

IMPLEMENTING WWWISIS FOR PROVIDING WEB ACCESS TO BIBLIOGRAPHIC DATABASES*

By Francis Jayakanth

Abstract: CDS/ISIS, an advanced non-numerical information storage and retrieval software was developed by UNESCO. With the emergence of WWW technology, most of the information activities are becoming Web-centric. Libraries and information providers are taking advantage of these Internet developments to provide access to their resources/information on the Web.

A number of tools are now available for publishing CDS/ISIS databases on the Internet. One such tool is the WWWISIS Web gateway software, developed by BIREME, Brazil.

This paper illustrates porting of sample records from a bibliographic database into CDS/ISIS, and then publishing this database on the Internet using WWWISIS.

Keywords: CDS/ISIS, WWWISIS, Web Access

Introduction

With the emergence of the Internet, the global connectivity of computer systems and access to electronic information have become a reality. The Internet has progressed magnificently to provide many user-friendly tools like Gopher, WAIS, E-mail, Telnet, FTP and World Wide Web (WWW) for information processing and access. The WWW, which is integrating all other access tools also provides a very convenient means for publishing and accessing multimedia, hypertext-linked documents stored in computers spread across the world. With the emergence of WWW technology, most of the information activities are becoming web-centric. Once the information is on the Web, one can access the information from any part of the world. Also, a web browser like Netscape or Internet Explorer would be used as a common user interface for accessing information/database. Libraries and information providers are taking advantage of these Internet developments to provide access to their resources/information on the Web.

CDS/ISIS is a menu driven generalized information storage and retrieval system. It allows you to build and manage structured non-numerical databases i.e. databases whose major constituent is text. The search language of CDS/ISIS is based on

* Paper presented for the GLOBAL 2000 Worldwide Conference on Special Librarianship, 16-19 October 2000 in Brighton, UK

Boolean algebra, which provides a convenient way of expressing logical operations between database fields

A number of tools are now available for making CDS/ISIS database accessible on the Internet/web. Some of these tools are 1) the WAIS-ISIS Server. 2) the WWWISIS Server 3) the IQUERY Server.

This paper explains porting of sample records from a bibliographic database into a CDS/ISIS database & then publish this CDS/ISIS database on the Web using the freely available CDS/ISIS Web gateway software called WWWISIS developed by BIREME, Brazil. WWWISIS versions for several operating system (OS) platforms are available I have used PC-Windows edition of WWWISIS 3.2c with Windows'95 and OMNIHTTPd Web Server software.

About CDS/ISIS

CDS/ISIS, an advanced non-numerical information storage and retrieval software is developed and maintained by UNESCO. It was developed to satisfy the need expressed by many institutions, especially in the developing countries, to streamline their information processing activities by using modern (and relatively inexpensive) technologies. The software was originally based on the mainframe version of CDS/ISIS. At present, UNESCO maintains three fully compatible versions of the software for different hardware/software platforms – MS-DOS, Windows & Unix.

The major features of the CDS/ISIS software are:

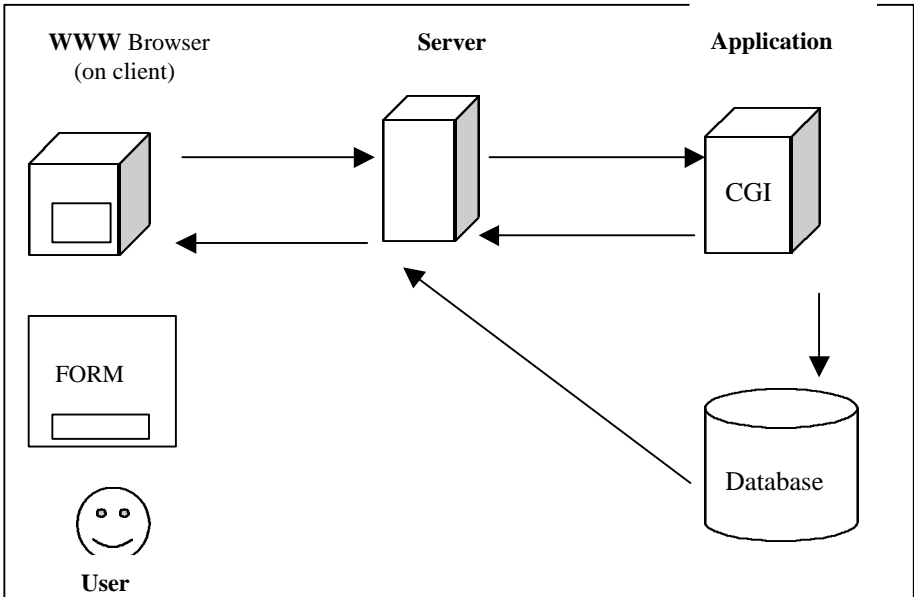
- the handling of **variable length** records, fields and **sub fields**, thus saving disk space and making it possible to store greater amounts of information;
- the handling of **repeatable fields**;
- a **database definition** component allowing the user to define the data to be processed for a particular application;
- a **data entry** component for entering and modifying data through user-created data base specific worksheets;
- an information retrieval component using a powerful **search language** providing for field-level and proximity search operators, in addition to the traditional and/or/not operators, as well as **free-text searching**;
- a powerful **sort and report** generation facility allowing the user to easily create any desired printed products, such as catalogues, indexes, directories, etc.;

- a data **interchange** function based on the **ISO 2709** International Standard used by leading data base producers;
- an integrated application **programming language** (*CDS/ISIS Pascal* and the *ISIS_DLL*), allowing the user to tailor the software to specific needs;
- functions allowing the user to build pseudo-relational databases.

Database Access on the WEB

The World Wide Web (WWW) is a client-server distributed hypermedia system initially developed at the European Laboratory for Particle Physics (CERN) in Switzerland, where it was envisaged as a means for sharing information amongst physicists. The process of providing Web access to database involves the interaction of Web server, Web browser, CGI scripts/programs and HTML forms. The use of HTML forms on the client side and CGI applications on the server side comprises the key mechanism to collect search requests and/or data entry and submit it to the database. Figure 1 depicts this mechanism.

Figure 1: Client/server interaction.



WEB Server

Web server is a computer system that provides Web services and Web pages to Intranet or Internet users. A HTTP (Hypertext Transmission Protocol) daemon program should be running on the web server to listen and respond to client's requests. A daemon is a process that runs in the background waiting for a client request. When the request is received, it processes it and then keeps waiting for other requests. A web server usually has a dedicated connection to the Internet/Intranet. Web server software is available for different OS platforms. The best way to access WWW is by using GUI based client programs like Netscape, Internet Explorer or Mosaic. These client programs are called web browsers. The GUI based browsers can interpret image and audio files too.

WEB Server and WEB Browser Interaction

A Web browser acts as an interface between a user and the Web server. Web browsers work by connecting over the Internet to the web servers, requesting specific documents and then formatting the documents they receive, for viewing on the local system. The Web pages on a Web server are text files encoded with HTML tags. These documents are usually named with the suffix "html" and are interpreted by the Web browsers that receive them. If you click the Open button on your browser and type any URL (eg:<http://www.ncsi.iisc.ernet.in/>), the browser generates a GET request, connects to the server at the specified URL and waits for a response. The server returns a response and closes the connection. This is the basic browser and server interaction.

Common Gateway Interface (CGI)

CGI is an essential tool for creating and managing interactive Web sites.. A Web server program with CGI support can invoke an application program, pass user-specific data to the program. The application program then processes the data received from the browser and the server then passes the program's response back to the Web browser. CGI is not a programming language in itself, but is a mechanism by which Web browsers interact with programs residing on the Web servers. Web servers use a special directory, commonly named **cgi-bin** to store CGI programs/scripts.

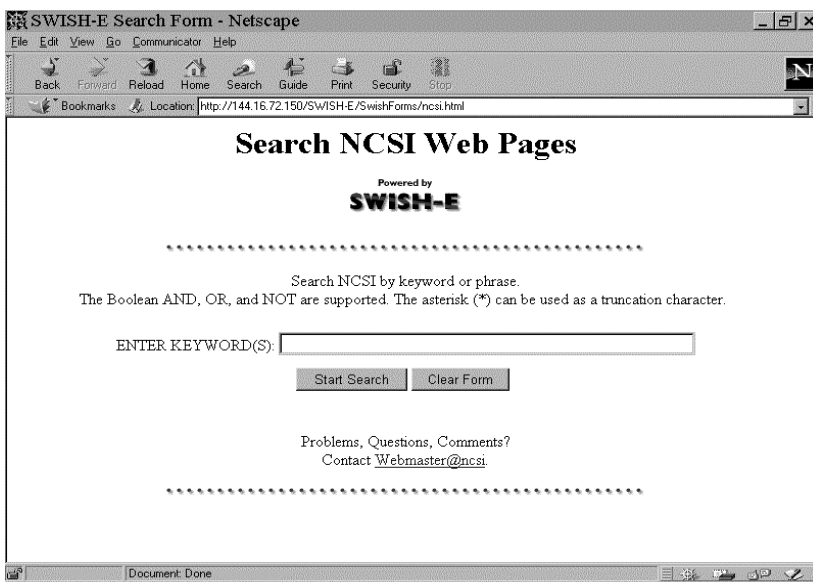
The Web server knows that the files stored in the cgi-bin directory are to be executed rather than simply sending it to the Web browser for display. CGI programs can be written in a variety of programming/scripting languages like MS-DOS batch file, Unix shell scripts, Perl scripts or higher level programming languages like BASIC , C or C++ etc.

HTML Forms

Forms are part of HTML that allows a user to supply information. Forms are used to accept inputs from a user. A simple form is shown in figure 2.

- This sample form has one **Text input field**: for entering keywords.
- The form also has a **submit button**. The user can click this button after filling the form. CGI program will act upon the information contained in a form only when the form data has been submitted to the CGI program. Clicking on the form's **Clear Form** button **resets** the form back to its default values.
- The form can also have **Multiple-choice fields** enabling users to select a specific item or items from a list.

Figure 2: A Sample HTML form.



The screenshot shows a Netscape browser window with the title "SWISH-E Search Form - Netscape". The address bar displays "http://144.16.72.150/SWISH-E/SwishForms/ncsi.html". The main content area features the heading "Search NCSI Web Pages" and "Powered by SWISH-E". Below this, there is a search instruction: "Search NCSI by keyword or phrase. The Boolean AND, OR, and NOT are supported. The asterisk (*) can be used as a truncation character." A text input field is labeled "ENTER KEYWORD(S):". Below the input field are two buttons: "Start Search" and "Clear Form". At the bottom, there is a link for "Problems, Questions, Comments?" and a contact email "Webmaster@ncsi".

WEB Interface to CDS/ISIS Databases

Following are some of the tools, which facilitate in providing Web access to CDS/ISIS databases.

- **WWWISIS** is one of the most popular softwares, which acts as a server for CDS/ISIS databases in a WWW environment. It was developed by BIREME and PAHO and is freely available on the Brazilian archive host; <ftp://ftp.brme.bireme.br/wwwisis> or <http://www.bireme.br/wwwisis2.htm> WWWISIS documentation is available at <http://www.bireme.br/wwwisis1.htm>. Searching, formatting and field updating languages of CDS/ISIS are supported by WWWISIS. WWWISIS is available for different O/S platforms like Windows'95, Linux, Sun Solaris etc. There are number of sites which are making use of WWWISIS to provide web access to their CDS/ISIS databases. The Middlesex University in UK (<http://158.98.53.13/>) is an excellent example usage site for WWWISIS.
- **IQUERY** is commercial software designed by Alejandro Ferreira Guido of Uruguay, to run under three platforms, Linux, SCO and DOS. The system was designed to act as an interactive search engine for Mini/Micro CDS/ISIS databases, it was developed taking the directives showed in the ISIS 3.0 version. The software can run in three different modes, as an interactive command shell, as a CGI to run under WWW servers, and as an interpreter taking the statements from a file. <http://www.gti.net/reu/iquery/>
- **WAISIS** freeware server distributed at the ftp site of GARR in Italy. <ftp://ftp.nic.it/pub/WAIS/WAIS-ISIS/>
- **Websis**, a CGI application written in C to search CDS/ISIS databases over the Internet using a Web interface. Websis was developed by the Infocentre and is distributed free of cost. <http://www.theinfocentre.com/>
- **JAVASIS** 3.0 consists of two Java applications: a Java Client and a Java Server. It is available for Windows 9'x and NT, Linux and other OS platforms. <http://web.tiscalinet.it/javaisis>

WWWISIS and its Operations

WWWISIS is a system specially designed to act as a server for CDS/ISIS databases in Intranet or Internet World Wide Web client server environment. It was developed and is being maintained and distributed by BIREME, the Latin America and Caribbean center on Health Sciences Information in Brazil. The current version of WWWISIS is version 3.3. It provides functions for searching, formatting and data entry operations over CDS/ISIS databases. It operates as a

CDS/ISIS server through the WWW Common Gateway Interface (CGI). The actual operation is driven by parameters.

Features of WWWISIS

- A key feature of this package is its capability to dynamically generate execution parameters, based on data sent to the CGI process from the WWW browser.
- WWWISIS may be called directly from the command prompt or as a command line inside a CGI script. An example of direct call to `wwwisis` in the command prompt:

```
wwwisis db=ccodw bool="plants and soil" pft=v200/
```

Here WWWISIS is called with three mandatory parameters, namely *db*, *bool* and *pft*. With the *db* parameter, name of the database (`ccodw`) is passed. With the *bool* parameter, the search expression ("plant and soil") is passed. With *.pft* parameter, the fields that are to be displayed from the retrieved record/s are included

- CDS/ISIS searching, formatting and field updating languages are supported by WWWISIS.
- Search expressions can be gathered through HTML forms, passed via CGI to WWWISIS and the results formatted and delivered to the client.
- WWWISIS output to the browser is generated via CDS/ISIS formatting language.
- WWWISIS uses CISIS interface library developed by BIREME to support CDS/ISIS searching, formatting and field updating features.
- No installation procedure is involved. When used as a Web gateway for CDS/ISIS databases, it should just be copied to the `cgi-bin` directory configured for the Web server.

Interfacing CDS/ISIS Database with WWWISIS

WWWISIS is freely available for different OS platforms. For the work reported in this paper, I have used PC-Windows version of WWWISIS 3.2c on a Pentium II system with Windows 95 OS and OMNIHTTPd Web server.

To load records into a CDS/ISIS database, a Field Definition Table (FDT) is created using a built function of CDS/ISIS. The FDT for the database, `CCODW` is as shown in the table below.

Field Definition Table (FDT)

Tag	Name	Length	Type	Rep	Delimiter
10	Author	200	X	R	
20	Title	200	X		
30	Source	200	X		
35	Year	25	X		
40	Kword	250	X	R	
50	Discipline	30	X		
60	Doctype	30	X		
70	Lang	20	X		
80	Address	200	X		
90	ISSN	9	X		
100	Publisher	200	X		
110	Abstract	1650	X		

Having defined the FDT (database structure), one can populate the database by loading records. The records can be loaded in any of the following three ways:

- Manually enter the data using worksheets, created during database design,
- Import records from a file, which conforms to ISO 2709 file format, or
- Use the CDS/ISIS Pascal programming interface to load records from a file.

I have written a Pascal program to load around 1000 sample bibliographic records into CDS/ISIS database. The program listing and a sample record are given in the appendix I and appendix III respectively.

The CDS/ISIS database thus created can be easily interfaced with WWWISIS to provide access to this database on the Internet. The database can also be queried using WWWISIS in the command prompt.

WWWISIS db=ccodw bool=elastic pft=mf, v20/

This command would search the CDS/ISIS database, *ccodw* for the term *elastic* & then display the record number (mf) and the *title* of all the references that match the search term, *elastic*. A sample record is given below:

000018 The influence of longitudinal vibrations of an elastic wall on natural frequencies of rectangular acoustic resonators

The following is a call to WWWISIS using the *cgi* parameter:

WWWISIS cgi=@ccod.cgi

The cgi parameter specifies the file, ccod.cgi, which contains a format specification that will generate the actual execution parameters, one per line.

```
'db=ccodw'  
bool=elastic'  
'pft=@ccodw.pft'
```

This call to WWWSIS with cgi parameter will retrieve the same set of records as with previous call. ccodw.pft is a CDS/ISIS display format file

However, it is through WWW CGI that WWWSIS acts as CDS/ISIS database server for Internet/WWW. CGI is an integral part of WWW servers that provides a mechanism to call other programs. Thus, requests from a WWW browser to a Web server may be sent to program or a script, which in turn may retrieve data from a database and send the output back to the Web server, that transfers it to the Web browser, usually in HTML format.

The most common way to gather data in a Web browser is through forms. Form is an element of the HTML language that allows the specification of input elements, such as field text, list box, check box, radio box, etc. A CDS/ISIS search, for example, may be typed into an input element of the type text and search limits may be collected via elements type list, check or radio boxes. Figure 5 is screen shot of a HTML form used for searching the CDS/ISIS database.

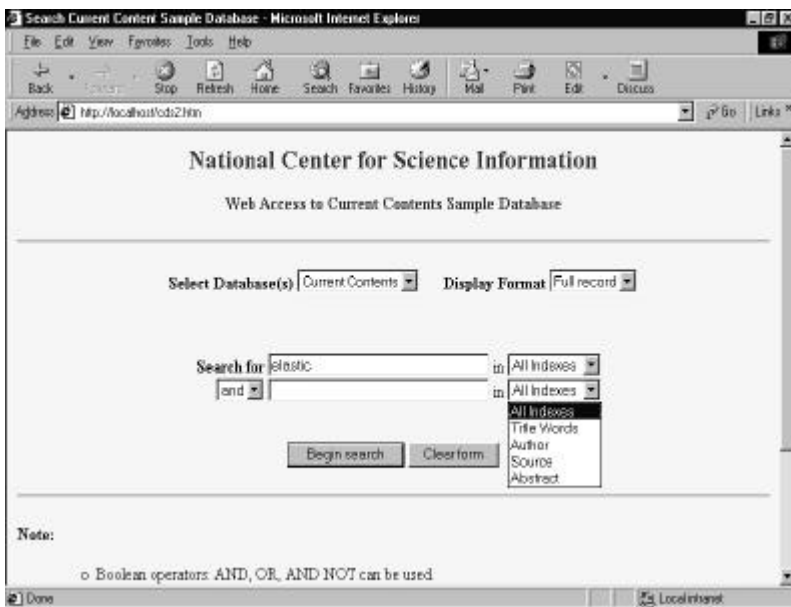


Figure 3: Screen shot of a HTML form used for searching CDS/ISIS database

Every FORM input element has two basic attributes associated; the name that identifies the input element instance and the value assigned to it. A key attribute of the FORM element is ACTION, which allows the specification of the name of the program or script to be activated via CGI.

To interface the sample database, ccodw, with WWWISIS in a WWW client/server environment, an HTML form as illustrated above is designed. The following is a sample specification of this form in HTML: The complete HTML code for the form is given in the appendix II

```

<html>
...
...
<form method=POST action="/cgi-bin/ccod.bat">
<select name="db^t5000">
  <option value="/isis/cds">CDS</option>
  <option value="/isis/data/ccodw">Current Contents</option>
</select>
<input type=text name="sterm^t5001" size="30" value=""> in
<select name="index^t5021">
  <option selected value="">All Indexes</option>
  <option value="/(24)">Title Words</option>
  <option value="/(70)">Author</option>
  <option value="/(30)">Source</option>
  <option value="/(110)">Abstract</option>
</select>
...
...
</form>
</html>

```

When WWWISIS starts, it first creates in memory an CDS/ISIS master file record called CGI-environment virtual record. This record is used by WWWISIS to store data sent by the Web server to the CGI process.

When name contains the sub-field identifier ^t followed by a number, this will be the data field tag (<input type=text name="sterm^t5001" size="30" value="">) number in which value will be stored in the CGI-environment virtual record. The search term entered in the form will be stored in data field tag 5001 and the name of the database will be stored in the field tag 5000.

When this form is submitted, the script file or the program file specified in the ACTION attribute of FORM element (<form method=POST action="/cgi-

`bin/ccod.bat">)` gets executed. Here, the batch file, CCOD.BAT gets executed. The batch file contains just the following one line:

WWWISIS CGI=@CCWEB.CGI

WWWISIS is being called with CGI parameter. CCWEB.CGI file contains following specifications:

```
'bool=',v5001/  
'db=',v5000/  
'pft=@cds1.pft/'
```

So, WWWISIS gets values for all the three mandatory parameters, namely db, bool and pft when the form is filled and submitted, WWWISIS would search the database, and then display the results according to the specifications in the pft parameter. Figure 4 illustrates the interaction of Web browser, Web server, CGI script/program, WWWISIS and the CDS/ISIS database.

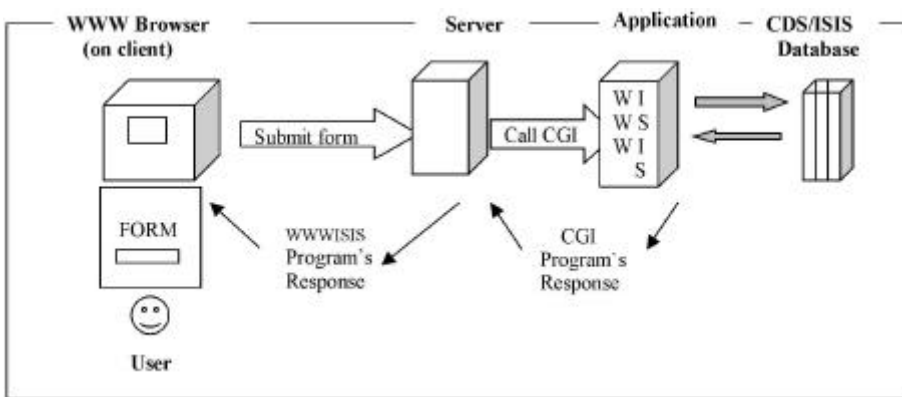


Figure 4: WWWISIS and its operation

Having received the three mandatory parameters, WWWISIS searches the CDS/ISIS database and the retrieved records are sent to the Web browser for display after formatting the records according to the specifications in the pft file. A screen-shot of records being displayed through a Web browser is shown figure 5.



Figure 5: Screen-shot of record display.

Conclusions

WWWISIS is a very convenient tool for providing Web access to CDS/ISIS databases. The task of interfacing CDS/ISIS databases with WWWISIS is quite simple. Any person with a working knowledge of HTML coding and basic knowledge of working of Web server software should be in position to publish the CDS/ISIS databases on the Web.

References

1. CDS/ISIS reference manual, UNESCO, 1989
2. Course material of Workshop on Indexing and Search Engines for the Web (WISE), 12-13 Jan 1999, NCSI, IISc, Bangalore-560 012.
3. Vatnal, R M, Project report on Circulation Management Module with CDS/ISIS for NCSI Library Books, NCSI, IISc, 1994-95
4. WWWISIS reference manual downloaded from BIREME site (URL <http://www.bireme.br/wwwisis2.htm/>)
5. Francis Jayakanth & Jayashree S, A Tutorial on Integrating CDS/ISIS Databases with World Wide Web. Information Today and Tomorrow, Vol. 18, Iss. 4, December 1999

Appendix I

program loadtit:

```
var
i,j,line_len,blank_pos,blank_pos2,trecs,no_recs,rec_cnt,tagval : real;
srce,delmtr,infile,outfile,inp_line,fldname,fldval,fldname1,next_rec : string;
delmtr_1 : string;

procedure delimiter(delmtr_1 : string;tagval : real);

var
k,l,m : real;

begin
k:=0.0;
l:=1.0;
m:=9.0;
fldname:="";
while k<= blank_pos2 do
begin
j:=position(fldval,delmtr_1,l);
if j=0 then
begin
fldname:=substr(fldval,l,blank_pos2);
k:=line_len+1;
i:=fldadd(tagval,nfields+1, fldname);
{ writeln('fldname1:',fldname);}
fldname:="";
end
else
begin
fldname:=substr(fldval,l,j-1);
l:=j+2;
i:=fldadd(tagval,nfields+1, fldname);
k:=j;
{ writeln('fldname2:',fldname,'j:',j);}
fldname:="";
end;
end;
end;

begin
clear;
rec_cnt:=0.0;
write('Pl. Enter Input File Name : ');
readln(infile);
{ write('No. of Recs. to load : ');
readln(no_recs);
assign('inp',infile);
readln(inp,inp_line); {read from input file}
```

```

while rec_cnt <= no_recs do
while not eof(inp)
begin
i:=newrec;
next_rec:=false';
while (next_rec <> 'true') do { loop till the next record }
begin
readln(inp,inp_line);
fldname:=substr(inp_line,1,6);
if fldname='Record' then
begin
update;{ update the database }
next_rec:=true';
writeln('New Rec:');
rec_cnt:=rec_cnt+1;
end
else
begin
line_len:=size(inp_line);
blank_pos:=0.0;
blank_pos2:=0.0;
blank_pos:=position(inp_line,chr(09),1); { find the first occurrence of a tab char }
{ writeln('line len:',line_len,' blank_pos:',blank_pos, line_len-blank_pos);}
blank_pos2:=line_len-(blank_pos+1);
fldname:=substr(inp_line,1,blank_pos-1); { sub str till the first tab }
fldval:=substr(inp_line,blank_pos+1,blank_pos2);
if fldname='Authors' then
begin
delimiter(', ',10);
end;
if fldname='Title' then
i:=fldadd(20,nfields+1,fldval);
if fldname='Full source' then
begin
j:=0.0;
j:=position(fldval,', ',1);
srce:=substr(fldval,1,j-1);
i:=fldadd(30,nfields+1,srce);
srce:=";
srce:=substr(fldval,j+2,blank_pos2-(j+1));
i:=fldadd(31,nfields+1,srce);
j:=0.0;
srce:=";
end;
if fldname='KeyWords Plus' then
begin
delimiter(', ',40);
end;
if fldname='Author keywords' then
begin
delimiter(', ',35);

```


Appendix III

A Sample bibliographic record.

Authors JDH Paul, F Appel, R Wagner
Title The compression behaviour of niobium alloyed gamma-titanium aluminides
Full source Acta Materialia, 1998, Vol 46, Iss 4, pp 1075-1085
KeyWords Plus OXIDATION BEHAVIOR; NB
TGA/Book No. YZ762
Discipline Metallurgy
Document type Article
Language English
Address Paul JDH, Gkss Forschungszentrum Geesthacht GmbH, Inst Mat Res, Max Planck Str, D-21502 Geesthacht, GERMANY
ISBN/ISSN 1359-6454
Publisher Pergamon-Elsevier Science Ltd, The Boulevard, Langford Lane, Kidlington, Oxford, England OX5 1GB
Abstract The underlying mechanisms behind the reported high strength of titanium aluminide alloys containing a large addition of niobium has been investigated by determining the flow stresses and activation parameters of plastic deformation. It has been found that alloys such as Ti-45Al-10Nb (at.%) and Ti-45Al-5Nb have 1.25% flow stress values in compression of >800 MPa at room temperature and > 500 MPa at 1173 K. When compared with values from a more conventional alloy, Ti-47Al-2Cr-0.2Si, they represent a considerable increase in strength. However, the activation volumes after 1.25% deformation are very similar to those of conventional alloys, particularly up to 973 K. This suggests that athermal dislocation mechanisms are responsible for the increased flow stress of the niobium containing alloys. By comparing the properties of the niobium containing alloys with different binary alloys it has been shown that the high strength is solely a result of the reduced aluminium content and that niobium plays no role in strengthening or work hardening.

Francis Jayakanth
National Centre for Science Information
Indian Institute of Science
Bangalore – 560 012
India
E-mail : franc@ncsi.iisc.ernet.in