assigned	___ 1	or Input Tag containing MFN	
Load/Merge/Update	L		
<b>Reformatting parameters</b>			
Reformatting FST	_____	Gizmo conversion file	_____
EDIT: <b>Replace</b>			Last page

Figure 4.20 : System import worksheet

## Discussion on Import/Export

The parameters used for import and export in both Micro-CDS/ISIS and WINISIS are almost identical, except few variations:

- in WINISIS it is possible to use any ASCII characters as record and/or field separators; and
- it is possible to export MFNs in a field in addition to others by specifying a tag without using *export FST*.

## 4.9 System Utility

### A. Of Micro-CDS/ISIS

Micro-CDS/ISIS allows users to create and/or edit menus and system worksheets. Print out of the system worksheets, menus, and system messages can also be obtained. This in turn enables users to get the hard copy equivalent of the worksheets which may be used for data preparation, and also of the messages which may be used for translation of the messages in designing a new language version of the software, for instance.

Using the **System Utility Services** users can tailor Micro-CDS/ISIS to their requirement.

For example, they can put their institution's name on the main menu, add special Pascal

programs to perform functions not provided in a menu (such as *global edit*), add HELP screens or *introductory screens*, change the terminology used by the system, add screens in languages other than the basic English, French, and Spanish provided at installation.

### **Creating/Modifying Menus**

In Micro-CDS/ISIS all menus are named in a systematic way. The first character of a menu name indicates the language: English is **E**, French is **F**, and Spanish is **S**. The second letter for every menu name is **X**. Up to three further characters follow which make a distinctive name for the menu. For example, EXISIS means English Menu ISISs.

Hence by selecting option **M-Create/Edit system menus** from **System Utility Services** and specifying name of a menu, users can do any of the following modifications to any existing menu:

- add new titles;
- change existing titles' text, screen attributes, and position within the screen;
- change option descriptions' text, screen attributes, and alignment;
- add new options to perform new functions (*menu exits*);
- change option identifier letter of an existing option; and/or
- change option identifier's screen attributes.

In addition, it is also possible to create HELP screen and link them to any standard menus of Micro-CDS/ISIS. For instance, users can create a new menu by the name EXHU1 which is called by the menu EXM1 in the System Utility Services. Also, users can create genuine new menus, with executable options, and use it for specialized applications written in ISIS Pascal.

## **Creating/Modifying System Worksheets**

Each individual prompt that appear on the four system worksheets is called a field. Using the **System Utility Services** facilities of Micro-CDS/ISIS, for any existing system worksheet it is possible to make any of the following permanent changes:

- change label of a field, default value of a field, its position, field type, any associated HELP messages, screen attributes; and/or
- create new titles, or modify existing titles.

Also, by copying existing system worksheets and making any of the above mentioned modifications users can create their own Import, Export, Sort, and/or Print worksheets, especially if the data to be specified in any of the above system worksheets is more or less the same whenever users use any of these system worksheets.

## **Changing System Screen Attribute**

If users are unhappy about the screen attributes used by Micro-CDS/ISIS for both color and monochrome displays they may change them. It is possible to change text attribute and background attribute of the work area (lines 1-22) and messages area (lines 22-24), and change the Micro-CDS/ISIS *definitions* of Normal Text, Reverse Video, Bold, Underline, Blinking, and/or Invisible which are used in the different parts of the Micro-CDS/ISIS screen. Normal attribute is used for displaying prompts and messages (other than those displayed in the message area) as well as for displaying field names (in worksheets and menus) and field values defined with the normal attribute.

If a designer wants to give different users different combinations of colors, he/she will have to direct them to different message files, since the system stores the values in the message file, which is activated when Micro-CDS/ISIS program runs.

### **Changing Message File**

Micro-CDS/ISIS messages are stored in a database for each language, EMSG is a message database file for English language messages, for instance. Each message in Micro-CDS/ISIS has a number by which it is known to the system. Messages are grouped into records in the message file, each containing ten fields (i.e., ten messages). Thus, message 38 is a eighth field in the fourth record, for instance.

Hence, by accessing the message file of a language and editing its records just like a normal database it is possible to modify existing messages. Before changing messages it is possible to print any message file for examining available messages, for instance.

### **B. Of WINISIS**

Each menu option and the commands of the database control-menu box are stored as a record in databases file called MNENDF.MST (for English menu) and MNITDF.MST (for Italian menu), so, each have a total number of nine records. By opening these databases as normal database and editing them, it is possible to modify system menus.

There is also another menu database called MNENSH.MST (for English language) which contains those menus which are intended as a search profile of WINISIS. This database can also be modified just like the main menus.

### **Modifying Message and Screen Texts**

Like menu options, error messages, button text, labels of drop-down boxes of the different dialog boxes of WINISIS, text that appear in the status bar when the mouse pointer points to some screen elements etc., are together found in databases called MSEN.MST (for English language) and MSIT.MST (for Italian language).

Hence, by opening and editing records of these databases it is possible to customize the different messages and texts that appear in the different dialogue/window buttons (such as

Ok, Cancel, Help, Retry, Run/Execute, Display, Clear etc.), boxes (text boxes, drop-down boxes, result boxes, etc.), title bars (database definition title bar, search history title bar, etc.), and/or status bars. Simply speaking, it is possible to edit all the messages and/ text that makeup the screens of WINISIS.

### **Discussions on Systems Utility**

Facilities provided by the **System Utility Services** of Micro-CDS/ISIS (such as creating/modifying menus/system worksheets and modifying screen attributes/system messages) can be done in WINISIS in a restricted manner, even if there is no direct interface which is prepared for such purposes.

Message modification is completely identical. Concerning menu creation/modification users of WINISIS can change only option descriptions; similarly for system dialog boxes users are allowed only to change the different descriptive texts used in the boxes dialog box. But modification of screen attributes can be done in another way for database record elements only: by defining colours, using colour groups, and using character formatting of Windows (such as bold, underline, italic) it is possible to display different elements of records in different colours and/or font attributes.

All the facilities provided by the **System Utilities Services** of Micro-CDS/ISIS are not provided by the current release of WINISIS: for instance, it is not possible to create a HELP screen to be attached to a menu or creating a genuine menu for an application developed by ISIS Pascal. However, some facilities which are provided by **System Utilities Services** of Micro-CDS/ISIS , can be done in WINISIS in another way even if there is no designed menu option for such purposes: for instance, messages files can be printed just like normal databases, since it is possible to access the print facility of WINISIS from the data entry window. Similarly the *Copy worksheets* and *Copy formats*

facilities of the **Systems Utility** (for creation of others by modification) can be done in WINISIS using the File Manger application of Microsoft Windows without exiting from WINISIS. Hence, some facilities provided by the **Systems Utility** of Micro-CDS/ISIS seems unnecessary in WINISIS.

However, there are new facilities in WINISIS, which are not available in Micro-CDS/ISIS, which enable users to control display of record elements and the interface: the way terms are displayed whenever a field is selected in the dictionary window (using parameter 104); it is possible to control how search results are displayed (using parameter 105); specify the default font to be used during data entry/modification (using parametr 109); default RTF header to be used for formats that comes from Micro-CDS/ISIS (using parameter 110); the deafult font that will be used for displaying and editing print formats (using parameter 111); users can control the display of TOOLBARS and STATUS BAR (using parameter 120 and 121 respectively); it is possible to store information about the size of the main window of WINISIS (using parameter 122); users can use parameter 123 to control how empty fields will be displayed during data entry/modification; and specify upto five files that will be displayed at the bottom of the FILE menu for easy opening.

## **4.10 Security Measures**

### **4.10.1. Backing Up and Restoring A Database**

#### **A. In Micro-CDS/ISIS**

In order to protect data from being lost by accident, either users' own or the computer's, users can make regular backups of their database(s). During backup any record which had been logically deleted from a database will not be included in the back up, and users will

get a report about these. When disaster strikes, users can restore their data from their recent backups.

The backup facility of Micro-CDS/ISIS will only backup the master file records. It will not include the four database definition files and the index files even if these files can be re-created after restoring records.

#### **4.10.2 Password Protection**

One can use passwords to protect Micro-CDS/ISIS from unauthorized use. Designers can give different individuals or groups of users passwords which will lead them to pre-arranged menus, worksheets and/or display formats (hence, hiding certain functions and/or parts of a database), to the read-only version of Micro-CDS/ISIS (**ISISCD**) as well as deny access to certain database files as a whole.

Password can be allocated by creating one SYSPAR.PAR containing generally applicable parameters for all individuals or group of users with parameter **0** and other system parameter files (where their names correspond to the different prepared passwords) which may be empty or possibly containing selected parameters specific to each individuals or groups of users. Hence, since Micro-CDS/ISIS checks SYSPAR.PAR whenever it starts operation, a user will prompted for a password (due to the line which starts by parameter **0**) when he/she tries to use the system, and accordingly he/she will be directed to an environment which has been designed for him/her.

#### **B. In WINISIS**

Like Micro-CDS/ISIS, it is possible to direct users to a modified menu or the search only menu of WINISIS and/or deny access to certain database files or part of them using

passwords and system parameter files. However, the backup and restore operations are not supported.

### **Discussion on Security**

In both Micro-CDS/ISIS and WINISIS data protection using passwords is identical. But, in Micro-CDS/ISIS to protect certain functions a designer has to create other menus by modifying the standard menu or use the **ISISCD** (which is a full system by itself); where as in WINISIS there is a ready made menu which denies certain functions and can be directly implemented using the new parameter (parameter **102**), even if it is possible to create others and use them for this purpose.

### **4.11 Advanced Programming Facilities**

Micro-CDS/ISIS includes a Pascal compiler and facilities to write a program and interface it to the existing system. CDS/ISIS Pascal (also called simply ISIS Pascal) is a programming language designed to develop Micro-CDS/ISIS applications requiring functions which are not readily available in the standard package. In order to use ISIS Pascal users must therefore be familiar with Pascal programming language and, obviously, with CDS/ISIS. The specificity of ISIS Pascal is in fact its library of pre-defined procedures, and functions which provide access to most Micro-CDS/ISIS functions in a convenient manner.

ISIS Pascal is an integral part of Micro-CDS/ISIS and consists of a compiler, an interpreter, and a library. The compiler produces pseudo-code which is then executed by the interpreter. Because the executable code is machine-independent, application programs written in ISIS Pascal are fully portable across the whole spectrum of computers supported by Micro-CDS/ISIS.

In Micro-CDS/ISIS, users may design programs in stand alone mode (which provide the equivalent of a *batch operation*) or as a *user exit* in certain Micro-CDS/ISIS functions. User exits provide powerful way to extend the functionality of Micro-CDS/ISIS (for instance, user exits can be used to check a password whenever a user tries to edit a database using the **Data Entry Services**).

User exits can not be executed from **Advanced Programming Services** even if it is compiled there. The execution of a user exits is controlled by the CDS/ISIS program or function for which it is designed. User exits can be divided into two according to the purpose for which the are prepared:

#### 1. **MENU exits**

*Menu exits* are those programs which are designed to add new functions to or modify existing functions of Micro-CDS/ISIS. Menu exits provide extra options for menus. A menu exit is activated each time users select its assigned option from a system menu. Major menu exits which are prepared to improve the functionality of Micro-CDS/ISIS are:

- a menu option to display the fields of a database;
- a menu option to display one record at a time on screen;
- a program to enable new databases to be created more easily;
- a program to modify a record by selecting a field from a table by tag number rather than by the usual worksheet;
- menu options to enable global edit; and
- a program to assist in the selection of worksheets, print formats and databases, by listing the options in a window.

## 2. Format Exits

A *format exit* is used to perform special formatting functions required by a particular application, which could not otherwise be obtained using the formatting language of Micro-CDS/ISIS. A format exit may be used to provide data to the standard formatting routine (for instance, bring only the first occurrence of a repeatable field to a print format). A format exit is invoked by calling it in a user display/print format.

### 4.11.1 ISIS Pascal Support of WINISIS

A file called ISISPAS.PIF enables the ISIS Pascal compiler used in WINISIS. ISIS Pascal programs are currently implemented only as format exits. The programs are fully compatible at **.PCD** level with Micro-CDS/ISIS.

Restrictions on string size and number of strings in the DOS version no longer exist. A string may now be up to 32K, including the input and output strings of a format exit. The following ISIS library routines are available for WINISIS:

ASSIGN	POSTS	FILEEXIST	RECORD
CHR	NXTPOS	WRITE(ln)	NPOSTS
COPYSTR	NXTPOST	GETMFN	SIZE
DATESTAMP	ORD	ENCINT	PATH
DBN	NXTTERM	MAXSET	SYSTEM
ENCREAL	POSITION	EXEC	POSTING
FIELD	READ(ln)	FIELDN	RECALL
FIND	SEARCH	FLDTAG	SETPOS
MAXMFN	SUBSTR	NOCC	VAL
NFIELDS	UC		

On some functions, however, the following implementation differences between the DOS and the Windows version exist:

#### **ASSIGN**

The **/k** switch (to prevent closing a file between calls) is not supported.

## **READ(In) and WRITE(In) procedures**

These functions are implemented only for files, i.e., Read(inp,...) and Write(out,...). Input from keyboard or output to the screen are not supported. This implies that programmers must always explicitly open the file to be read or written by using the **ASSIGN** procedure.

## **EXEC procedure**

1. All programs called through EXEC must be of the same type as the original caller (e.g. if program A is a format exit and it calls program B then B must also be a format exit);
2. If the calling program (A) is a format exit, then the called program (B) receives as input string (s1), the current value of the output string (s2) of A at the time the EXEC function is executed. Program B returns a value in the output string (s2). This allows programmers to pass an input parameter to, and to receive output from the executed program.

## **SYSTEM procedure**

This procedure may execute both DOS and Windows programs. If programmers are executing a DOS program and they want it to execute in a window (rather than in full screen mode), they must also create a corresponding PIF file.

The program called by the SYSTEM procedure executes synchronously and control returns immediately to the calling program. This feature may be used, for example, to play a sound file (e.g. a description of the record) while the record is displayed. Assume for example that field 1 contains the name of a .wav file associated with the record. Then the following format, and the associated format exit, will play this file while displaying the record:

```

Format: ..... &play() .....
Program play(s1: string; lw, occ: real; s2: string);
var wav: string
begin
    wav:=field(fieldn(1,1));
    if (size(wav)>0) then system('mplayer/play/close '|wav|'.wav');
end.

```

### **Discussion on Advanced Programming**

In both Micro-CDS/ISIS and WINISIS the functions and procedures found in ISIS Pascal are almost identical except minor implementation differences in some functions. Currently in WINISIS ISIS Pascal programs can be implemented only as format exits. But, in WINISIS menu exits are not supported. Hence, users can not enhance functionality of the system. However, some of those programs that are developed by UNESCO and other users to be used as user exits, such as to list files, display one record at a time during data entry etc., are not necessary in WINISIS since the system itself gives such facilities.

#### **4.12 New Features of the Formatting Language of WINISIS**

Microsoft Windows, as well as most graphics environments, offers a rich array of text-writing capabilities. In particular, Microsoft Windows lets users choose the font to be used for text output.

Similarly, WINISIS gives users some *keywords* to manipulate the display of records using various proportional and non-proportional fonts. Users of WINISIS can add these *keywords* to their formats and hence, giving them a new look. In fact such *keywords* extend the formatting language of Micro-CDS/ISIS when they are used in WINISIS. Moreover WINISIS adds some new functions for handling numerical data and strings.

## 1. New database functions

The following are new functions that have been added to the formatting language of Micro-CDS/ISIS to be implemented in WINISIS:

### A. Field occurrences

Users may now access individual occurrences of a repeatable field. For example:

v10[1]	retrieves the first occurrence of field 10
v10[3..]	retrieves the 3 <sup>rd</sup> through the <i>last</i> occurrence of field 10
v10[1]^a	retrieves subfield <b>a</b> in the 1 <sup>st</sup> occurrence of field 10

### B. New numeric functions

**LR(format)** - like the **L** function, **LR** searches the inverted file for a term defined by *format*, but returns *all* the postings of the term. For example: **REF(LR(V10),V1,V2)** will retrieve fields 1 and 2 from all the records posted under the term contained in field 10.

**NOCC(Vtt)** - returns the number of occurrences of field *tt*.

**NPST(format)** - searches the inverted file for the term defined by *format*, and returns the number of postings of the term.

**SIZE(format)** - returns the size of the string generated by the ISIS format *format*, (which depends on the MODE used in the *format*).

**TYPE** - test the type of a string:. For example:

**TYPE(TYPE, FORMAT)**

where, *type* may assume the following values: 0 - pattern; 1 - alphanumeric; 2 - alphabetic; 3 - numeric ;4 - decimal (optionally signed) integer; 5 - decimal number (including scientific notation) and *format* is a string resulting by a format. So, for instance, users may check if the field v40 contains a number or not as:

**IF TYPE(1,V40)=1 THEN 'YES' ELSE 'NO' FI**

**OCC** - returns the number of the current occurrence in a repeatable field. It may be used to produce numbered lists. For example, the following format: **(v70/)** produces one line for each occurrence in the field 70. Using the function **OCC** users may write:

**(IF P(V70) THEN F(OCC,1,0),'. ' FI,V70/)**

producing a numbered list as the following:

1. First Author
2. Second Author
3. Third Author

### *C. New string functions*

**DATE(exp)** - returns the current date and time in the format specified by the numeric expression *exp*. But, at present only format 1 is implemented, which returns a date stamp identical to the DATESTAMP function of ISIS Pascal (i.e. an 18-byte string of the form MM-DD-YY HH:MM:SS).

**DB** - returns the name of the current database.

### *D. Alternate database*

The **REF**, **L** and **LR** functions may now specify a database qualifier to refer to an alternate database different from the current one. For instance:

**REF->BIB(L(v10),v200)**

retrieves the specified record elements from the BIB database even if it is not the currently opened one.

## 2. Graphic-based Environments

In WINISIS results of display/print formats are displayed using the Windows non-proportional font Courier New to simulate the DOS character-based screens, in order to give the right meaning to indentation commands of Micro-CDS/ISIS. However, users can implement non-proportional fonts (like Times New Roman, Aril, etc.) for displaying and printing records. But using proportional fonts in Windows environment gives no meaning to the indentation commands of Micro-CDS/ISIS, since each character has a different width (which is proportional). Hence, WINISIS gives some new *keywords* (or indentation commands) to control the indentation of paragraphs. In any case, users of WINISIS can implement different fonts, font sizes, and font attributes.

### A. New Visual Functions

#### *The font table*

In WINISIS users can list fonts at the beginning of a display/print format and can use it afterwards in the format (otherwise WINISIS uses the default font - Courier New- from the default table - which contains only two fonts). A group of listed fonts is called *font table*.

A font table begins with the control word **fonts**. A font is defined by a font family and its name. The following are some font families including some of their fonts:

<u>Family</u>	<u>Fonts</u>
nil	Unknown or default fonts (default)
roman	Roman, proportionally spaced serif fonts.
Swiss	Swiss, proportionally spaced sans serif fonts.
modern	Fixed-pitch serif and sans serif fonts (Courier, ...)
script	Script fonts (Cursive, ...)
decor	Decorative fonts (Old English, Zap, Chancery,....)
tech	Technical, symbol, and mathematical fonts.

Each font definition must be enclosed between parenthesis as shown in the following example:

```
fonts((roman,Times New Roman),(swiss,Arial)...) 
```

The order by which fonts are listed represents their referring number, i.e., font numeration is automatically given in ascending order starting from 0 (zero). Hence, in the previous example the font 0 (zero) is the Times New Roman while the font 1 is the Aril one.

However, if users do not define any font table the default font table which has the following form will be used by WINISIS whenever it displays and prints records:

```
fonts((modern,courier),(swiss,Arial)) 
```

Also, even if users define their own font table but do not specify any usage among the listed fonts, WINISIS applies the first font from their font table.

In addition, using the new parameter (parameter 109) it is possible to specify a font name with a size which will be used for displaying field contents in any data entry/modification window.

### *Character formatting properties*

The following five control words (or visual functions for character formatting) have been introduced in WINISIS for character formatting of record elements:

<u>Control word</u>	<u>Meaning</u>
<b>b</b>	Bold Italic
<b>ul</b>	Continuous underline
<b>fn</b>	Font number
<b>fsn</b>	Font size

Any character formatting control word preceding plain text turns on the specific attribute. Some control words can be applied to specific parts of the format only. For example, **b** turns bold on for the rest of the format, while **b(v24)** turns bold on for field 24 only. In addition, the control words **fn** and **fsn** are required to be followed by a value (indicated by an "n" following the control word), specifying, for example, the size of the current font (for instance: **fs24**). However, in WINISIS font size must be indicated by the double of the points users want it to be. So, to obtain a 12 point sizes character, users have to type: **fs24**.

### ***New Indentation and tabulation commands***

The following two control words (or functions) have been introduced in WINISIS to control indentation and tabulation of record elements:

<code>m(li, fi)</code>	sets the indentation from the left margin and the indentation for the first line of the paragraph.
<code>tab(twips)</code>	inserts a tab character that goes to the specified distance from the left margin.

Instead of using the normal indentation commands of Micro-CDS/ISIS, users of WINISIS can define tabulators, as if they are working with a word processor application. Then, when they use a tabulator control word, they will be able to position their text at a certain specified distance from the left margin. In the same way, it is possible to define different indentations for the first line of a paragraph than for the rest using the first keyword. However, the measurement unit to be used in WINISIS *twips*, which is one twentieth of a point, instead of characters.

### ***The color table***

In WINISIS, users can implement different colors, which are supported by Microsoft Windows, while displaying their records. Color information are contained in the *color-*

*table group*. The control word **cols** begins this group. Here users must specify for each color a triple of values (red, green, blue from 0 to 255) corresponding to the color indexes used by Microsoft Windows to define the amount of red, green, and blue that makes up a color. Normally, records are colored black when displayed. Users can define more colors to give their records a more attractive look. Colors are made by a mixture of three base colors: Green, Red and Blue. The extremities are the black color (0,0,0) and white (255,255,255). Since colors have no name, users refer to a color by its index number. Indexes start from 0 (zero). Each definition must be enclosed within parenthesis. For instance, to define three colors: black, dark blue, and dark red users have to write the following color table and insert it at the top of their display/print format:

```
cols ((0,0,0), (0,0,128), (128,0,0))
```

The control word **cn**, where *n* is the colour number, enables users to select any of the defined colours. However, if users do not specify any colour number from their table, WINISIS will automatically implement the first of their table (colour 0) for displaying all record elements. In addition, if a user tries to use a colour which is not defined, WINISIS will consider the Windows default palette. For instance, in the above example if users try to use color number 3, which is not defined there, WINISIS will take color 3 of Windows Palette, which corresponds by default to the green.

### **The Stack**

All the keyword commands take effect immediately; so, for instance, if users turn the bold on all the rest of their format result will appear bold. Hence, to control switching properties WINISIS implements braces {}. In particular, a left brace signals that subsequent text may have different attributes and the consequent right brace cancels any

change by restoring the attributes that were active before the left brace. This procedure is called a *stack*, and particularly used to apply formatting properties to specific objects only.

#### 4. Adding Hypertext features to a format

In WINISIS users can write interactive formats where they will be able to perform some operations by clicking on hyper-textual terms. Such terms can be defined by using the following keyword:

```
link((descriptor),action)
```

where, the *descriptor* is a format whose output will be displayed using the foreground color 2 (cl2) of Windows and single underlined (green color by default), and the *action* term represents the command to be executed when user clicks onto its descriptor. It will not be displayed at all. The following are example commands that can be used in the action part of a hyper-textual commands:

- **CMD string** - Runs command line (just like the "File-Run" menu option in the Program Manager) . For example the following action:

```
'CMD pbrush c:\\winisis\\ab_jan97.bmp'
```

executes the Windows Paintbrush (\windows\pbrush.exe) and lets it open the c:\winisis\ab\_jan97.bmp bitmap file.

- **GOTO mfn-number** - enables one to go to a specified MFN number. The expression may be a format but it should return a string containing a number. For example:

```
link(('test'),'GOTO 1')
```

the action goes to the first record in the database, and in the following second example:

```
link(('test'),'GOTO 'f(1('water'))')
```

the action goes to the first occurrence of "water" in the current database.

- **TEXTBOX** *format* | **FILE** *filename* - executes the specified format and shows the result in a text box. Then users can keep the window opened for comparisons. For example, an action:

```
'TEXTBOX v24'
```

will display field 24 in a text box, while the command:

```
'TEXTBOX FILE cds001.pft'
```

opens a text box with the current record displayed using the cds001.pft format. This gives a second simultaneous view of the record.

- **FORMAT** *format-file-name* - changes the current display format into the specified one. Such format may have not been specified in to the FDT. For example the action:

```
'FORMAT test.pft'
```

redispays the current record using the specified format.

- **BROWSE** *database file-name* - opens a specified MST file in a new MDI window.

For instance the action of the following hyper-textual command:

```
link(('test'),'BROWSE thes.mst')
```

will open THES.MST as a new window.

## 4.13 Summary

Table 4.1 presents a comparison of Micro-CDS/ISIS and WINISIS based on features and facilities which are common to most text retrieval packages.

*Table 4.1 : Comparison of Micro-CDS/ISIS and WINISIS*

Characteristics/Property	Micro-CDS/ISIS	WINISIS
<b>I. General/System</b>		
Version	3.07	1.0
Operating Systems	MS-DOS	Windows 3.1 or superior
Minimum internal memory	512 KB	4MB
Minimum external memory	400 KB	2MB
Fast processor recommended	no	yes <sup>1</sup>
Multi-user version	yes <sup>2</sup>	yes <sup>2</sup>
Maximum number of files per program	unlimited	unlimited
Maximum number of records per file	16Million (within 500MB)	16Million (within 500MB)*
Maximum number of fields per record	200	200*
Maximum number of characters per record	8KB	8KB
Maximum number of characters per field	8KB	8KB
<b>II. Security</b>		
File and/or users	yes <sup>3</sup>	yes <sup>3</sup>
Group of functions	yes <sup>3</sup>	yes <sup>3</sup>
Hiding certain fields	yes <sup>3</sup>	yes <sup>3</sup>
<b>III. User interface</b>		
Screen interface language	Eng., Fre., Spanish	Eng., Italian
Reference manual/Training manual	both	not available yet
Manual Language	Eng., French	not available yet
Well organized	yes	not available yet
Comprehensible	partly	not available yet
Table of Content	yes	not available yet
Presence of index	yes	not available yet
Presence of HELP messages	no <sup>4</sup>	yes
HELP always present/on request	no <sup>4</sup>	both
HELP can be chosen	no	yes
HELP Level can be changed	no	no
HELP in context	no	yes
HELP useful and comprehensible	no	yes
On-screen instructions	no	yes
Error messages presence	yes	yes
Error messages pertinent	yes	yes
Error messages helpful	partly	yes
<b>IV. Input and maintenance of files</b>		
Internal screen editor	yes	yes
Deletion of records	yes	yes
Deletion of records reduces file size	yes <sup>5</sup>	yes <sup>5</sup>
Accept extended ASCII	yes	yes
Field structure possible	yes	yes
Field structure required	yes	yes
Variable field length	yes	yes
Fixed length fields	yes <sup>6</sup>	no

Control of input data	yes <sup>7</sup>	yes <sup>7,8</sup>
Obligatory input in certain fields possible	no	yes <sup>8</sup>
Pre-defined record structure present	yes <sup>9</sup>	yes <sup>9</sup>
Batch input from structured ASCII file	yes	yes
Downloaded data must be converted	yes	yes
Prescribed field order in batch file	no	no
Multiple occurrence fields in batch file	yes	yes
Field contents must be partitioned into lines	no	no
<b>V. Reorganization of files</b>		
Addition of fields	yes	yes
Removal of fields	yes	yes
Change of field names	yes	yes
<b>VI. Indexing</b>		
Searching via index	yes	yes
Maximum number of indexes	1	1
Maximum number of indexed characters per term	30	30
Stopword list	yes	yes
Maximum number of stopwords	799	799*
User defined stopwords	yes	yes
Indexing of words tagged during input	yes	yes
Indexing single words	yes	yes
Indexing compound words	yes	yes
Indexing complete (sub)-fields	yes	yes
Type of indexing can be changed	yes	yes
Index can be built or removed afterwards	yes	yes
Batch-indexing possible	yes <sup>10</sup>	yes <sup>10,11</sup>
Controlled vocabulary	yes <sup>12</sup>	yes <sup>13</sup>
Thesaurus with relations (NT,BT,RT,STN)	yes <sup>14</sup>	yes <sup>13</sup>
<b>VII. Searching</b>		
Boolean combinations	yes	yes
Combination of search terms in one statement	yes	yes
Combination of sets	yes	yes
Complex combination with parentheses	yes	yes
Left truncation	yes <sup>15</sup>	yes <sup>15</sup>
Right truncation	yes	yes
Internal truncation	no	no
Interval/Range searching	yes	yes
Proximity searching	yes	yes
Field tags required	no	no
Field tags optional	yes	yes
Multiple field searching	yes	yes
Definition of default field or Basic Index	no	no
Index can be displayed	yes	yes
Index shows number of occurrences	no <sup>16</sup>	yes
Terms can directly be selected from index	yes	yes
Sequential search of non-indexed text	yes	yes
Saving search strategy	no	yes
<b>VIII. Outputs</b>		
User defined formats	yes	yes
Records can be sorted on multiple fields	yes	yes
Field tags can be included in output	yes	yes
Search terms can be highlighted in output	no	no
Output formats for direct exchange with other programs	no	yes <sup>17</sup>
Output bibliographies according to standard journal-style	no	no

- 
- <sup>1</sup>Pentium at 100 Mhz, 16 Mbyte RAM, and Super VGA Screen (800x600).  
<sup>2</sup>Even if it is not a separate multi-user version it can be used in multi-user environment.  
<sup>3</sup>Password can direct users to a particular menus, worksheets, and/or display formats which hide certain functions and/or deny access to certain database files or part of them.  
<sup>4</sup>But users can define HELP for some operations.  
<sup>5</sup>Only after re-indexing and/or reorganization of the Master file.  
<sup>6</sup>For PATTERN filed types.  
<sup>7</sup>By field types defined at the FDT.  
<sup>8</sup>Subfield delimiters and repeatability are also checked. In addition users can define field and/or record validation criteria.  
<sup>9</sup>Including conversion module.  
<sup>10</sup>That is the only indexing method.  
<sup>11</sup>It can be made automatic using parameter 130.  
<sup>12</sup>Using the THES program of UNESCO.  
<sup>13</sup>THES program of UNESCO can also be used in WINISIS even if it can not be used as a menu exit.  
<sup>14</sup>Using the THES program of UNESCO  
<sup>15</sup>Only through free-text search.  
<sup>16</sup>But when printed.  
<sup>17</sup>ASCII and RTF.

\*Even if most system restrictions of the current release of WINISIS (May 1997 version) are almost identical to that of Micro-CDS/ISIS, the following extracts of a recent discussion with Mr. Davide Storti, member of the Developers Team of CDS/ISIS, are worth mentioning:

*The ISIS core has been designed to support multi-giga byte databases. Doing this, databases will no longer be compatible with the DOS version and will need some conversion tool. This is also for the record size that will be increased to 32KB very soon..*

Concerning indexing and inverted files, he stated that:

*Inverted file generation is now much faster than before (up to 3 times depending on the FST). We are going to introduce two new parameters in the DBNPAR.PAR file to set the inverted file terms length so that users may specify a term length as they want.*

Concerning display format, Mr. Storti also stated that:

*a format may now be up to 10KB, and soon it will become 16KB. A format may output a string upto 64 KBytes*

## CHAPTER V

### CONCLUSIONS AND RECOMMENDATIONS

#### 5.1 Conclusions

A text retrieval software is a software that is specifically designed to support the creation, maintenance, and use of database of textual data. Some of such data may also cater for numerical, tabular, graphical data and image.

Today a number of text retrieval packages are available in the market and they show a large variety in what they do how they do it. Indeed, the flexibility and potential of a text retrieval software package differs from one package to another.

Micro CDS/ISIS is an information storage and retrieval package developed by UNESCO specifically designed to handle structured non-numerical databases. Since the release of version 1.0 in May 1985 a number of efforts have been made to improve the functionality of the system. However, even if efforts have been made to improve the features and facilities of Micro CDS/ISIS there are many problems reported by different scholars and users. Hence, considering these problems and taking in to considerations the increasing popularity of Microsoft Windows as an operating system for micros UNESCO has released a Windows based Beta version of CDS/ISIS, called WINISIS, in July 1996.

Since July 1996 several Beta (test) versions have been released by UNESCO in different times: October 1996, January 1997, April 1997, and a pre-release version in May 1997.

As it is observed in this study (using the pre-release version of WINISIS and version 3.07 of Micro CDS/ISIS), and taking in to considerations the basic purposes for which both packages are prepared by UNESCO, WINISIS provides the following new (or modified) facilities

which are not totally available (or available in different ways) in the standard version of Micro CDS/ISIS:

1. there is on-line and context-sensitive help during every single operation;
2. automatic creation of SYSPAR.PAR for further modification and save during database definition is possible;
3. an almost finished data entry/modification worksheet based on a defined FDT will be provided during database definition;
4. there is no field type: PATTERN; and for any of the three field type a pattern can be defined, which can also be repeatable;
5. no 'pages' of data entry worksheets;
6. incorporation of field and/or record validation criteria definition during database definition;
7. avoidance of field length definition during database definition, and hence during worksheet definition;
8. it is possible to access individual occurrences of a repeatable field, and hence, there is no need of using format exits for such purpose;
9. other databases in addition to an active one can be accessed for retrieval purpose using the REF function, hence integration of databases is improved;
10. modification of display/print formats by observing the effect on the same screen as the records is possible, i.e., What You Modified Is What You Get (WYMIWYG) facility is added;
11. restriction on size of display format has increased to 10 KBytes with its output to 64 Kbytes, hence more record elements can be outputted by a display format;

12. global addition and deletion of field values in a range of records (by specifying MFNs or using search) is possible;
13. during data entry/modification terms can be selected from the terms dictionary;
14. it is possible to select (mark) records (from records which are results of a search) for further purposes: modification or print;
15. using parameter 130 of the SYSPAR.PAR it is possible to automatically update the Inverted file whenever changes are made in the data entry window;
16. data/information submission for system prompts are mostly by selection from listed ones;
17. there is a fundamental change in search interface: there are two search interfaces one for novice users and another for experts;
18. *search only* menu is available and can be used whenever the system starts using the new parameter (parameter 104), and this feature can enhance the system security feature;
19. from the terms dictionary users can select only those terms that are from a specified field;
20. in a field is display from the dictionary window WINISIS correspondingly shows number of occurrences of each term in the field;
21. search results can be saved for further search, hence this feature can simplify the operations needed for CAS and SDI;
22. opening of a database for use automatically makes it ready for browse;
23. during browse or display of search results:
  - only one record is displayed per screen,
  - 'record up' is possible,
  - different records can be displayed in different formats, and
  - it is possible to access the print facility.

24. more than one database can be opened for browse or search side by side in the same screen (but, of course, searching and browsing is one at a time);
25. usage of search history is much more simplified. For instance, only part of a previously submitted search expression can be copied to formulate a new search expression;
26. users can specify their print requirements in a much better ways. For instance: usage of proportional fonts for various purposes, having a sample print preview, and changing scale of printing are possible;
27. it is possible to modify existing print formats just before printing, and even save it for future purpose;
28. it is possible to print in ASCII formats;
29. it is possible to copy selected records to the Microsoft Windows clipborad in either ASCII or RTF, and hence paste in other applications for further processing;
30. it is possible to save a print setup that users have arranged in a session for future purpose.
31. modifying a data entry worksheet is improved. For instance, there is no need of making a room for inclusion of a new field among existing ones;
32. display of records using different proportional or non-proportional or a mixture of the two, colors, and font attributes as well as applying of these to only selected elements in a display/print format using braces (the *stack*) is possible;
33. users can write interactive display formats which enables them to perform some operations such as displaying images and/or playing some sounds/music while displaying records, i.e., multimedia facility is included; and
34. hyper-textual error debugging and suggestion is included, i.e., whenever users make a mistake while writing a display/print formats, clicking the hypertext term takes the insertion pointer (or the cursor) around the formatting commands that causes the error. In

addition as explained before, after trying to fix the error it is possible to see the modified format result on the same screen as the records.

Finally it can be concluded that WINISIS incorporates a number of features in addition to those which are in Micro-CDS/ISIS and in other text retrieval packages.

Considering the compatibility of files created by these two packages, however, database files and applications developed in Micro-CDS/ISIS can easily be ported by copying the files onto appropriate directories.

## **5.2 Recommendations for Applications and Improvement**

Almost all text retrieval facilities which are available in Micro-CDS/ISIS are found in WINISIS. Hence, the software can be observed as a continuation of Micro CDS/ISIS in Windows platform. Taking into considerations the *old* facilities, the newly incorporated facilities, and the strength of the underlying operating system - Microsoft Windows - it is possible to apply WINISIS in the following operations:

1. for library catalogue databases;
2. for management of conventional bibliographic and text databases;
3. since WINISIS has very good printing capabilities, it can be used for maintaining data and printing them out in different formats, which may be directories for publications or labels for postage;
4. for management of referral and factual databases relating to persons, institutions, projects etc. and make it more descriptive by inclusion of pictures, and images;
5. since saved search results can be used for current search session, WINISIS can easily be adopted for SDI and CAS services; and

6. since sound and image files can be incorporated into fields, applications that need multimedia facility such as :

- multimedia catalogues
- museum databases
- flora and fauna databases etc.

are possible.

In addition, since UNESCO is trying to increase restrictions of field/record size to 32 KB and enabling creation of multi-gigabyte databases, there can be possibilities for applying WINISIS as a full-text retrieval package, to some extent in the near future.

To improve the usability and functionality of WINISIS from the beginning, this study suggests the following improvements in the future versions:

- inclusion of vocabulary control devices;
- support of menu exits; and
- inclusion of full hypertext browsing and navigation facilities.

## BIBLIOGRAPHY

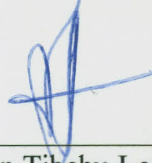
1. Abebe Rorissa, et al. ABNCD+: a prototype for an integrated information storage and retrieval system. Microcomputers for Information Management. 9(3).1992.
2. Buxton, A. and Hopkins, A. 1994. The CDS/ISIS Handbook. London. Library Association Publishing.
3. Chowdhury G.G. 1996. Text Retrieval Systems in Information Management. India. New Age International Publishers.
4. Citroen C. L. 1989. Microcomputers software for information storage and retrieval: how to make a well-founded choice. Program 23(3). 141-150.
5. ECA, PADIS. 1988. The Use of Microcomputers in Documentation and Information Centres in Africa. Fifth Session of the Joint Conference of African Planners, Statisticians and Demographers held at Addis Ababa, 15-20 March.
6. ECA, PADIS. 1990. The Use of Microcomputers in Documentation and Information Centres in Africa. Sixth Session of the Joint Conference of African Planners, Statisticians and Demographers held at Addis Ababa, 21-28 March.
7. Harter S. P. 1986. Online Information Retrieval: Concepts, Principles, and Techniques. LA. Academic Press.
8. Kenneth S. B. and Bruce B. A. 1988. Records design and methods. A process approach. USA. Mayfield publishing company.
9. Kimberly, R. and Rowley, J. 1990. Text Retrieval: a directory of software. England: Gower Publishing company Limited: 1-15.

10. Lancaster F.W. 1986. Information Retrieval Systems: Characteristics, Testing, and Evaluation. USA. John Wiley & Sons Inc.
11. Lancaster, F.W. 1986. Vocabulary Control for Information Retrieval. 2nd ed. Arlington: Information Resources Press.
12. Lauro, A.P. 1990. Manual for Preparing Records in Microcomputer-based Bibliographic Information Systems, Ottawa:IDRC
13. Mahamood K.M. 1996. An Introduction to Library Automation and Management Program (LAMP). INFOISIS 2(1) 7-16.
14. Molla Hunegnaw. 1993. Enhancement of User-Friendliness of Micro-CDS/ISIS for Search and Retrieval. M.Sc. thesis. Addis Ababa University, School of Information Studies for Africa. Addis Ababa
15. Muyoyeta Simui. 1995. A Micro-CDS/ISIS Information System for supporting Research in the Institute For African Studies, University of Zambia: Msc. thesis. Addis Ababa University, School of Information Studies for Africa, Addis Ababa.
16. Naber .G. 1996. DRAiN: A Computerized Library Management System Using CDS/ISIS Software. INFOISIS 2(2) 45-50.
17. Neelameghan, A. A Home-grown non-bibliographic object oriented databses:NIT'92. Fifth International conference on new information techniques held at Hong Kong, 30-Nov - 2 Dec. 1992.
18. Neelameghan, A. 1991. Concept categorization and knowledge organization in specialized databses: A case study . Int. Classification. 18(2).
19. Nieuwenhuysen, Paul. 1990. Computerized storage and retrieval of structured text information: CDS/ISIS Version 2.3. Program.25(1):1-18

20. NSTCDIC. 1991. Proceedings of the national policy seminar on information systems and services. 27-29 November. Addis Ababa
21. Perera, P. 1992. Micro CDS/ISIS: a critical appraisal of its search interface. Program 26(4): 373-386
22. Rowley, E. R. 1988. Organising knowledge. An introduction to information retrieval. Gower publishers.
23. Sisay Fisseha. 1995. Online Tutorial for Micro CDS/ISIS Software. M.Sc. thesis. Addis Ababa University, School of Information Studies for Africa. Addis Ababa.
24. Siverts, E.G. et al. 1991. Software for information storage and retrieval tested, evaluated and compared: part I - General Introduction. The Electronic Library 9(3). 5-19.
25. Siverts, E.G. et al. 1991. Software for information storage and retrieval tested, evaluated and compared: part II - Classical retrieval systems. The Electronic Library 9(6).300-316.
26. Siverts, E.G. et al. 1992. Software for information storage and retrieval tested, evaluated and compared: part III - End-users software. The Electronic Library 10(1).5-18.
27. Spinello R.A. 1994. Full Text Retrieval Systems for Document Management. Information Management, Strategy, Systems, and Technologies 2:3-17. Boston. Averbach Publications.
28. Steve Smithson. 1994. Information Retrieval Evaluation in Practice: A case Study Approach. Information Processing and management 30(2) 205-220.
29. Unesco.1989. Mini-Micro CDS/ISIS(Reference Manual Version 2.3), Paris Unesco.

## DECLARATION

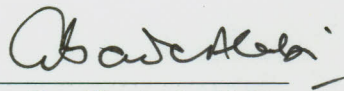
This thesis is my original work and has not been presented for a degree in any university.



---

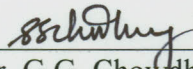
Samson Tibebe Legesse  
May 24, 1997

The thesis has been submitted for examination with approval as university advisor.



---

Dr. Gbade A. Alabi  
Advisor  
May 24, 1997



---

Dr. G.G. Chowdhury  
Advisor  
May 24, 1997

