Application of WWWISIS: Web-based CDS/ISIS Bibliographic Databases for Journal Articles

Sridhara B., Veena Makhija and Ajay Kumar Pandey

Defence Scientific Information and Documentation Centre, Delhi-110 054

ABSTRACT

With the advent of database connectivity to the web servers, it is now possible to generate dynamic web pages using database as the back-end and a web browser as the front-end. This article focuses on implementation of WINISIS and WWWISIS on Microsoft Windows XP using professional Microsoft's web server (Internet Information Server 5.1) for generating an information retrieval system for defence science and technology articles database developed by Defence Scientific Documentation & Information Centre, DRDO. The database provides an instant access to information on defence-related articles to DRDO scientists. The paper also focuses on generating a current awareness service, Current Literature in Defence science and technology using Defence Science and Technology Articles Database.

Keywords: CDS/ISIS database, WINISIS, WWWISIS, information retrieval system, scripting language

1. INTRODUCTION

A traditional computer application can be created to work on a web browser. The main advantage is the ability to share data and to allow user access to the system from anywhere, i.e., home, office, or internet café. Since the first diffusion of the world wide web (www), concerted effort have been made for the creation of interfaces for using the existing library databases, particular those based on CDS/ISIS. When a user clicks on a web page, the web server determines which elements are needed and then searches for those elements in the database. The required information is then transferred from the database to the web server, thus making database a web-enabled database. In contrast, web applications dynamically generate a series of web pages in a standard format supported by common browsers such as HTML/XHTML with client-side scripting language. Librarians and information providers are taking advantage of these technologies to provide access to their resources/information on the web.

All the versions of the CDS/ISIS software were developed before the world wide web technology came into being. During the early 1990, with the growing popularity of the web, several ISIS developers tried to enhance CDS/ISIS with a web interface. The following developments made CDS/ISIS usable on the web:

✖ Development of common gateway interface (CGI) program for web servers to drive standard CDS/ISIS programs for UNIX

✖ Development of a CDS/ISIS server like wwwisis or wwwi32.exe by combining CGI and CDS/ISIS features for UNIX and Windows NT

✖ Development of CDS/ISIS server for ISIS-Free WAIS (Z39.5) and interface with the web through a web-enhanced WAIS client for UNIX

Some of the tools like WAIS-ISIS server, WWWISIS server IQUERY server, and so on are now available for making CDS/ISIS database accessible on the web.

2. OBJECTIVE

Defence Scientific Information and Documentation Centre (DESIDOC) under Defence Research and Development Organisation (DRDO) is one of the premier centre of excellence in defence scientific information processing and dissemination. The Defence Science Library (DSL) of DESIDOC is a central resource for providiing scientific information to the scientists working in various DRDO labs/estts and to the top management at DRDO HQrs. The core competence of DESIDOC is the development of information resource, design and development of application software for library automation, development of content, digitisation, multimedia and web-based software applications, database creation, and technical editing and publishing.

To fulfil the objective to provide the resources of the DSL to the top management and the scientist of DRDO, a bibiliographic database, defence science and technology articles database (DEFSTAD), from core journals on defence science and technology subscribed by the DSL, has been developed using WINISIS.

3. WWWISIS SOFTWARE

WWWISIS is a tool to serve as a web-based CDS/ISIS server. WWWISIS brings out many interesting points about the development of CDS/ISIS software package for internet or intranet users to make databases available for global access on net. It is a system developed and distributed by BIREME (the Latin American & Caribbean Centre on Health Sciences Information), PAHO (the Pan-American Health Organisation), and WHO (the World Health Organisation).

WWWISIS is a program specially designed to act as a server for ISIS databases in a web client-server environment through the WWW CGI of hypertext transport protocol (HTTP). Starting from version 4.0, also known as WXIS, it is driven by an XML-based script language called IsisScript. IsisScript output to web browser can be generated using ISIS formatting language. It is possible to develop a fully graphical, client-server search interface for ISIS databases either in internet or intranet environment, as it is a part of the ISIS software family.

It is a complete C library for CDS/ISIS and performs nearly all CDS/ISIS functions. Search specifications can be gathered through HTML web forms, and then passed via CGI to WWWISIS, the results are formatted and delivered back to the client. Similarly, it is possible to implement database-updating operations. By default, WWWISIS presupposes a multiuser environment and HTML can be customised through CGI.

Salient features of WWWISIS are:

✖ WWWISIS operations are driven by parameters that can be generated dynamically.

✖ Mandatory parameters are: database name (db), search expression (bool) and format specification for display(pft).

✖ The records of an ISIS database can be dynamically converted to HTML documents, that is to say, the results of a search or a partial master file can be retrieved and sent back to the web client using ISIS display formatting language.

✖ WWWISIS may be called directly or as command line inside a CGI script.

- CDS/ISIS searching, formatting, and field updating languages are available.

- Search specifications gathered through forms are passed via CGI to WWWISIS, and the results are formatted and delivered to the client.

3.1 WWWISIS Scripting Language

Scripts are computer-program codes written in relatively simple programming language. It is nothing but files that initiate routines like generating web pages dynamically in response to user input. A client-side script is a program that may accompany an HTML document or be embedded in it. The program executes on the client's machine when the document loads, or at some other time such as when a link is activated. HTML's support for scripts is independent of the scripting language.

Scripts offer authors means to extend HTML documents in highly active and interactive ways. For example:

- Scripts may be evaluated as document loads to modify its contents dynamically.

- Scripts may accompany a form to process input as it is entered. Designers may dynamically fill up parts of a form based on the values of other fields. They may also ensure that input data conforms to predetermined ranges of values, that fields are mutually consistent, etc.

- Scripts may be triggered by events that affect the document such as loading, unloading, element focus, mouse movement, etc.

- Scripts may be linked to form controls (e.g. buttons) to produce graphical user interface elements

4. COMMON GATEWAY INTERFACE

WWWISIS was conceived and implemented to serve as an interface between ISIS database and web servers using CGI.

The CGI permits interactivity between a client and a host-operating system through a world wide web via HTTP. It is a standard for interfacing external applications with information servers, such as HTTP or web servers.

The CGI programs need to reside in a special directory called/cgi-bin so that the web server executes the program rather than just display it to the browser. CGI is not a language, it is a simple protocol that can be used to communicate between web forms and the program running on the server. A CGI program is preferably a script that can be written in any language that can read standard input, write to standard output, read environment variables, and allows these to be executed on the system, such as C/C++, Fortran, PERL, TCL Any Unix, shell Visual Basic and AppleScript.

WWWISIS static web pages are stored on the server in the root directory or in subdirectories of that directory. Typically the root directory is called wwwroot. Any file to be sent to the client, whether text pages, images, sound files, etc., need to be stored in wwwroot. When the web server software is configured, the root directory has to be specified and that determines what the website will consist of. The web server also recognises another directory, usually called cgi-bin (for CGI binary files). This directory contains program files and can run on the server by specifying these in the URL.

Briefly, the CGI programs collect request from the browser and send data back to the browser in HTML format.

5. WEB SERVER

Web servers are computers on the internet/intranet that host websites serving pages to viewers upon request. This service is referred to as web hosting. Every web server has a unique address so that other computers connected to the internet know from where to find it on the vast network. WWWISIS works with various web server software like Omni HTTPd, Apache, etc.

6. PROCESS OF DEFSTAD DEVELOPMENT

The process of development of DEFSTAD and Current Literature in Defence Science and Technology (CLDST) is described below.

6.1 Installation of Internet Information Server for DEFSTAD

Microsoft's web server Internet Information Server (IIS) 5.1, on Microsoft Windows XP Professional platform was used for the purpose. IIS 5.1 adds the power of web computing to Windows XP. It makes easy to publish information on the internet/intranet. It includes a broad range of administrative features for managing websites and web servers. With programmatic features like active server pages (ASP), one can create and deploy scalable and flexible web applications.

IIS 5.1 cannot be installed by default and can only serve 10 simultaneous client connections and does not leverage all the features of the server versions.CGI-BIN and WORK virtual directories were created with execute permissions to install IIS.

While the virtual content directory CGI-BIN contains all scripts (.xis files) as well as .htm forms, the virtual content directory: WORK contains files generated by DEFSTAD.

6.2 Journal(s) Selection

DSL is subscribing to more than 450 journals related to defence science and technology. Some of the core journals selected by the information scientists are included in DEFSTAD.

6.3 Scanning of Content Pages

Content pages of selected core journals are scanned. Relevant and most important records are then edited as per the bibliographic format. New records are added using database worksheet, and records which conforming to ISO 2709 file format are imported from the file.

6.4 Keyword(s) Indexing

Articles selected for DEFSTAD are indexed and maintained for the information storage and retrieval systems. An authority file has also been developed for maintaining the consistency using controlled vocabulary tools such as Thesaurus of Scientific and Technical Terms and DRDO Thesaurus of Scientific and Technical Terms.

6.5 Subject Headings

Certain subject headings have been identified for the selection of journal articles to meet the requirements of DRDO scientists. The subject headings are: Aeronautics, armaments, combat vehicles and engineering, electronics and instrumentation, defence policies, information and computer science, life sciences, management, materials, missiles, naval research and development, general science and technology.

6.6 Field Definition Table

CDS/ISIS validates each field entered according to the field type defined in the field definition table (FDT) of the database. It does not allows more than one occurrence of a non-repeatable field, and checks the validity of subfield identifiers. To load records onto the DEFSTAD, a FDT has been created using a built-in function of WINISIS software.Thirteen fields describing the bibliographic content of the articles have been identified for data input.

6.7 Field Select Table

CDS/ISIS allows selective creation of inverted files for each database, based on selected subfields or elements thereof. In addition, by specifying appropriate options, individual words, phrases or descriptors from selected fields can be extracted. Searchable elements have to be defined by means of a field select table (FST), which contains the fields to be inverted and the indexing technique to be used for each field. CDS/ISIS provides a virtually unlimited number of access points for each record through

the creation of a special inverted file, but may not automatically update the inverted file whenever records are added, modified or deleteted from a master file.

Following FST was used to generate the inverted file for DEFSTAD.

Field tag	Indexing technique	Format
200	4	mhl,v200
300	0	mhl,(v300/)
10	0	mhl,v10
490	0	mhl,v490
491	0	mhl,v491
492	0	mhl,v492
493	0	mhl,v493
442	0	mhl,v442
201	4	mhl,(v201/)
610	0	mhl,v610
620	0	mhl,(v620/)
22	0	mhl,v22
101	0	mhl,v101

6.8 Designing HTML Forms

An HTML form is a section of documents containing normal content, markup, special elements called controls (checkboxes, radio buttons, menus, etc.) and labels on those controls. Users generally complete a form by modifying its controls (entering text, selecting menu items, etc.), before submitting the form to an agent (web server, mail server, etc.) for processing. HTML defines the following control types: Buttons, checkboxes, radio buttons, menus, text input, file select, hidden controls, and object controls.

The most common method of collecting data from the browser is to use the HTML form. The form tag contains another element called Action where one can specify the action to be invoked. A brief syntax of the form tag is given in *Appendix 1*.

6.8.1 DEFSTAD Simple Search Form

A simple search form has a text input field for entering keywords, a submit button, and a reset button. The user can click the submit button after filling the form. CGI program understands the information contained in

the form when the form data is submitted. Clicking on the form's reset button resets the form back to its default values. The form also contains multiple-choice fields, thereof enables users to select a specific item or items from the list. Users can select a database, and display formats such as full record, citations, or titles, as shown in Fig. 1.
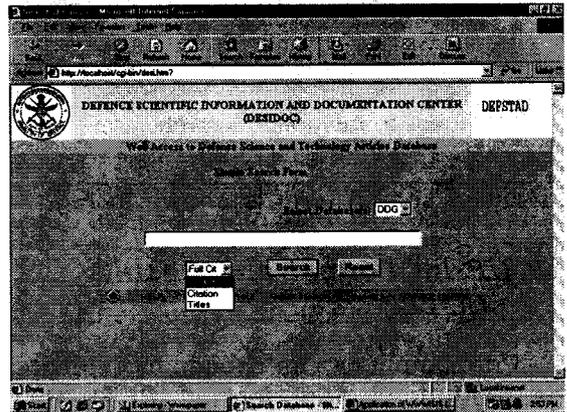

Figure 1. Simple search form

The $ sign for truncation, and the symbols * for AND, + for OR, ^ for AND NOT and . (dot) for adjacency or to form a phrase are used to frame search queries. Scripts and HTML code for simple search are given in *Appendices* 2 and 3.

6.8.2 DEFSTAD Advanced Search Form

A search term may be a word or a phrase under which user may be looking for information in the database. Multiple search terms can be combined with Boolean operators AND, OR, AND NOT, for retrieving relevant documents.

Advance search enables combining Boolean operators with the choice to specify the field in which the search is to be carried out. The fields namely, all indexes, title words, author, keyword, source, year, may be selected for searching the term from anywhere in the records.

The use of paranthesis resolves any ambiguity in the search expressions. An
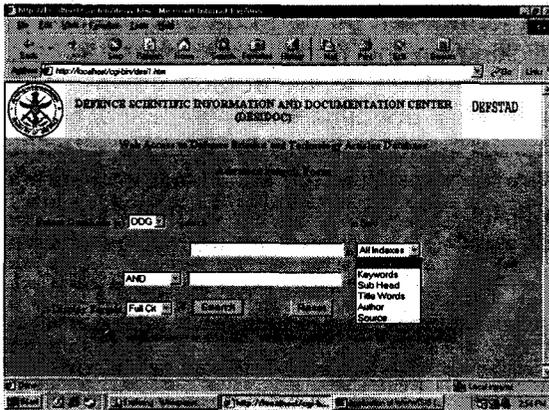
Figure 2. Advanced search form

Figure 3. Result display: Full record

advanced search form is shown in Fig. 2.

6.8.3 Print Format and Results Display

Print output can be chosen by selecting the predefined print format (PFT). PFT presentation of CLDST to retrieve bibliographic articles and keyword index is given below. CDS/ISIS allows modification in existing print format shortly before printing and can save the modified format.

At present, DEFSTAD have more than 26,000 bibliographic articles. To display results, three print formats have been designed and written in text editor and saved in WORK virtual directory. These formats are:

❑ PFT for Full Record

To display full record, fields chosen are title, author, source, volume number, issue number, year, and keywords. An advanced search PFT for full record is shown in Fig. 3. The script for full record is:

'<B><p>'mfn,'</B>
','Title:',v200, '<br>','Author:',v300+|;|,'<br>','Source:'v201, '<br>','Volume No:'v491,'<br>','IssueNo: 'v492,'<br>','Year:'v490,'<br>','Keywords: 'V620+|; |, '<HR>

❑ PFT for Citations only

To display citations only, fields chosen are title, author, source, citation, and keywords. An advanced search PFT for citation only is
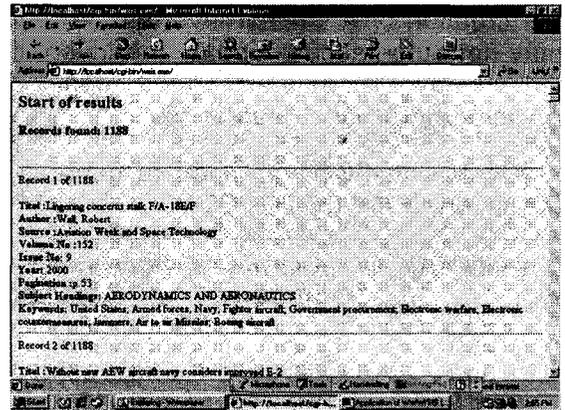
shown in Fig. 4. The script for citation is:

'<B><p>'mfn,'</B>
','Title:' v200,'<br>', 'Author:'v300+|;|,'<br>','Source:'v201,'<br>', 'Citation: 'v491,|(|v492|)|', 'v442' 'v490,', 'v493,'<br>','Keywords:' V620+|;|,'<HR>'

❑ Print format for Titles only

To display titles only, title field has been chosen. An advanced search PFT for titles is shown in Fig. 5. The script for this format is:

'<B><p>'mfn,'</B>
','Title:',V200,'
<p> <Font Face=arial Color=red>','</FONT>', '<HR>'

6.9 Interfacing DEFSTAD with WWWISIS Software

With BIREME ISIS it is possible to create a 100 per cent web-based CDS/ISIS system, complete with web-based data entry, editing, indexing, and retrieval features, or create and update databases on the fly from user input over the web. WWWISIS operates as a server through the WWW CGI. The actual operation is driven by the input database name, Boolean search, and an output format specification.

A key feature of WWWISIS is its capability to dynamically generate its execution parameters based on data sent to the CGI
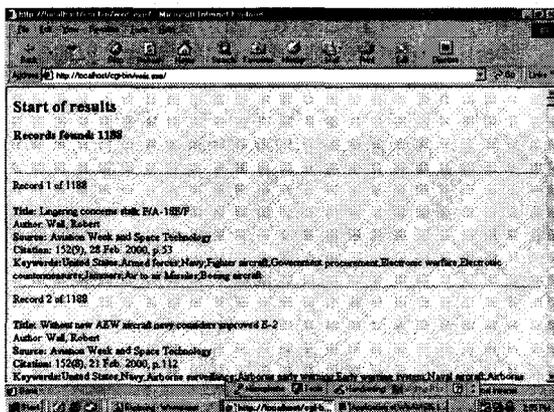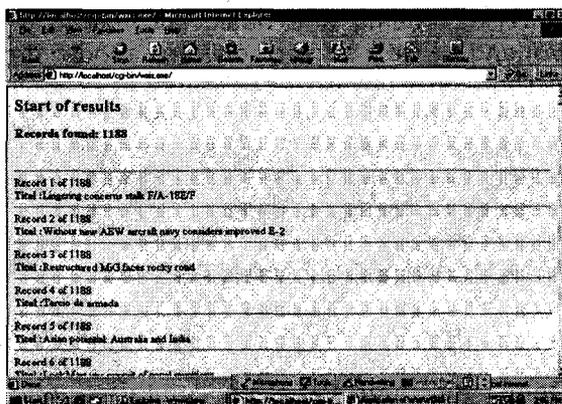
Figure 4. Result display: Citations only

Figure 5. Result display: Titles only

process from the WWW.

The database structure in WWWISIS is the same as in CDS/ISIS (DOS and Windows). Databases are developed using WINISIS 1.4 version and then enabled for searching using WWWISIS. Data between DOS/Windows and Unix can be exported from one version with ISO file and then imported into the other. The databases can be loaded or stored in any directory on the server, provided that cipar parameter in IsisScript tell wxis where to find the database, and the directory has suitable permissions set for the database-intended actions (eg. read or write).

All DEFSTAD files with the following drive/path (.cnt, .fdt, .fst, .ifp, .101, .102, .Lkl, .Lnl, .Ln2, .mst, .nOl, .n02, .xrf, .fmt,.pft) created by WINISIS are copied into WORK virtual directory in the IIS for search and retrieval. However, all the HTML files wxis.exe, and all .xis scripts are saved in CGI-BIN in IIS. The database thus created can be easily interfaced with WWWISIS to provide access to the database locally or through intranet, and can also be queried using WWWISIS in the command prompt.

7. CLDST: A CURRENT AWARENESS SERVICE USING WINISIS

Current awareness service has been described as delivering the right information to the right user at the right time in the right format covering the right sources at the right

cost and with the right amount of effort to keep the user up to date. Current literature in defence science and technology is a monthly current awareness service developed by DESIDOC using Windows version 1.4 of CDS/ISIS as back-end and WWWISIS version 5.4 as front-end for searching databases on Windows XP platform. It is a collection of bibliographic articles from selected core journals subscribed by DSL.

The service provides details of the articles pertaining to new scientific research, new technologies, products, reviews, and editorials published in leading scientific, technical, and military science journals. This current awareness service is meant to keep the DRDO scientists abreast with the latest developments in the field of military science and technology. The keyword index provides an access tool to the readers to know specific information on the topic of their interest. The bibliographic details of the articles and keyword index are generated by WINISIS through the PFT's. The current awareness service are available on DRONA, (DRDO rapid online network access).

8. CONCLUSION

It has been concluded that it is easy to access single or multiple CDS/ISIS database using WWWISIS. It is a very convenient tool for providing web access to CDS/ISIS databases. Only thing required is a sound knowledge of ISIS formatting language, HTML and CGI. Any person with a working knowledge

of HTML coding and basic knowledge of working of web server software should be in a position to publish the CDS/ISIS databases on the web. However, care must be taken to design proper FST, which contains the fields to be inverted and the indexing technique to be used for each field.

ACKNOWLEDGEMENTS

FURTHER READING

1. Prasad, A.R.D. Using multimedia database with WWWISIS on Win9x/NT. *In* Paper Presented at Workshop on Multimedia and Internet Technologies, 26-28 February 2001, Documentation Research and Training Centre, Bangalore.

2. http://ncsi.iisc.ernet.in/raja/netlit/iirml/franc2.ppt (visited on 22 May 2006)

3. http://www.isisnetz.de/download/dokumentationen/wwwisis-handbook.pdf (visited on 22 May 2006)

4. ftp://ftp.bireme.br/wwwisis/doc/wwwisisl2.doc (visited on 24 May 2006)

5. CDS/ISIS for windows: Reference manual: Version 1.3. UNESCO, November 1998.

6. WWWISIS: A world wide web server for CDS/ISIS databases. *Asia-Pacific POPIN Bull.* 1999, **11**(3), 19-22.

7. http://itt.nittat.tripod.com/itt9904/itittut.htm (visited on 24 May 2006).

8. www.geocities.com/devadason.geo/index-websis.htm (visited on 24 May 2006).

9. http://www.w3.org/TR/REC-htm40/interact/scripts.html (visited on 24 May 2006).

A BRIEF SYNTAX OF THE HTML FORM

```
<HTML>
<HEAD>
<TITLE>Search Database</TITLE>
<style>
<!--
.offsmall {
FONT: 8pt Arial, Helvetica, sans-serif, verdana, ; TEXT-DECORATION: none;
}
-->
</style></HEAD>
<BODY TEXT="#000000" BGCOLOR="AAAACC" LINK="#0000EE" VLINK="#551A8B"
ALINK="#FF0000" topmargin="0" leftmargin="1">
<center dir="ltr">
<form method=post action="http://localhost/cgi-bin/wxis.exe/">
<INPUT type=hidden name=IsisScript value="desi.xis">
<center>
  <p> </p><p> </p>
  <p>Search: <input type="text" size=55 maxlength=253 name=expression></p>
</center>
<blockquote>
    <blockquote>
        <blockquote>
        <center><p>       </p>
     <p>    <input type=submit name="tag6021" value="Search"
style="font-weight: bold">      ;
   <input type=reset name="tag6022" value="Reset" style="font-weight: bold"></p>
   </center></blockquote></blockquote></blockquote>
   </form>
```

ISISSCRIPTS SCRIPT USED FOR SIMPLE SEARCH

```
<parm name=cipar>
<pft>
'TEST.*=d:\inetpub\wwwroot\WORK\TEST.*'/
'2005.*=d:\inetpub\wwwroot\WORK\2005.*'/
'full.*=d:\inetpub\wwwroot\WORK\full.*'/
'Title.*=d:\inetpub\wwwroot\WORK\Title.*'/
'citn.*=d:\inetpub\wwwroot\WORK\citn.*'/
</pft>
</parm>
<IsisScript name=desi>
<display><pft>'Content-type: text/html'##</pft></display>
<display><BODY BGCOLOR="FFFFCC"></display>
    <field action=cgi tag=3000>db</field>
    <field action=cgi tag=3010>expression</field>
    <field action=cgi tag=3020>format</field>
<display><H2><center>Start of results</H2></center></display>
<do task=search>
    <parm name = db><pft>v3000</pft></parm>
    <parm name=expression><pft>v3010</pft></parm>
    <parm name=format><pft>v3020</pft></parm>
    <field action=define tag=1001>Isis_Current</field>
    <field action=define tag=1002>Isis_Total</field>
<loop>
<field action=import tag=3020>3020</field>
        <display><pft>if val(v1001) = 1  then
            '<b><H3>Records found: ',v1002,'</H3><br><hr></b>'
          fi </pft></display>
<display><pft>'Record ',v1001,' of ', v1002 </pft></display>
<display><pft>if v3020='Full' then @full.pft
            else if v3020='Short' then @citn.pft
            else if v3020='Title' then @Title.pft
            fi fi fi
```

```
</pft>
</display>
</loop>
        <display><pft>if val(v1002) = 0 then
        'No records found' fi</pft></display>
</do>
<display><H2><center>End of results</H2></center></display>
</IsisScript>
```

HTML CODE USED FOR SIMPLE SEARCH IN DEFSTAD

```
<HTML>
<HEAD>
<TITLE>Search Database</TITLE>
<style>
<!--

.offsmall {
FONT: 8pt Arial, Helvetica, sans-serif, verdana, ; TEXT-DECORATION: none;
}
-->
</style>
</HEAD>
<BODY TEXT="#000000" BGCOLOR="AAAACC" LINK="#0000EE" VLINK="#551A8B"
ALINK="#DFD9CB">
<H2 ALIGN=center>Defence scientific information and documentation center</H2>
<center>
<h4>
 Web access to Defence science and technology articles database</h4>
<p>
<i><b>Simple Search Form</b></i></p>
<FONT COLOR="#00356A">
<hr>
</font>
<form method=post action="http://localhost/cgi-bin/wxis.exe/">
<INPUT type=hidden name=IsisScript value="desi.xis">
Select Database(s)</b>
<select name="db" size="1">
<option value="d:\inetpub\wwwroot\WORK\TEST" selected>TEST</option>
<option value="d:\inetpub\wwwroot\WORK\2005">2005</option>
</select>     
  <p><input type="text" size=55 maxlength=253 name=expression></p>
</center>
```

```html
<blockquote>
    <blockquote>
        <blockquote>
        <center><p>
        <select name=format
                size=1>
                    <option value="Full" selected>Full Record</option>
                    <option value="Short">Citation</option>
                <option value="Title">Titles</option>
            </select>
                    <input type=submit name="tag6021" value="Search" style="font-weight:
bold"> &nbsp    
                    <input type=reset name="tag6022" value="Reset" style="font-weight:
bold"></p>
    <table border="0" width="530" cellspacing="0" cellpadding="0" align="center" bgcolor="#A0956D">
    <tr>
    <td width="52" align="center">
        <font face="verdana" size="2"><small>  <a href="help.htm">HELP</a></small></
font></td>
    <td width="478" align="center">   <font face="verdana" size="2">
    <font face="Arial" size="1">
     </font>   </font>
    <b><font> ;           
</font></b><font                                                            face="verdana"
size="2"><small>           
Search by :</small>
    </font>
        <font face="verdana" style="font-size: smaller">
        <a style="text-decoration: none" href="desi1.htm"><font color="#0000FF">ADVANCED
FORM</font></a></font></tr>
    </table>
                </center>
            </blockquote>
        </blockquote>
    </blockquote>
    </form>
    <FONT COLOR="#00356A">
    <hr>
    </font>
```

Contributors

Shri Sridhara B. obtained Masters in Library and Information Science from Kuvempu University, Karnataka. His areas of interest include: Digital libraries, database, and digitisation. Currently he is working as a Scientist at DESIDOC.

Smt Veena Makhija obtained Masters in Physics from the University of Delhi, Delhi and Associateship in Information Science from Indian National Scientific & Documentation Centre, New Delhi. She worked as a Documentation Officer at National Institute of Entrepreneur and Small Business Development, Ministry of Industry during 1989-90, and as a Scientific Officer at National Informatics Centre, Ministry of Planning in 1991. Her areas of interest include: IT database, digitisation, and digital libraries. At present, she is working as a senior Scientist at DESIDOC.

Shri Ajay Kumar Pandey obtained Masters in Library Science from Indira Gandhi National Open University, Associateship in Information Science from National Institute of Science Communication and Information Resources, New Delhi. His areas of interest are: IT database, digitisation, and digital libraries. Currently he is working as a Senior Technical Assistant at DESIDOC.