Creating a Greenstone Digital Library with Winisis Database

K RAJASEKHARAN
K M NAFALA

Library & Information Division
Kerala Institute of Local Administration
Mulagunnathu Kavu – 680 581, Thrissur, India
Email: librariankila@hotmail.com
Phone: 91-487-2204097

Preface

The purpose of this manual is to serve as a self-instructional guide to those who want to convert Winisis archive into greenstone collections with Greenstone digital library software 2.72 independently.

The manual can be used as a training hand-out for exposure training that long for one or two days. The basic approach of the manual is to include the essential and easy-to-learn things and to exclude everything else that cannot be learned by new users.

In exposure training, please desist from showering a lot of information of complex nature to the new participants who cannot digest all of them. That many dispel some of them from greenstone.

Practice reminds that learning happens in the form of layers one over the other. When the first layer is not stable, you cannot build up the subsequent layers. This manual deals with the first layer of knowledge on implementation of greenstone. Advance features are not dealt with in this manual.

We express our gratitude to Dr. John Rose for the enthusiasm he has shown in clarifying our doubts and helping us learn newer changes, without which the document would have more infirmities.

Mulakunnathukavu   *K. Rajasekharan*
20-02 2007   *K.M Nafala*

Creating a Greenstone Digital Library with Winisis Database

~~~~~~

K Rajasekharan* **& K M Nafala****

1. Introduction

A digital library is an organized collection of digital documents, in any format such as text, image, audio or video, that would permit easy access by information users and proper maintenance by librarians. Digital libraries reinforce the concept of a library as a repository of knowledge with the notion of it as a provider of information services to users. In this age of exponential growth of information materials in print as well as in digital form, both physical and digital libraries are essential and can work in unison, supplementing each other.

The *Greenstone Digital Library Software* (*GSDL*) offers exciting ways to build and distribute digital document collections. It helps us to publish digital collections on the Internet or on CD-ROM. Within a few minutes time, one can build full-text *search indexes* and *browsing classifiers* for any collection of digital documents. Once initiated, the collection building process will take place mechanically, running into several hours or days for a very large collection.

1.1. Objectives of this Guide

The purpose of this guide is to help the users to convert a full-text digital document archive created with Winisis into a greenstone digital library.

Winisis database, consisting of bibliographical records alone, can be converted into greenstone collection[1] and you can put it on the web or on CD-ROM.

---

*   Librarian, Kerala Institute of Local Administration, Mulagunnathu Kavu – 680 581
** Library Computer Operator, Kerala Institute of Local Administration, Mulagunnathu Kavu –680 581
[1] This Winisis into greenstone conversion is described as "as is" conversion and the modalities of it is described in the document titled "Creating Digital Libraries based on CDS/ISIS Database" referred to at the end of this guide under "References"

1.2 Software Installation

Software Download

The *GSDL 2.72* software can be downloaded from http://www.greenstone.org/ or http://greenstonesupport.iimk.ac.in/our_mission.htm

Java2 Runtime Environment can be downloaded from http://java.sun.com/

ImageMagick can be had from http://www.imagemagick.org/

Ghostscript can be obtained from http://www.cs.wisc.edu/~ghost/

Winisis 1.5 can be downloaded from www.unesco.org/isis/files/Winisis15_3.exe or www.unesco.org/isis/files/winisislicense .html

Installation of Software

Install **Java2 Runtime Environment** in your computer as a pre-requisite for installing GSDL software. Then install the **GSDL 2.72 (windows version)** software in your computer. Choose the recommended **Local Library** mode for installation.

You may install **ImageMagick** and **Ghostscript** softwares, required to build image collections and to do advanced conversion of PDF and Postscript documents, respectively.

1.3 Creating a Winisis Database of Fulltext Documents

Collect a set of digital documents and place them in a folder in your computer. Create a database with the metadata of the documents such as Author, Title & Keywords in Winisis[2]. In the database, one field[3] (40 Fulltext) should contain the exact location[4] (D:\Sample\001.pdf) of the fulltext document. This Winisis database and the associated fulltext documents can be built into a greenstone collection. The conversion to greenstone offers some advantages[5].

---

[2] Winisis is the Windows version of the CDS/ISIS software of the Unesco
[3] Filed is *Fulltext*, Tag is *40* here
[4] Location of the first document provided in the field 40 Fulltext is D:\Sample\001.pdf
[5] The advantages are that you can have fulltext searching of associated documents and it can be easily converted into a CD-ROM or be put in the web.
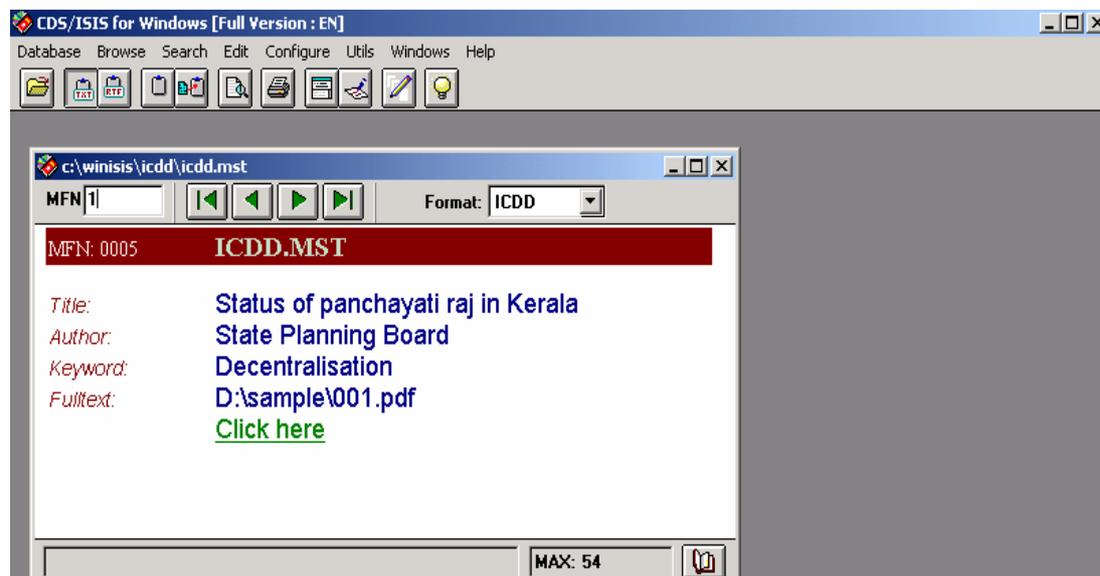
The above database and documents can be converted into a greenstone collection. But if you insert the format line[6] **link((Click here), 'OPENFILE ',v40)** in the print format of the Winisis database, the database will be linked up to the appropriate fulltext document and you can use it as an independently working Winisis archive.

The adding of the format line in the print format will create a *hyperlink*, from the words **Click here** appearing on the database display, to the appropriate fulltext document. A click on the **hyperlink** will result in opening up the document automatically. You can search for document with winisis, retrieve it and open the document by a mouse click on the link.

The sample collection used here contains 54 papers of the *International Conference on Democratic Decentralisation* (*ICDD*).

The Winisis archive would appear as follows:-

Figure 1 Winisis Archive of Digital Documents

---

[6] The meaning of the command is that, when you click on the link **'Click here'**, the click will automatically result in opening the digital document, denoted in the field V40, in a new window.

2. Creating the Collection with Greenstone

Open the Greenstone Librarian Interface (GLI)[7] **Start ⇒ Programs ⇒ Greenstone Digital Library Software v2.72 ⇒ Greenstone Librarian Interface**.

To create a new collection, choose **New** from **File** menu
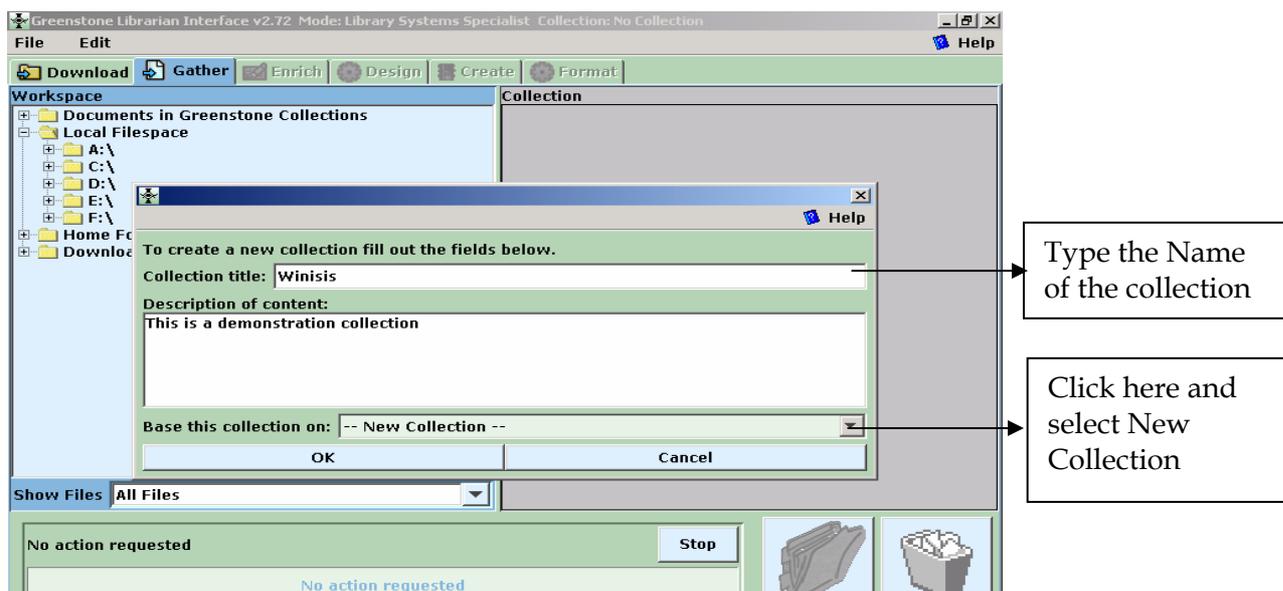
Figure 2 Starting a New Collection

Give a title for the collection (say **Winisis)** and give a short description about the collection as shows in the following figure.

---

[7] GLI can be used to perform the following basic activities, while building up a collection: -

1. **Gather** and associated documents for building up the collection
2. **Explode** the database by adding metadata
3. **Design** the collection, its appearance and the access facilities
4. **Format** the appearance of the digital library
5. **Create** the collection
6. **Convert** the digital collection into a CD-ROM library

Figure 3 Collection Name & Description

Choose **New Collection**[8] in the **Base this collection on** dialogue box and click **OK.**

2.1 Gather the Database

Now the **Gather** panel will become active. Locate **.mst, .fdt** & **.xrf** files from your Winisis database (normally at *C:\Winisis\data*) and drag them into the **Collection** pane as in the following screen:-
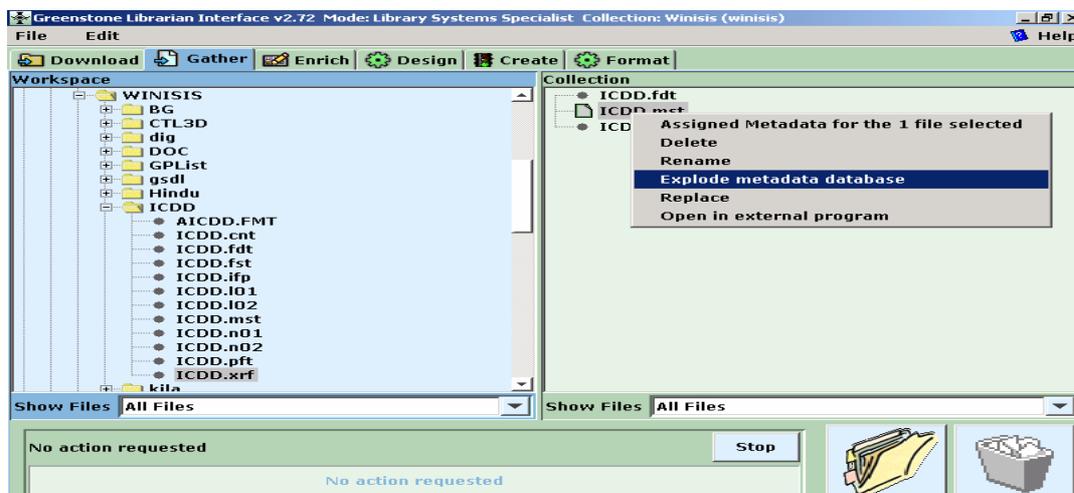
Figure 4 Gathering the Files

---

[8] This indicates that your collection to be built now will have entirely new structure. If you want your present collection to follow the pattern of an already existing collection, select the name of that collection from the drop-down menu.

2.2. Exploding the Metadata Database

Each Winisis record contains metadata elements of documents in the collection. If you explode the database, the Winisis database will be converted into individual documents with associated metadata.
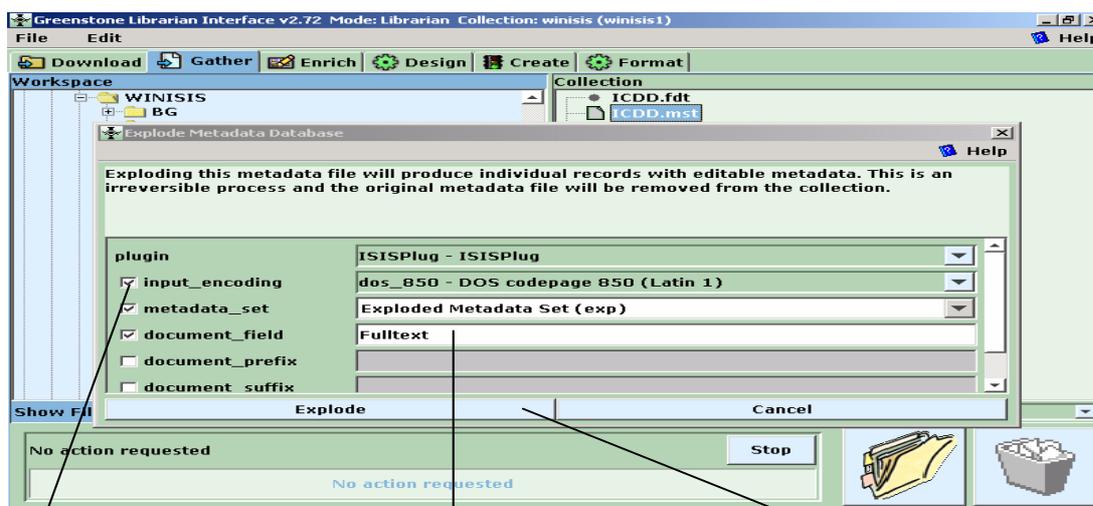
Right click the **.mst** file and choose **Explode metadata database**.

Figure 5 Explode Metadata Database

Then the following window will appear. Provide explode parameters as follows:-

Figure 6 Explode Parameters

Put tick mark to select Input_encoding and select DOS_850 as above	Type the name of the Winisis field containing fullpath of the document	Finally click here to Explode

In the above screen select:-

Input_encoding - dos_850-DOS codepage 850 (Latin 1)

document_field - Fulltext (it provide the *field name* in Winisis that holds the full path
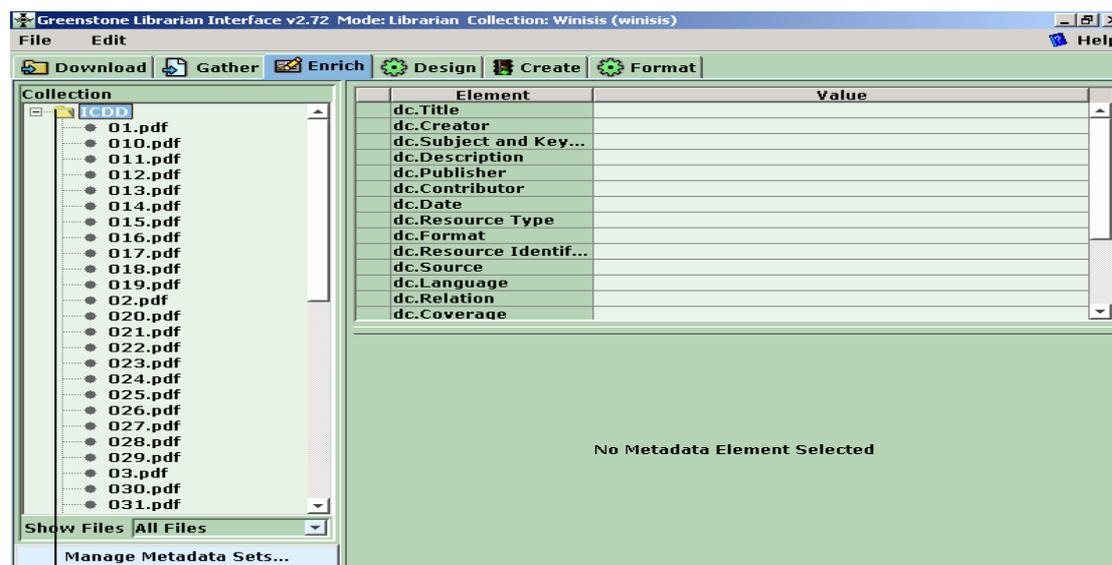
document_prefix –leave it as blank[9]

document_suffix - leave it as blank

Then click on **Explode** and click on **Ok** when explode is completed.

View the files in the Enrich Panel

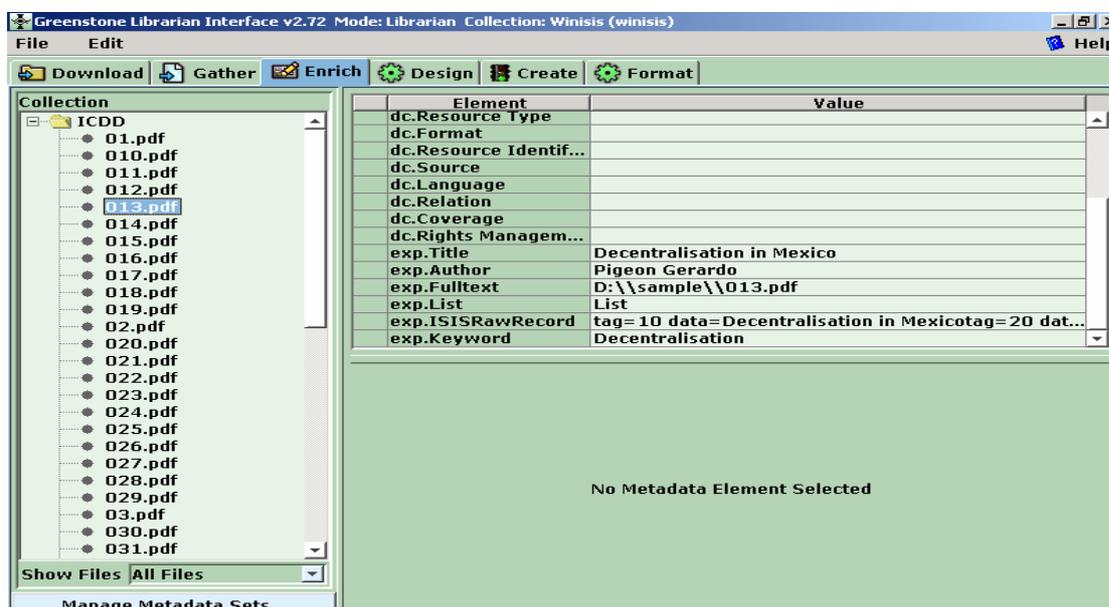You can view the fulltext documents in the **Enrich** panel as follows:-

Figure 7 View the Documents

Double click here to view the documents

---

[9] If the document field contains the document name (eg.ram) alone, provide the part of the path to be prefixed (eg. D:\sample\) in document_prefix and the file extension (eg. .pdf) in the document suffix. If you don't provide full path of the document in the Winisis database field, it is better to provide document name with extension in document_suffix (eg.ram.pdf) and the folder name (eg.D:\sample\) in the doument_prefix. Whatever be, the full path of the document location should be concatenated from document_field, document_prefix and document_suffix (eg. D:\sample\ram.pdf)

If the documents in the **Collection** pane of **Enrich** panel appears as dummy files with **nul** extension (.nul), the files are not gathered and something has gone wrong in collection building. So ensure that the **Collection** pane shows the exact documents file name with extension (eg. 01.pdf).

Enrich panel is meant for adding /viewing metadata. You can view the metadata[10] of each document by clicking on it in the left pane.

Figure 8 Metadata set of each document

2.3. Designing the Collection

Then design the collection by choosing the following features given under the **Design** menu. Collection design consists of many facets enlisted in the left pane.

2.3.1 Create Search Indexes

Choose the **Search Indexes,** shown on the left pane, for creating *Search Indexes*. Search Indexes determine whether to confine the collection search to paragraph, chapter or the entire text of the document. If you need fulltext search, keep the *Search Index* for *Text* and retain it as default index.

---

[10] The metadata elements with prefix **exp.** (exp.Title, exp.Author etc.) are only when you explode the database. The dublin core metadata such as dc.Title, dc.Creator etc. are of no consequence.

Remove the Default Indexes

Remove the default indexes for *ex.Title* and *ex.Source* by selecting the index description under *Assigned Indexes* and then by clicking on the **Remove Index** button. Do not remove the *search index* for **text [Default Index]**[11].
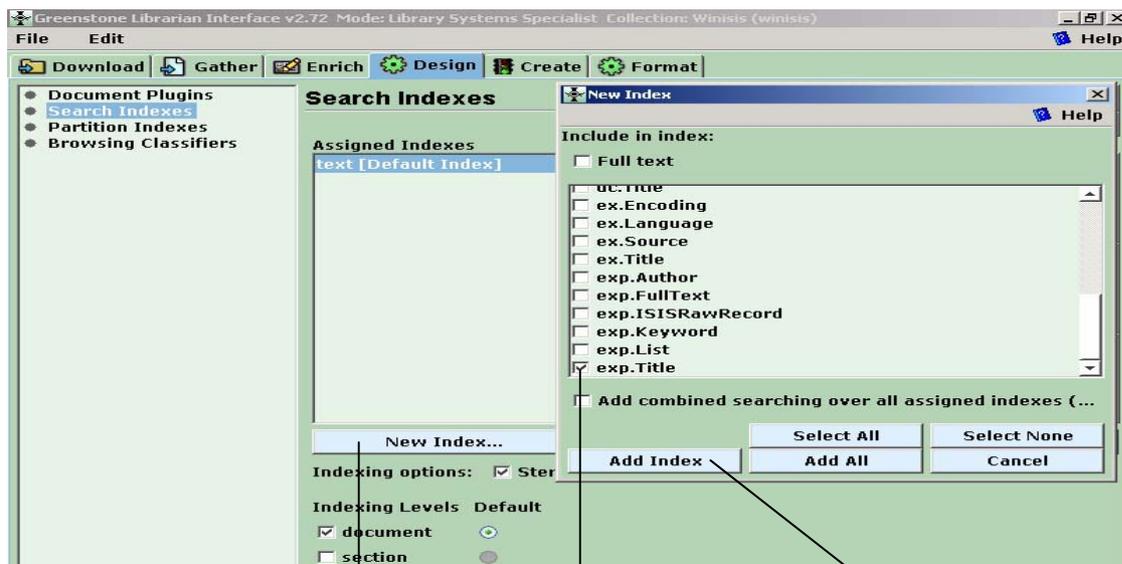
Figure 9 Removing Index

Adding New Indexes

Click on the **New Index** button. Select the **exp.Title**, **exp.Author** and **exp.Keyword**, by tick marking on the check box one by one, and add them one by one by clicking on the **Add Index** button. That means select *exp.Title* first and add it, then select *exp.Author* and add it, then select *exp.Keyword* and so on.

---

[11] If you don't provide the *search index* for text, you cannot search the entire text of the document. Make it default, preferably.

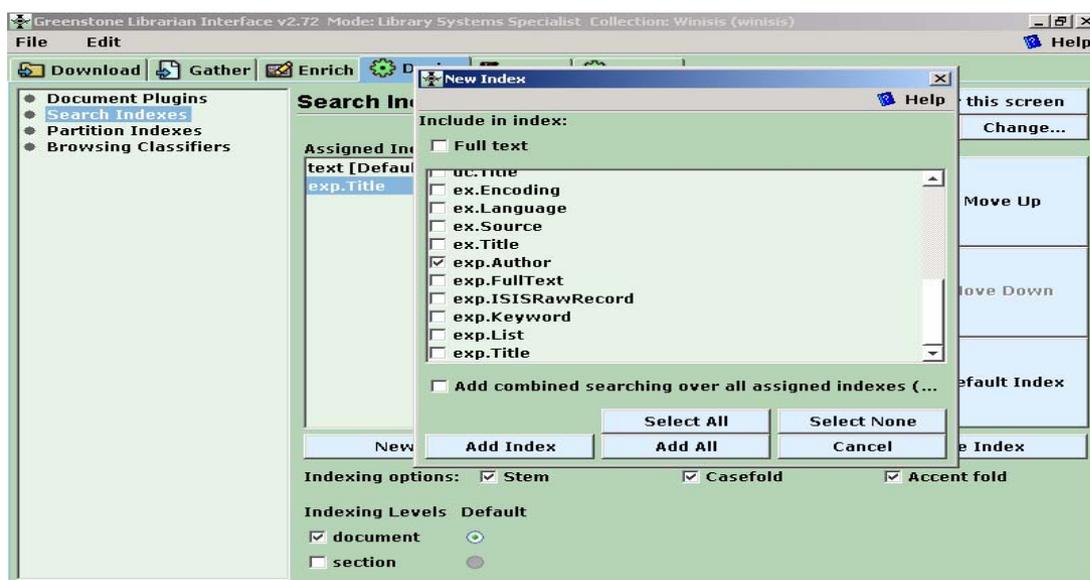**Figure 10 Adding Index for exp.Title**

Click here to add New Index

Put tick mark to select Index

Click here to Add Index

Select and add the Search Index for *Author* by choosing the metadata (*exp.Author*) as follows:-

Figure 11 Search Index for exp.Author

Select and add the Search Index for *Keyword* by choosing the metadata (*exp.Keyword*) as follows:-

Figure 12 Search Index for exp.Keyword

At the end, all the three indexes will be added one by one as in the following figure.

Figure 13 Indexes Selected for the Collection

You can move any index up or down to set its order of appearance by selecting the index and clicking on the buttons on the right side. Likewise you can set any *index* as *default index* by selecting the index and clicking on the **Set Default Index** button.

2.3.2. Browsing Classifiers

If you want to browse on a metadata element, you must set up a Browsing Classifier, independently of creating a search index on this metadata element. The metadata used for the database are *Author, Title & Keyword* and we need to create browsing classifiers for them. The default Browsing Classifiers are as follows:-
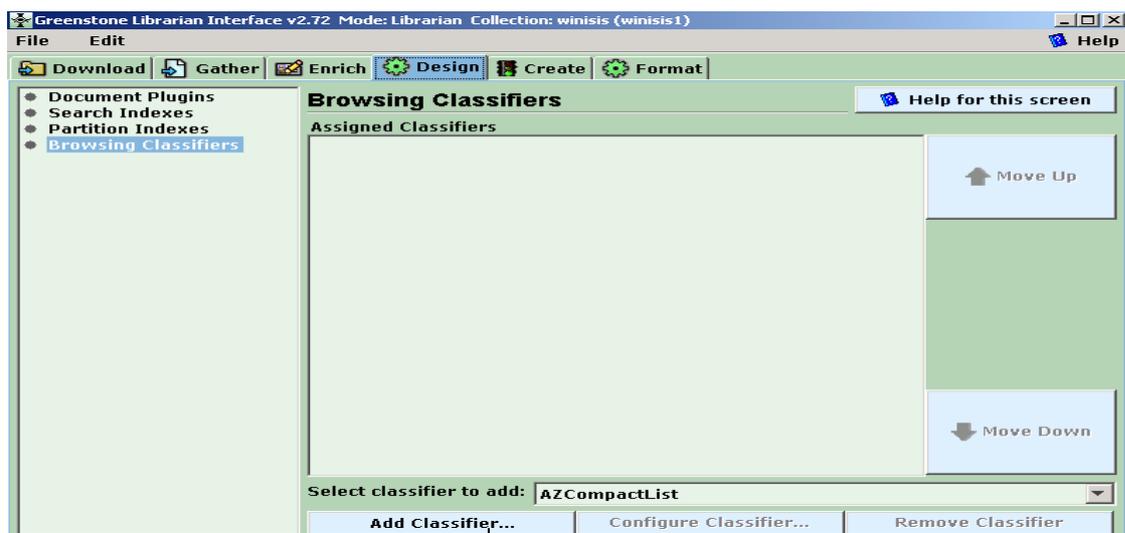
Figure 14 Default Browsing Classifiers

Click here to select Browsing Classifiers

Click here to Remove Classifier

Select **Browsing Classifier** by clicking on it and then click on the **Remove Classifier** button for removing the default classifiers[12] one by one. The resultant screen would be as follows:-

---

[12] Default classifiers *ex.Title* and *ex.Source* are of no use in this case

Figure 15 Adding Classifier

Click here to Add Classifier

Adding Browsing Classifiers

Now, choose **Select classifier to add** pull down list and select **A-Z List** or **A-Z Compact List** [13].

Then click on **Add Classifier** and add the Browsing Classifiers[14] for **Title**, **Author** and **Keyword** by one by one.
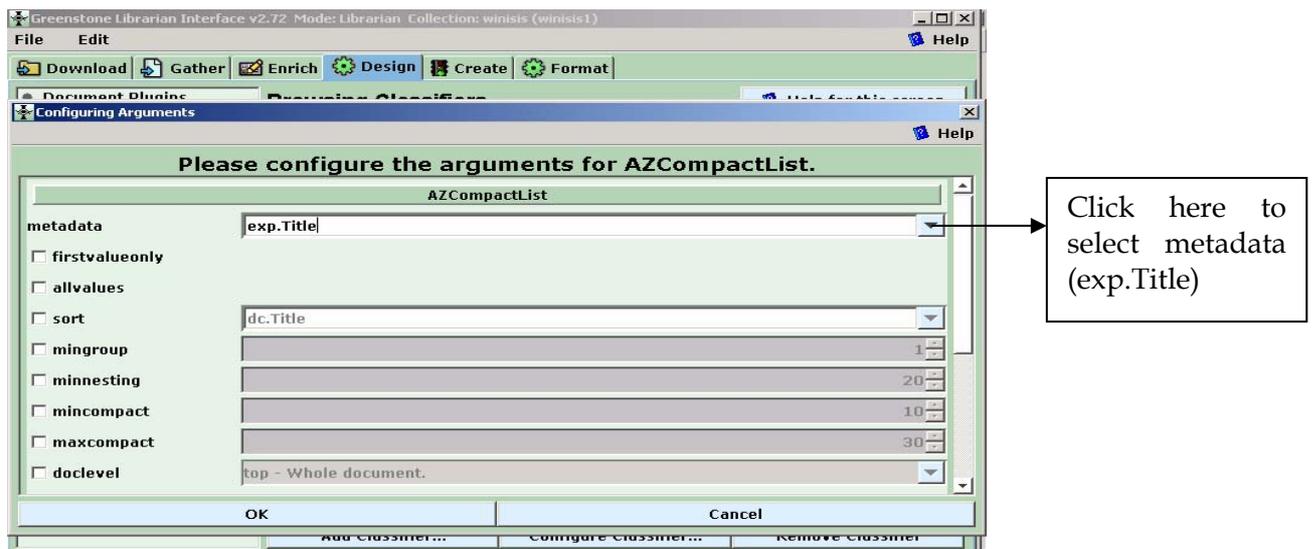
When you click **Add Classifier** button in the above screen, you will get the window for choosing the **Browsing Classifier**.

Select the browsing classifier for *Title* by choosing the metadata option (exp.Title) as follows:-

---

[13] Use of AZCompactList Classifier brings bookshelf icons under the browsing classifier. It groups together the documents that appear multiple times with same metadata and does not differ with AZList in any other manner.
[14] If you want to give a button name other than the button names like **Titles, Creators** etc, check the buttonname box in the "Configuring Arguments" window, type the desired button name and click **Ok**.
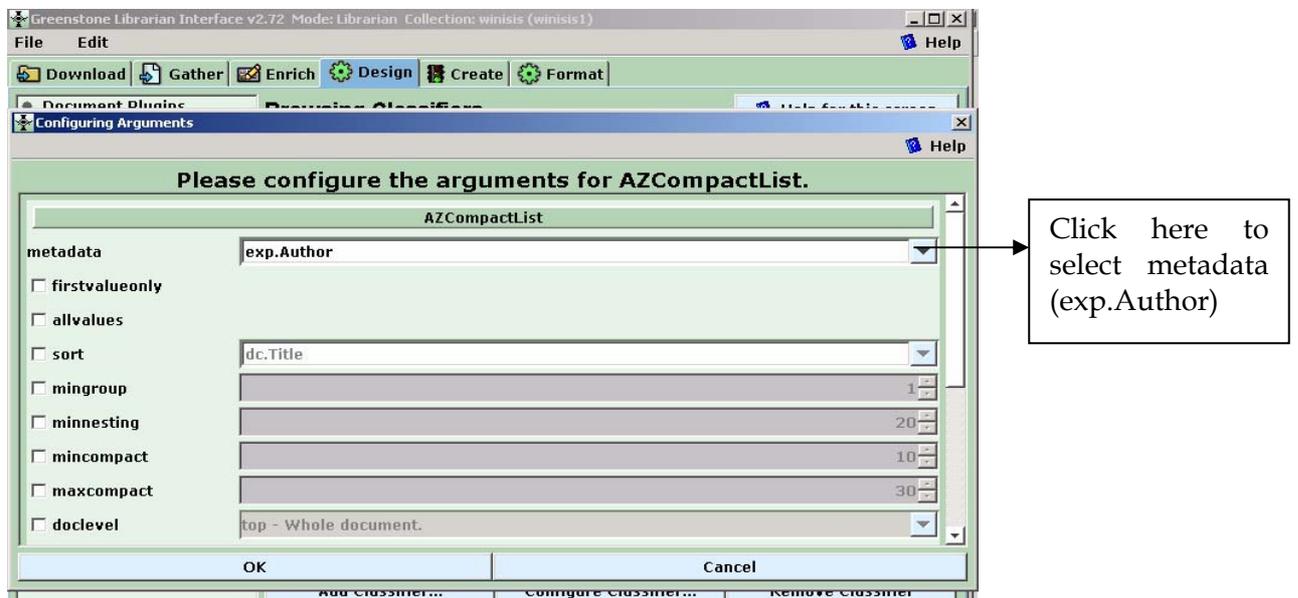
Figure 16 Adding Classifier CL1 for Title - exp.Title

Click **OK** button to add *Title* classifier.

Then select the browsing classifier for *Author* by choosing the metadata option (exp.Author) as follows :-
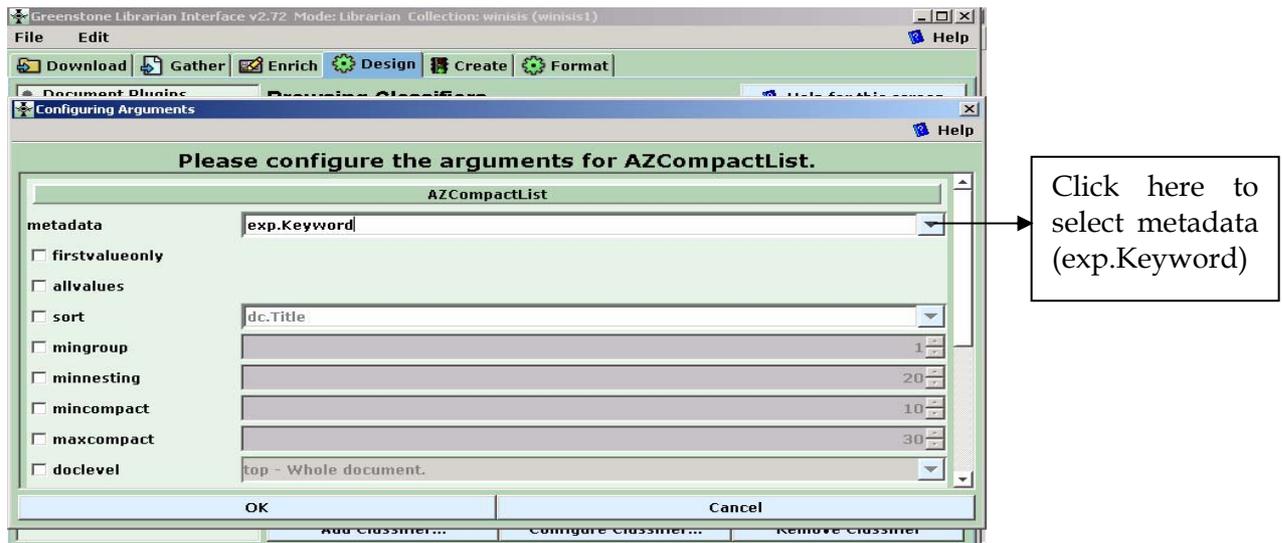
Figure 17 Adding Classifier CL2 for Author - exp.Author

Click **OK** button to add *Author* classifier.

Then select the browsing classifier for *Keyword* by choosing the metadata option (exp. Keyword) as follows :-
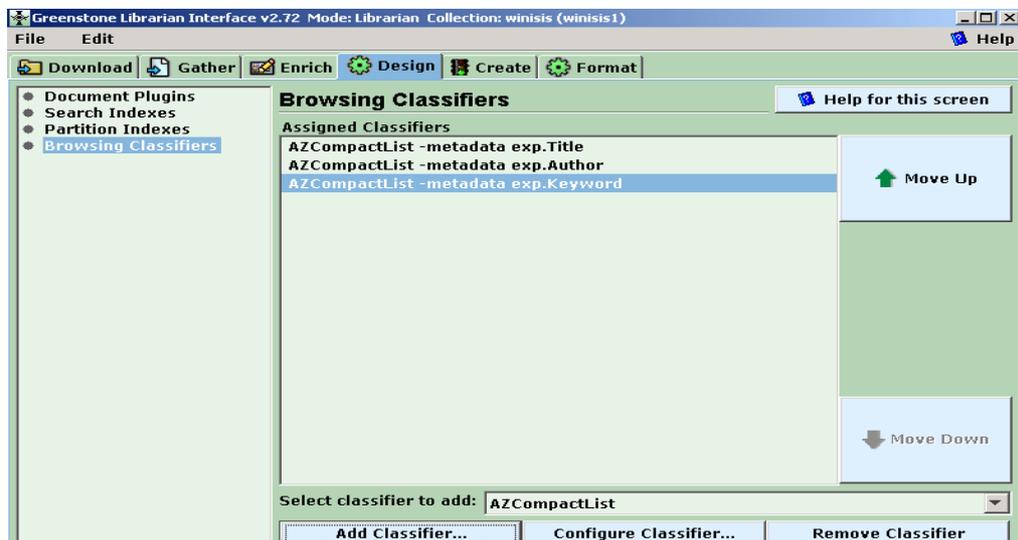
Figure 18 Adding Classifier CL3 for Keywords

Click **OK** button to add *Keyword* classifier.

The resultant screen, that shows the three added Browsing Classifiers, will appear as follows:-

Fig 19 Browsing Classifiers assigned for the Demo Collection

2.4. Format

Many aspects of the page-display of the resultant digital library are governed by the features provided under the **Format** menu.
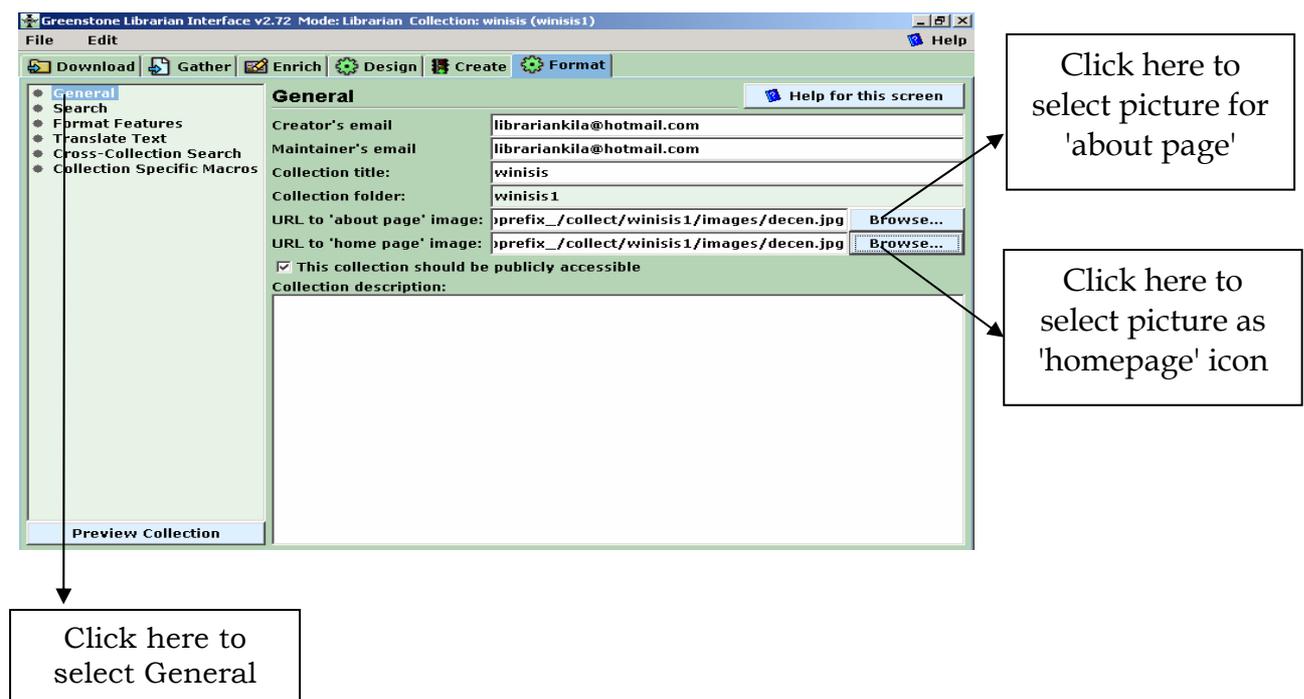
2.4.1 General

Choose the **Format** Tab and select **General** to provide the general information about the collection.

Put a picture as a collection icon

Choose a small picture of around *100*100* pixels that needs to appear as an icon of the collection in the homepage, by clicking on the **Browse** button on middle right and selecting the picture. On choosing the picture, the fullpath of the picture will appear in the address box lying against the **Browse** button. You can provide the same picture or a different picture as the image for the 'about page' also. About page is the first page about an individual digital library collection that may contain a short description about the collection.
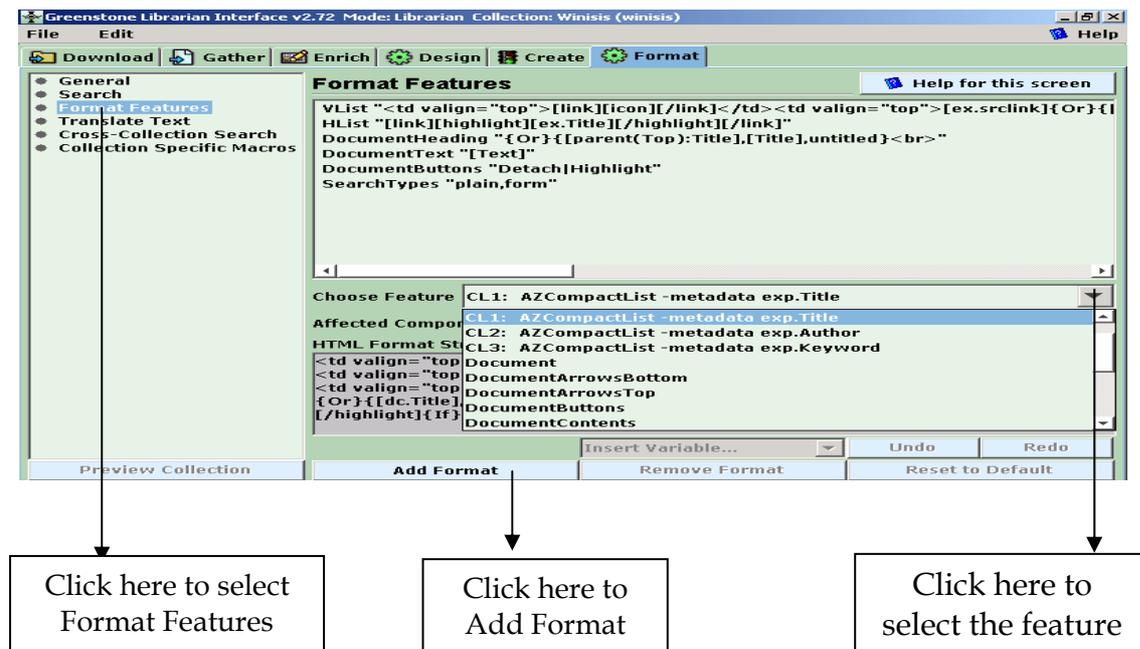
Figure 20 Designing the Collection – General Information

2.4.2 Format Features

Format Features[15] allow unlimited choices in modifying the display features of the collection. Click on the **Format Features** section of the **Format** panel on the left pane as follows: -

Fig 21 Modifying the Format Features

| Click here to select Format Features | Click here to Add Format | Click here to select the feature |

If you want to change any format feature, select the appropriate one (eg. CL1 AZCompactList- metadata exp.Title) from the **Choose Feature** pull down list and add the format string to the **Format Features** box by clicking on the **Add Format** button. You can select any feature appropriate for changing the display of the pages in the digital library. Fresh users may find it difficult to learn the format features in

---

[15] There are two types of Format Statements - General statement applicable in general to all and specific statement applicable to specific classifier or search list. Specific statements, if present, would override the general statements.
Let us examine some statements
VList - applies to all vertical lists in all classifiers
Search VList - applies to all search result lists
CL2HList - applies to all hierarchy lists in classifier 2
The value of any metadata can be interpolated by placing the metadata name in square brackets. eg. [Title] in the format statements.
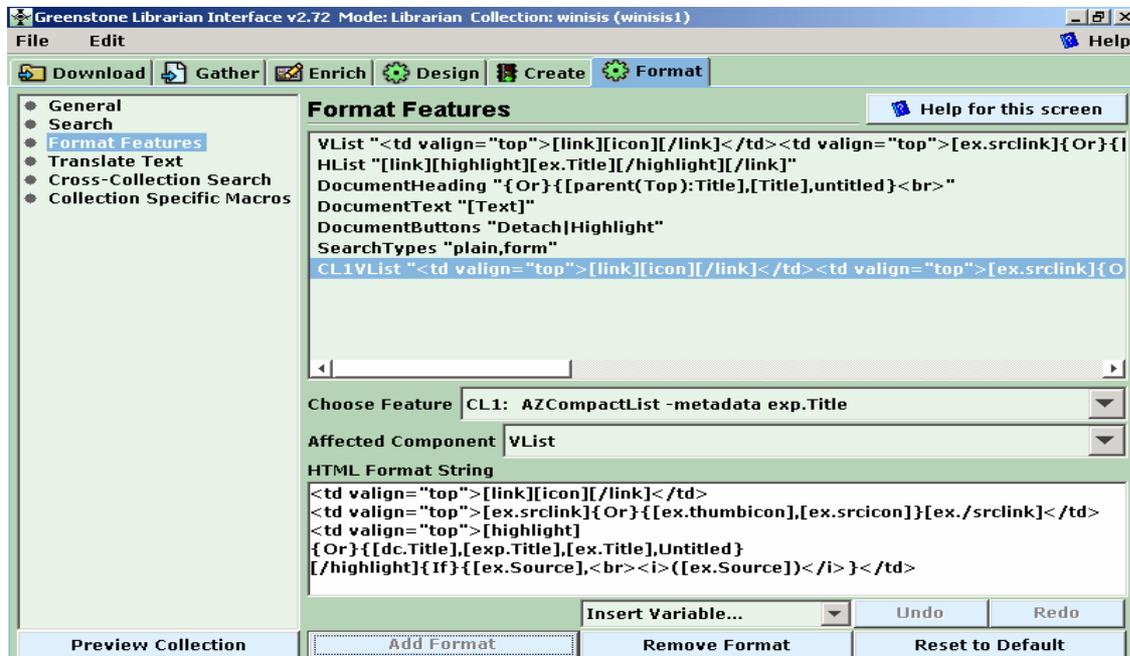[link] [icon] [/ link] provides a link from icon to the HTML version of the document.
[link] [Title] [/link] provides link from the Title of the document.
[srclink] [srcicon] [/srclink] provides a link from the word/pdf icon to the word/pdf version of the document.

the beginning stage. But if properly understood, it is so easy to manage. The CL1 Browsing Classifier[16] for Title can be added as shown in the following screen.

Figure 22 Adding CL1 Browsing Classifier for exp.Title

Similarly select other classifiers **Author** and **Keyword** from **Choose Feature** box and click **Add Format** button for customizing it as in the above screen. You can edit the HTML strings under the **Format Features** box by selecting it and editing them in the **HTML Format String** box below.

---

[16] CL1 denotes the first browsing classifier exp.Title and CL2 denotes the second one exp.Author and so on.

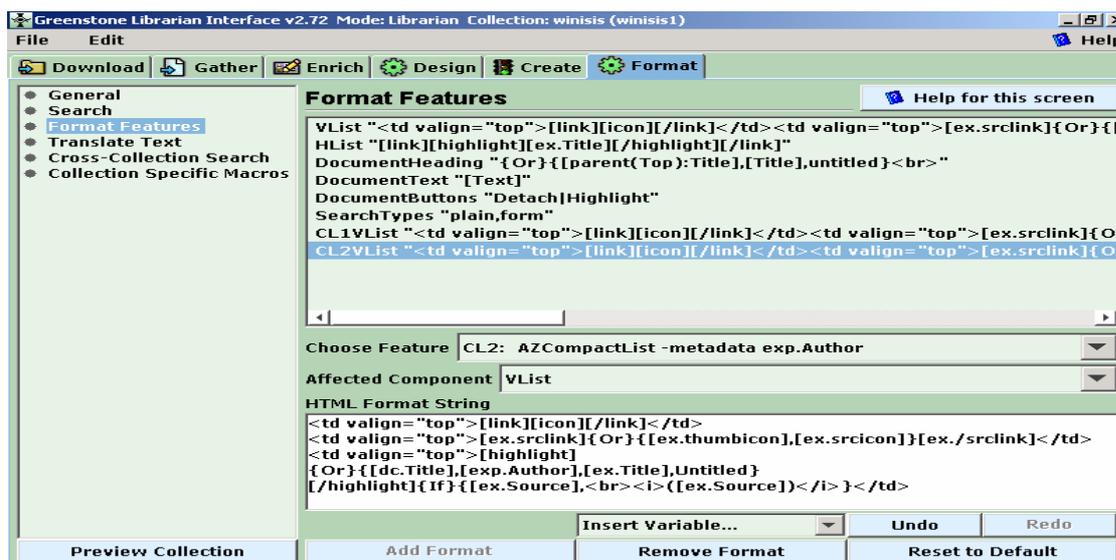**Figure 23 Adding CL2 Browsing Classifier for exp.Author**

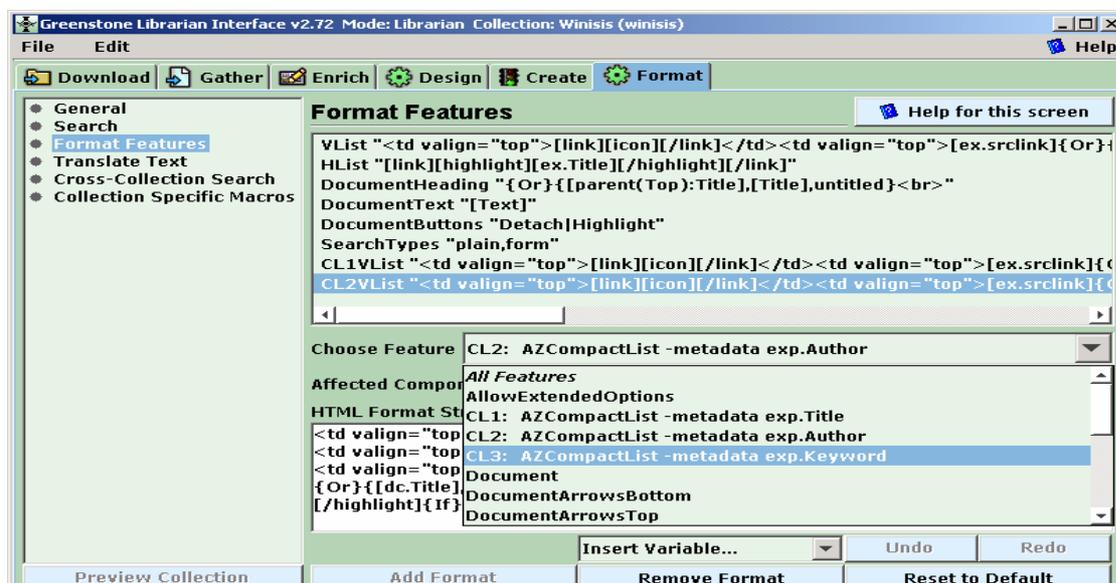**Figure 24 Modifying the Format for exp.Author**

Click here and replace *exp.Title* as *exp.Author*

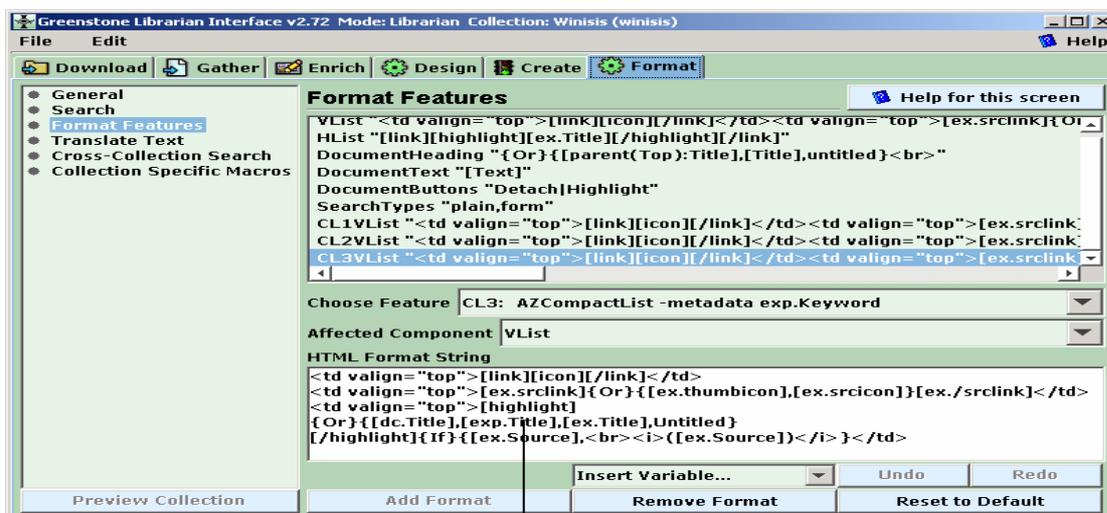**Figure 25 Modified Format for exp.Author**

To add format feature for **Author** select the **CL2: AZCompactList-metadata exp.Author** in the **Choose Feature** and then click on the **Add Format** button. Then change in the **HTML Format String** *exp.Title* as *exp.Author*

Figure 26 Adding CL3 Browsing Classifier for exp.Keyword

Choose the **CL3: AZCompactList-metadata exp.Keyword** and then add it by clicking on **Add Format** button.
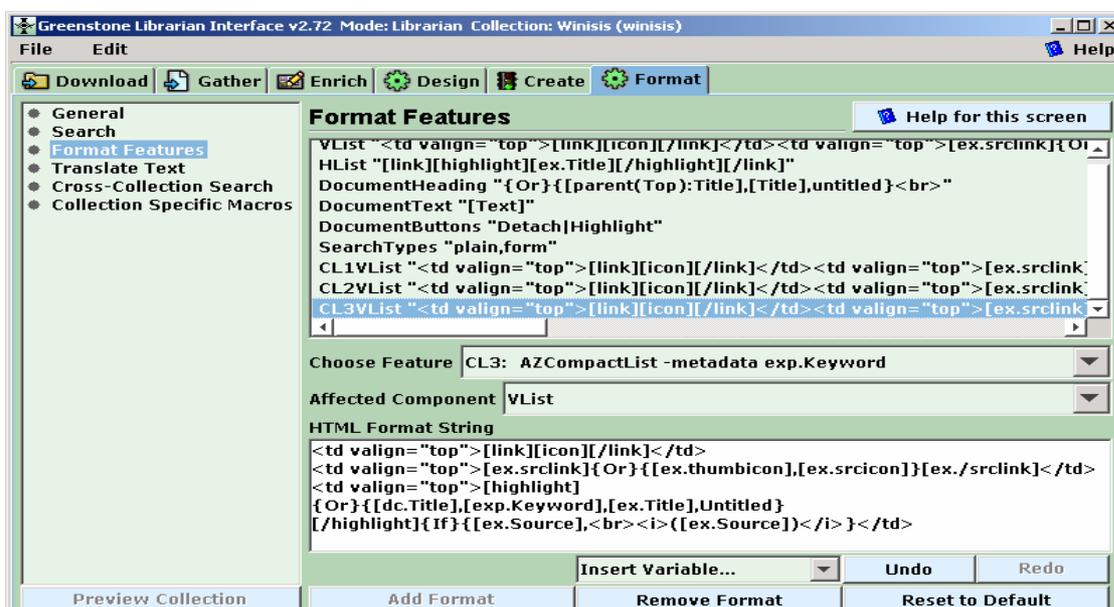
Figure 27 Modifying the Format for exp.Keyword

Click here and replace *exp.Title* as *exp.Keyword*

Then change the *exp.Title* as *exp.Keyword* as follows:-
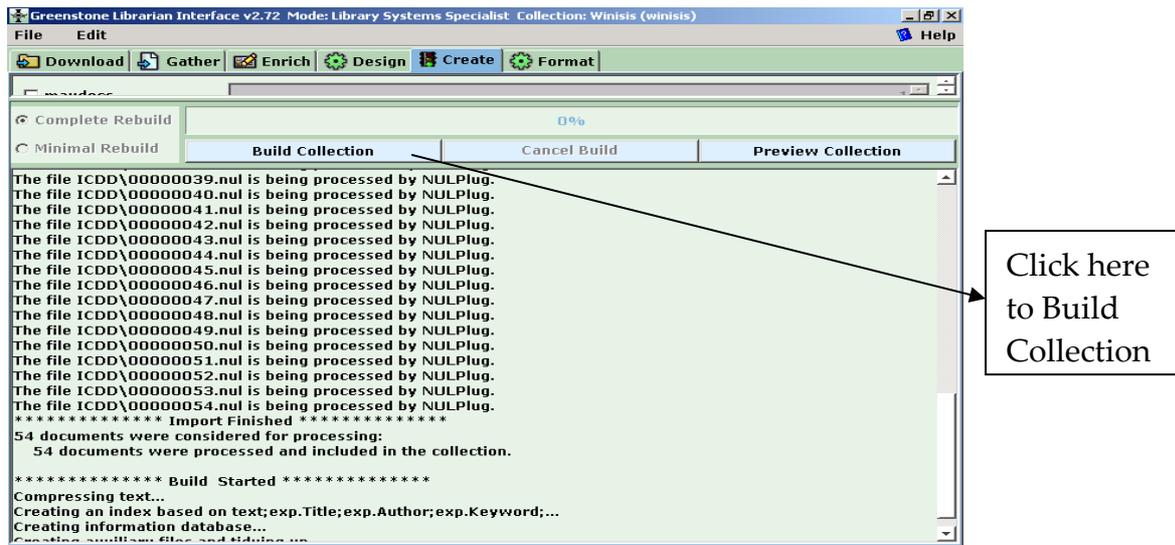
Figure 28 Modified Format for exp.Keyword

2.5 Build the Collection

Then go to the **Create** panel. Click on the **Build Collection** button and the progress-window will show the progress in building the collection.
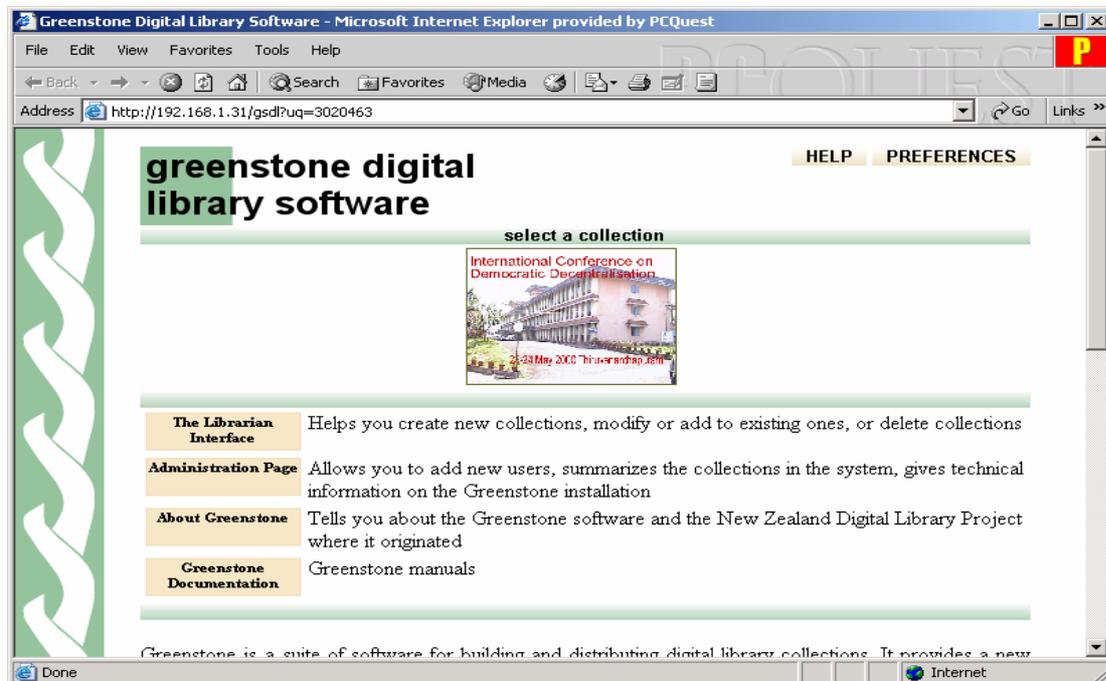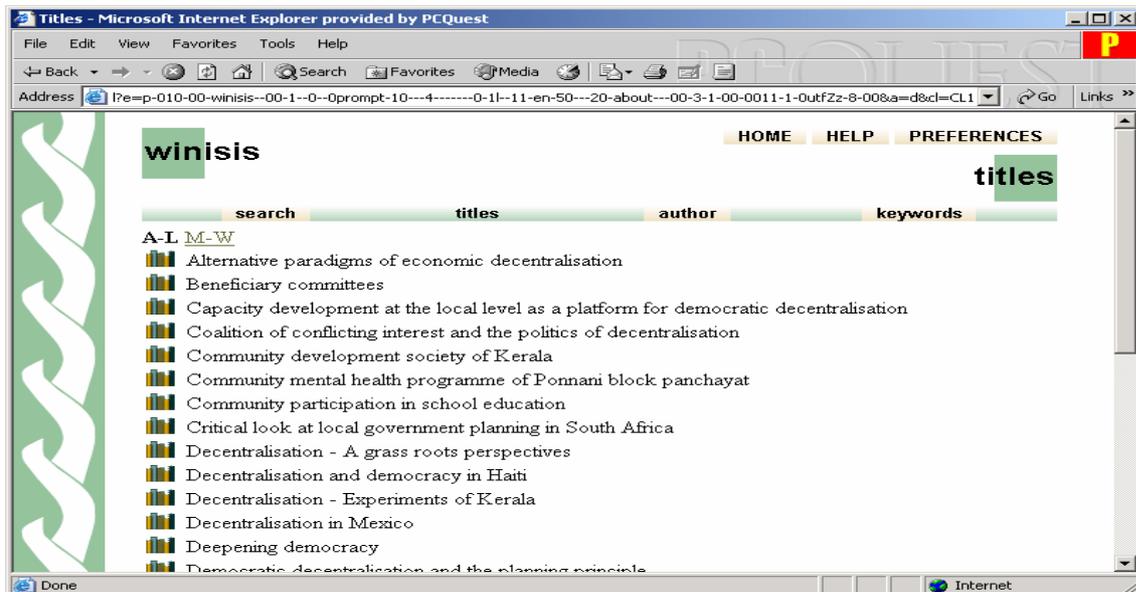
Fig 29 Building a Collection

2.5.1. Preview the Collection

At the end of the building process, Click on the **Preview Collection** button to view the collection built.
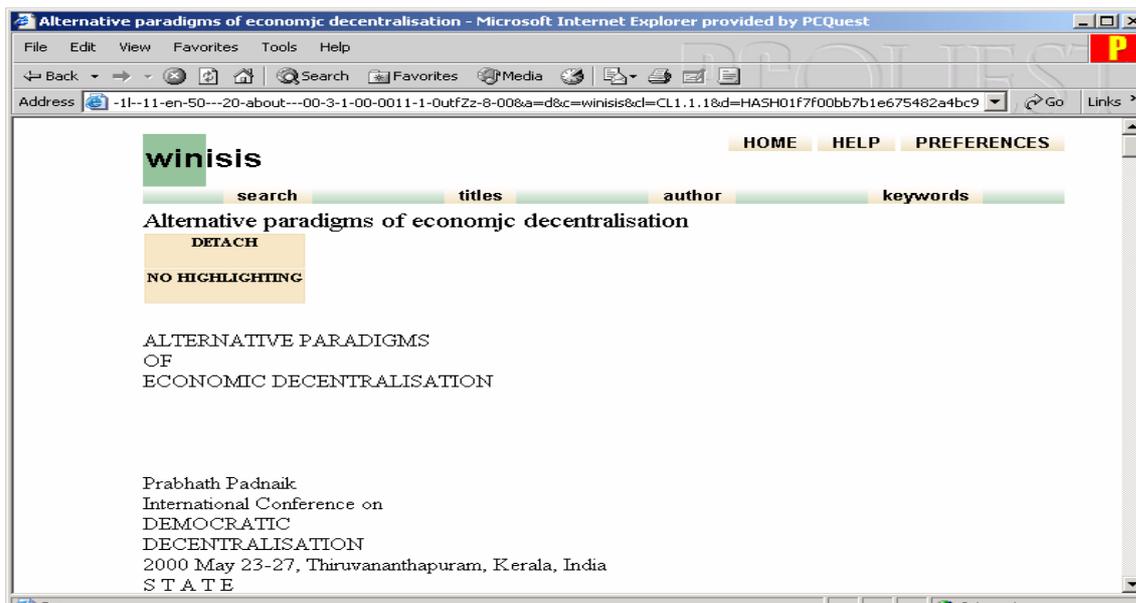
Figure 30 Home Page of the Demo Collection

Fig 31 Browse the Collection by Title Classifier
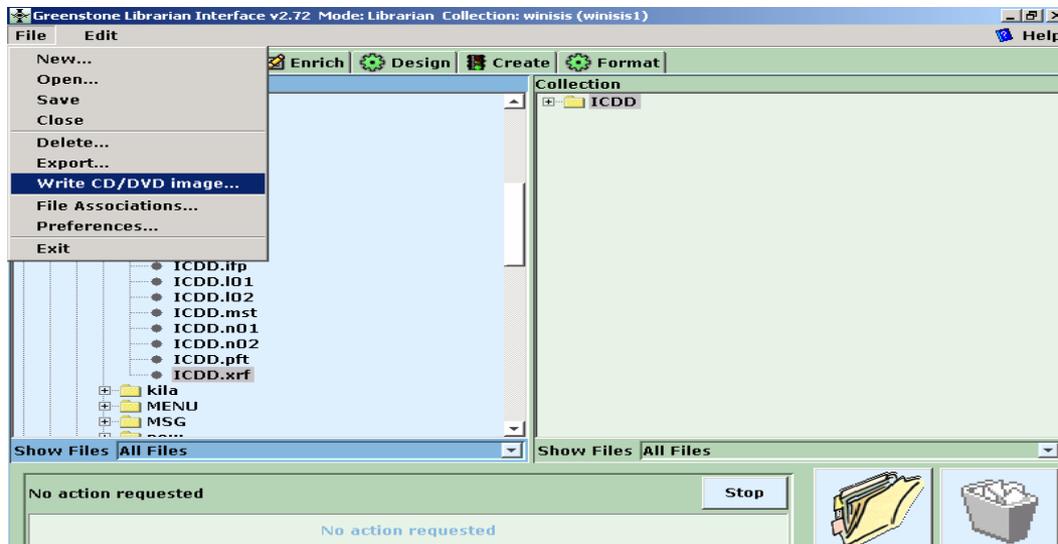

Fig 32 Open the Fulltext Document

2.6. Converting the Collection in to a CD-ROM Library

You can write the collection/collections to a CD-ROM, if you want to convert the Greenstone collection into an installable CD-ROM for distribution among wider audience.
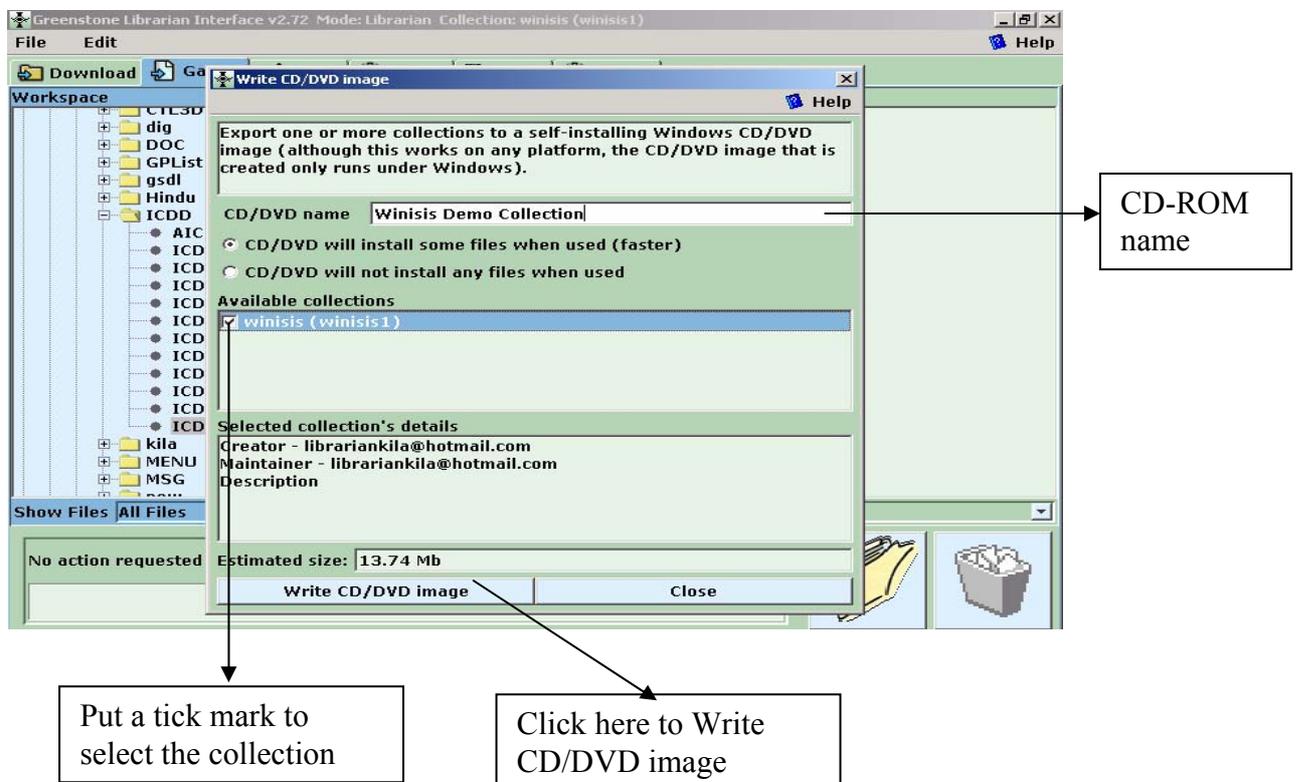
Click on **File ⇒ Write CD/DVD image......**
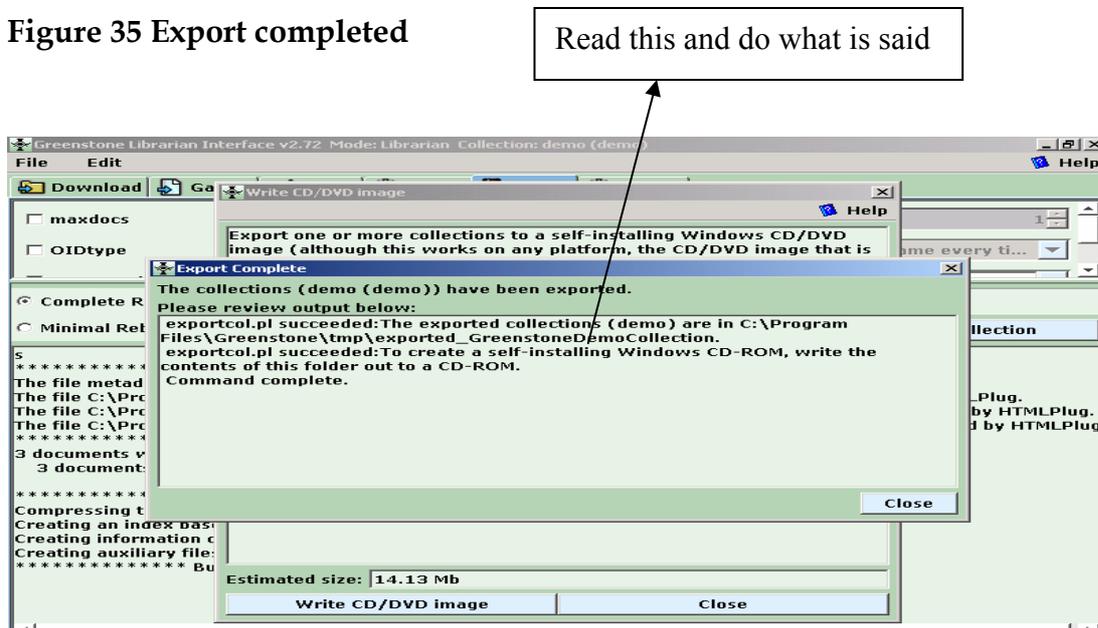
Figure 33 Starting the Writing Process

A pop-up window will appear as follows:-

Figure 34 Giving a Collection Name for CD-ROM

CD-ROM name

Put a tick mark to select the collection

Click here to Write CD/DVD image

Provide a name for your CD-ROM, mark the check box pertaining to the collection and click *Write CD/DVD image* button.

Figure 35 Export completed

Read this and do what is said

Click **Close** button.

Then write the contents of the folder (C:\ProgramFiles\Greenstone\ tmp\exported_Winisis) into a blank CD-ROM for creating self-installing Windows CD-ROM library.

3. Conclusion

The basic greenstone collection with normal *look and feel* can be created without much time. The digital library collection created thus can be customized later in a variety of ways. The customization can be done, whenever you like, by opening the collection, which can be done by clicking on the menu, **File ⇒ Open** and selecting the appropriate collection in the GLI**.**

You can radically alter the style of the pages generated, by changing the *macro* files included in the 'macro' folder of the Greenstone installation.

A general-purpose digital library like greenstone is a useful tool to provide information services in our libraries. Absence of knowledge on how to use it, should not come in the way of exploiting the advantages it offers. This documentation may be used as a tool to bring in more people to the growing constituency of greenstone users. We, the librarians, can improve our capabilities, as knowledge managers, if we are particular in learning the information technology tools like greenstone and use them for managing knowledge resources. We should learn, utilise, promote and propagate greenstone to make our libraries better and transform ourselves as knowledge mangers rather than mere book keepers.

References

1. Witten, Ian H and Brainbridge, David(2003) : *How to build a digital library*. Amsterdam, Morgan Kaufmann

2. Witten, Ian H and Boddie, Stefan (2004): *Greenstone Installer's Guide*. New Zealand Digital Library Project, New Zealand.
   http://prdownloads.sourceforge.net/ greenstone/Install-en.pdf

3. Brainbridge, David *et al* (2004): *Greenstone Digital Library Developer's Guide.* New Zealand Digital Library Project, New Zealand.
   http://prdownloads.sourceforge.net/greenstone/Develop-en.pdf

4. Witten, Ian. H. *et al* (2006) *Greenstone User's Guide*. New Zealand Digital Library Project, New Zealand.
   http://prdownloads.sourceforge.net/greenstone/User-en.pdf

5. Loots, M. *et al* (2001) *From Paper to Collection*. New Zealand Digital Library Project, New Zealand.
   http://prdownloads.sourceforge.net/greenstone/Paper-en.pdf

6. Morete, Pablo and Rose, John: *Creating Digital Libraries Based on CDS/ISIS Databases*.
   http://greenstonesupport.iimk.ac.in/Documents/CDS-ISIS_to_DL.pdf

7. Tutorial exercises on Greenstone at
   http://greenstone.sourceforge.net/wiki/index.php/Tutorial_exercises

****