

**Procedures for Maintaining
The Online Bibliographic Database
Health and Retirement Study**

**Michael A. Nolte
Research Area Specialist Lead**

**September 2010
Revised June 2012**

Table of Contents

I. Apply Appropriate Search Methodology.....	3
II. Update the EndNote Database.	4
III. Extract and Process EndNote Content for Export	6
Step 1: Export EndNote Records.	6
Step 2. Run the Parser Script to Create Access Input Files.	7
IV. Perform Microsoft Access Work.	8
Step 1. Import and Process Files	8
Step 2. Review Keyword Processing.....	8
Step 3. Check Your Work.....	9
Step 4. Export Bibliographic Database Summary Information	10
Step 5. Generate a Static Web Page Listing.	11
Step 6. Update Bar Chart.....	12
V. Final Steps.	12
Step 1: Check summary tables and graphs.	12
Step 2: Check current status of local copies of researcher papers.....	13
Step 3: Make Additions/Changes to the document database	13
Appendix A: Medical Keywords To Eliminate	14
Appendix B: Overview of the Bibliography Update Process	15
Appendix C: Content Area List.....	16
Appendix D: Setting Up a ProQuest Email Alert.....	17
1. Initiate your search.....	17
2. Review results; if satisfactory, select Set Up Alert.....	17
3. Answer the pop-up window questions	18

I. Apply Appropriate Search Methodology.

Keeping the HRS Bibliographic database up-to-date depends on the development and use of an effective search methodology that adds new and updated content to on a regular basis.¹ There are two primary goals that the search methodology must meet. First, it must ensure that as many of the HRS-related publications in existence are catalogued. Second, the process (including search terms) must be clear and understandable so that future searches can easily replicate the searching method. Recommended procedures are summarized below:²

1. [University of Michigan Library](#)³
 - Select: *Search tools: Articles and Databases* > {login with ITD username} > *Power Search*
 - Choose *advanced search*
 - Select database set from *My Databases*⁴
 - Check box next to each desired database (if no box appears, the meta-search feature is not available for the database)
 - Search terms (all fields): **health and retirement study** AND Year: {desired year}; **assets and health dynamics** AND Year: {desired year}
2. [Cambridge Scientific Abstracts](#)
 - Use *Quick search*
 - Select databases to be searched (most hits will be produced by *AGELINE*, *MEDLINE*, and *CSA Sociological Abstracts*)
 - Choose *continue to search*
 - Enter search terms: "**health and retirement study**" or "**assets and health dynamics**"
 - The number of citations will be listed for each database and each publication category.
3. [Web of Science/ISI Web of Knowledge](#)
 - On the Web of Science page, uncheck *Arts and Humanities Citation Index*
 - Use *Quick search*
 - Search terms: "**health and retirement study**" or "**assets and health dynamics**"
4. [Science Direct](#)
 - Uncheck *Preferred web sources* and *other web sources*, leaving *All journal sources* checked
 - Check *exact phrase*
 - Search terms: **health and retirement study** or **assets and health dynamics**
5. [EconLit \(via SilverPlatter\)](#)
 - Search terms: "**health and retirement study**" or "**assets and health dynamics**"
 - Limit search to *published after 2000*

¹ We are indebted to Jonathan Stone for the original version of this search tool summary. See his report entitled *The Benchmarking of HRS Publications* (working paper, Summer 2003) for further details.

² Be sure to log in to your University of Michigan Mirlyn account before attempting to follow any of these links

³ The UM library searches can be configured to provide email alerts; see Appendix D for details

⁴ We assume that the *My Databases* feature has been set up for your uniqueness, and that you are logged in to Mirlyn

6. [ProQuest Research Library](#)
 - o Choose *advanced search*
 - o Search terms: "**health and retirement study**" or "**assets and health dynamics**"
 - o Select desired databases
 - a. Interdisciplinary — Research Library
 - b. Interdisciplinary — [Dissertations and Theses](#)
 - c. Interdisciplinary — Gender Watch
 - o Limit by date if necessary
 - o Limit search if necessary
7. [JSTOR: The Scholarly Journal Archive](#)
 - o Choose *advanced search*
 - o Exact Phrase: "**health and retirement study**" or "**assets and health dynamics**"
 - o Select disciplines: **Business, Economics, Law, Political Science, Population Studies, Psychology, Public Policy & Administration, Sociology** (note that the lists of journals overlap)
 - o Limit by type (articles, reviews, opinion pieces, other) as desired
 - o Limit by date if necessary (JSTOR lags current date by 2 years)
8. [PsycINFO](#)
 - o On EBSCO/Host page select *Advanced Search*
 - o Select *in PsycINFO*
 - o If you *Select a Field* choose the TX option
 - o Search terms: "**health and retirement study**" or "**assets and health dynamics**"
 - o Limit by date if necessary
9. [Family & Society Studies Worldwide](#)
 - o Defaults to *Basic Search*
 - o Select *contains phrase*
 - o Search terms: **health and retirement study** or **assets and health dynamics**
 - o Specify *publication year* to limit by date
10. [Lexis-Nexis](#)
 - o Date: **Previous two years** (change to limit results)
 - o Search terms: **health and retirement study** or **assets and health dynamics**
11. [Popline](#)
 - o Choose *advanced search*
 - o Search terms: **health and retirement study** or **assets and health dynamics**
 - o Check *Peer reviewed journals* to limit search results
12. [Google Scholar](#)
 - o Search terms: "**health and retirement study**" or "**assets and health dynamics**"

II. Update the EndNote Database.

Make all needed changes to the EndNote database located in the V:\biblio folder. All entries should have keywords, abstract (if available) and a unique ProCite style reference number⁵ in addition to

⁵ ProCite reference numbers are stored in the EndNote LABEL field. They are used as reference links to the PDF images of research papers stored in the Document Database (See Section 5, Step 3 in this document).

HRS Website Documentation Series - 2

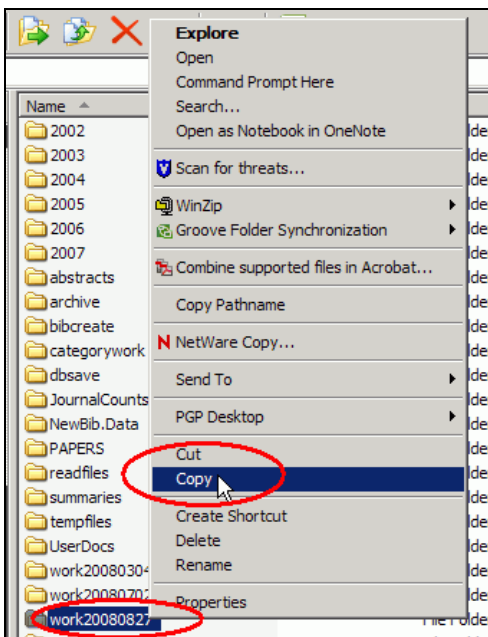
the default record identifier assigned by EndNote. If at all possible, use existing keywords as provided by EndNote. Remember, if you add new keywords, you'll need to reconcile them against the broad categories and sub-categories used in the HRS online bibliography; see table keyword_work2 in BibWork.accdb for details.

When you have completed your EndNote work, be sure to mark all records for export. You should remember the record count so that you can compare it to the number of records that appear in the new BiblioData table.

1996	Journal Article	https://hrsweb.isr.umich.edu/extern/pub:
2001	Journal Article	https://hrsweb.isr.umich.edu/extern/pub:
Duplicatess		
907 Records Marked	907 Records in List	907 Records in Database

Additional preparation work:

1. Create a new work subdirectory



2. Paste the copy and rename it to the current date

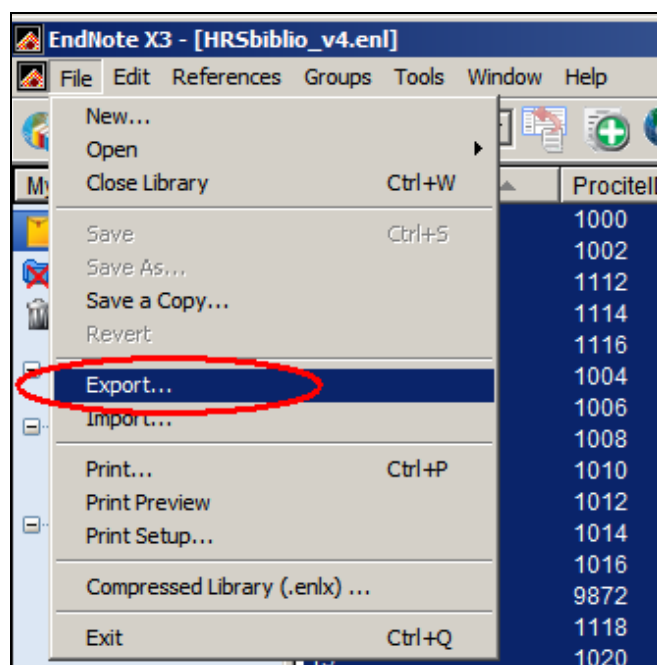
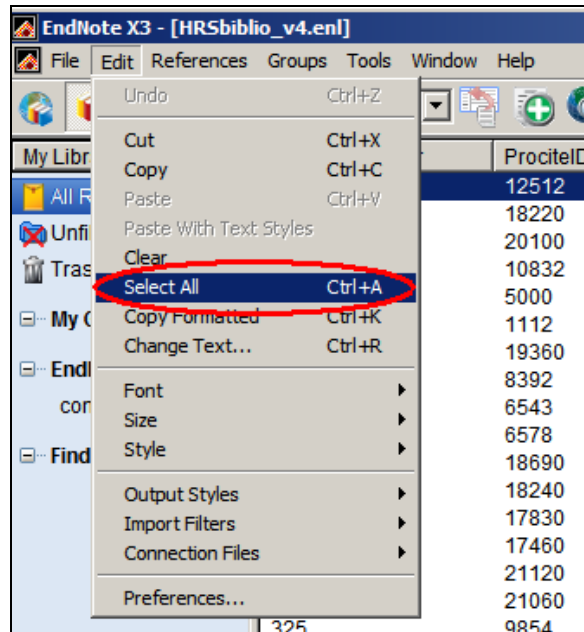


3. Copy key files and folders from the last work subdirectory to the new one
CumulativeWorkformV2_files
BibWork.accdb
CumulativeWorkformV2.htm
CumulativeWorkformV2.vsd

III. Extract and Process EndNote Content for Export

Step 1: Export EndNote Records.

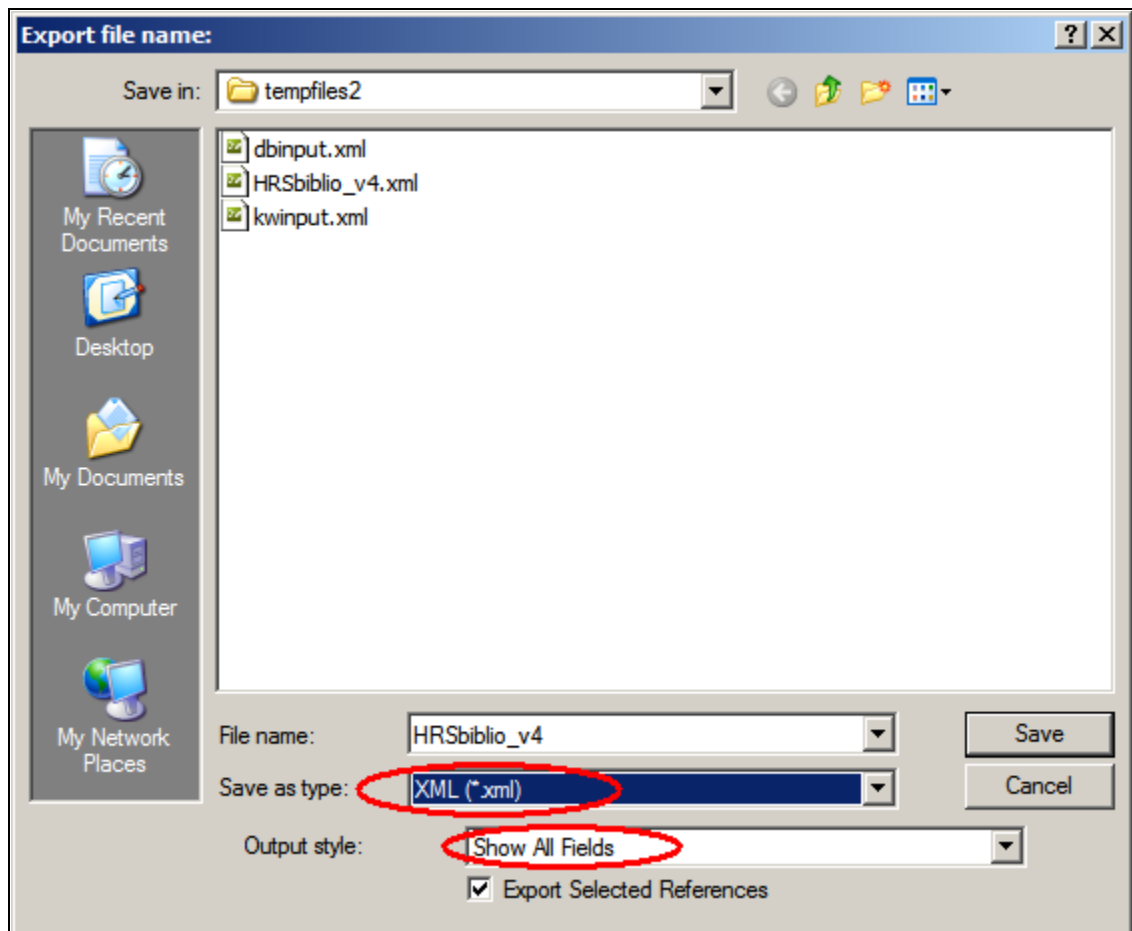
Export EndNote record fields (in XML format) for further processing and eventual inclusion in the Access work database. Be sure to select all records from the open database (*Edit > Select All*). To export the selected records choose *File > Export*.



HRS Website Documentation Series - 2

The records extracted from the EndNote database in this step will be used to produce both the online (dynamic) version of the Bibliography as well as the static Web page version.

Save the exported records in XML format as V:\BIBLIO\tempfiles2\HRSbiblio_v5.xml.



Step 2. Run the Parser Script to Create Access Input Files.

At the MSDOS command line prompt, run the batch file containing Pearl parser scripts (runparse.bat):

```
[perl] xmlcleanv3.pl V:\BIBLIO\tempfiles2\HRSbiblio_v5.xml
```

These programs will produce the following files in V:\BIBLIO\tempfiles2

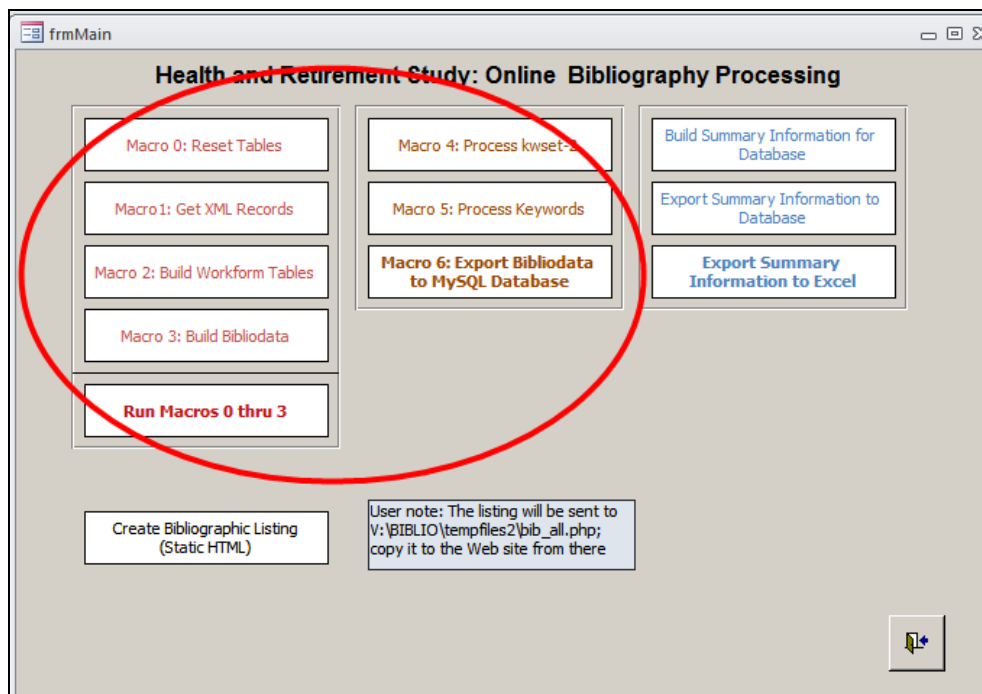
```
dbinput.xml  
kwinput.xml
```

IV. Perform Microsoft Access Work.

Step 1. Import and Process Files

In theory, this step is quite straightforward: First open the working Access database (e.g., V:\BIBLIO\work\ymmdd\BibWork.accdb). Now open form frmMain and run the following **macros** to convert EndNote information to a format suitable for use on the Web:

- 0 *Reset Tables* – resets workform and keyword tables from templates.
- 1 *Get XML Records* – imports the XML files (bibliography and keyword information) produced by the parser program.
- 2 *Import Tables* – build workform tables from imported XML information.
- 3 *Build BiblioData* – creates the BiblioData table.
- 4 *Process kwset-2* – first state in keyword creation.
- 5 *Process keywords* – second stage in keyword creation.
- 6 *Export Bibliodata to MySQL* – creates the Web site version of the BiblioData table.



Step 2. Review Keyword Processing

After you run Macros 4 and 5 and before you execute Macro 6 you should update the tables used in keyword standardization. If you have added a keyword to any new or existing EndNote record, the keyword value must be standardized and linked to a content area. To obtain a list of keywords requiring standardization open the query *Check keyword status*.

What's happening underneath: Macro 4 creates table kwset-2 from the keyword file V:\BIBLIO\tempfiles2\kwinput.xml which was previously extracted from an XML dump of the EndNote database (HRSbiblio_v5.xml) by Perl script xmlcleanv3.pl. Macro 5 standardizes

(recodes) these keywords and links them to content areas that match the retrieval options in the Dynamic Bibliography. For this recoding process to work properly, an additional table called `Keyword_work2` is used to establish the cross-reference between the original EndNote keywords and the HRS categories. An excerpt from this table appears below:

keyword	Kupdate	xcode	FirstIDNu	contentarea
2007-2009 recession	Great Recession			Public Policy
401(k) participation	401(k) participation and balances			Pensions
401(k) participation and balances	401(k) participation and balances		6648	Pensions
401(k)s	401(k) participation and balances			Pensions
401k	401(k) participation and balances		1128	Pensions
ability	Ability			Other
Absenteeism	Absenteeism		4385	Employment and Labor Force
academic achievement	Education			Demographics

Adding the New Cross-Reference Information: Begin by making a backup copy of tables `Keyword_work2` and `Keyword_work2_update`. Then use query *Make:Keyword_work2_update* to create a permanent version of the records found by *Check keyword status*. Edit this table, adding appropriate cross-reference information in the `kupdate`, `xcode` and `contentarea` fields (field `FirstIDnum` can be left blank). For example, the addition of the keyword “401(k) procedures” would be standardized as follows:

401k procedures	401(k) participation and balances		1128	Pensions
-----------------	-----------------------------------	--	------	----------

Once you have completed your editing of `Keyword_work2_update`, append the contents of the table to `Keyword_work2`.

Here are some guidelines for assigning the `kupdate`, `xcode` and `contentarea` field values for a new keyword:

- If the new keyword is not unique, (it seems close to another keyword) assign an existing `kupdate` value.
- If the new keyword is unique, give `kupdate` the value of keyword
- Content area values are fixed; you must fit the new keyword into one of them (see Appendix C for a list).
- If the new key has an alphanumeric suffix (e.g. “Adult Children (D009500)”), copy it to `xcode`.
- Some keywords can be dropped (see Appendix A)
- If you can’t make a decision, talk to a domain expert.

Step 3. Check Your Work

Before you update the web server database you need to be sure that you have created a correctly populated and formatted version of the `BiblioData` table. Make sure that you have the correct number of records, and that each record contains all necessary information. You may wish to select a sub-sample of records (4 or 5 percent) and compare them to the EndNote original. If all seems well, then execute the final macro: *6 Export BiblioData to MySQL* in order to update the Web server database:

Step 4. Export Bibliographic Database Summary Information

Once the EndNote database records have been imported and processed, and BiblioData has been exported to the MySQL database, you should complete the process by building and adding summary information.

Three tables are created and exported in this process:

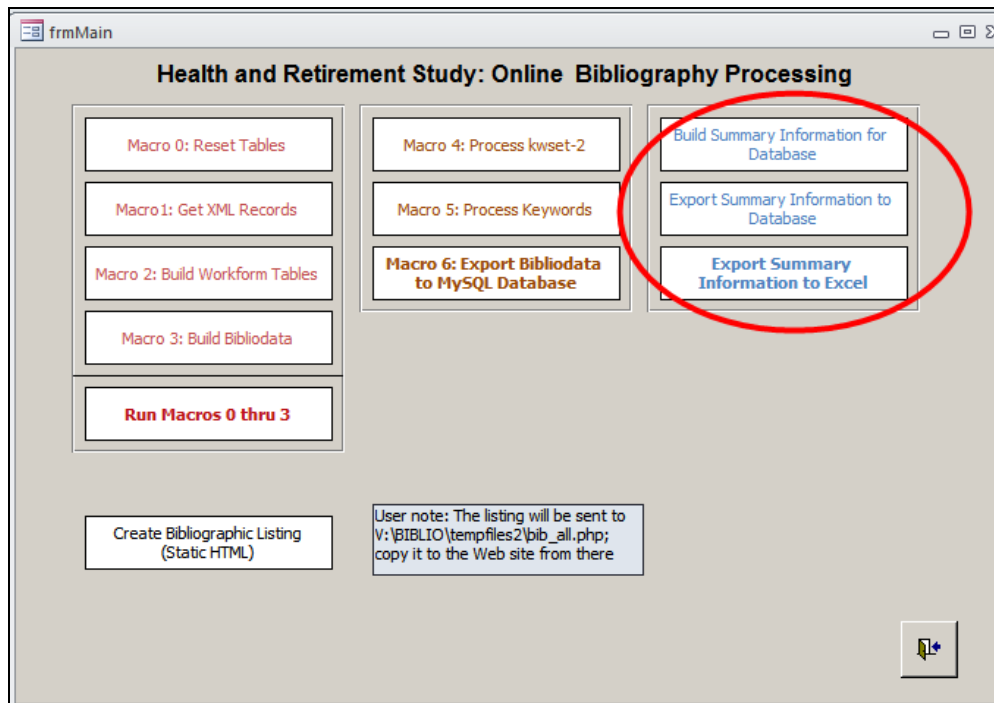
- BiblioData_Workform-Year
- BiblioData_Year-Workform
- InterestAreas_Year

These tables summarize the content of the database by year (beginning in 1993) for selected *workforms* and *interest areas* (see below).

Workform	Interest Area
Book Chapter	Economics
Book Whole	Family
Dissertation	Health
Journal Article	Health Services
Magazine Article	Labor
Newspaper	Methods
Report	

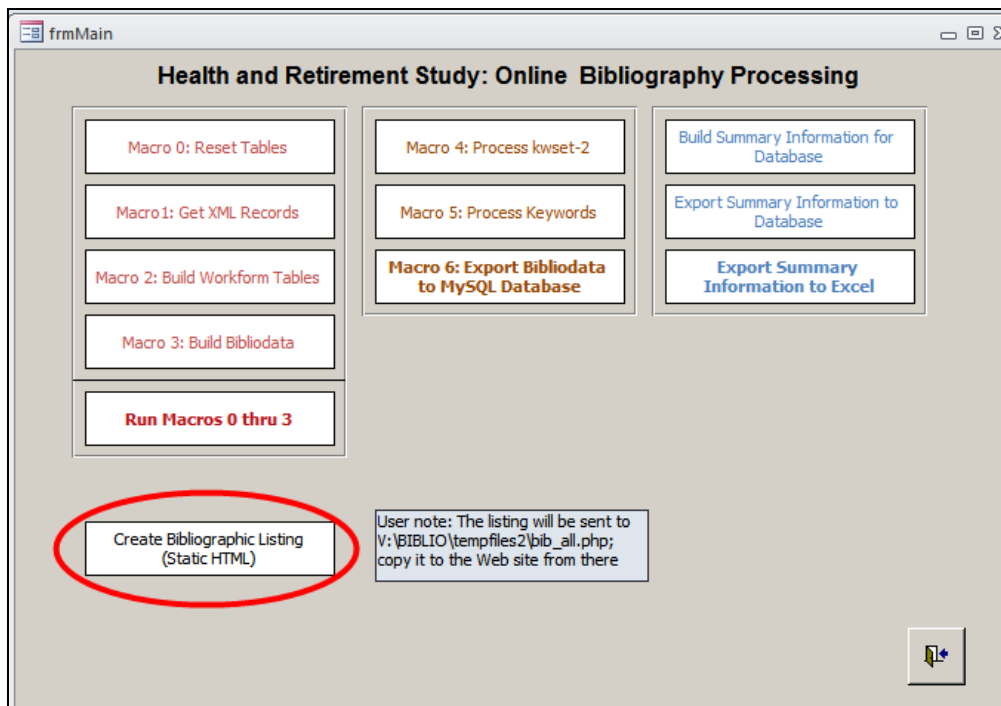
- *Build Summary Information for Database* – create the tables noted above
- *Export Summary Informatin to Database*⁶ – drop the existing versions of BiblioData_Workform-Year, BiblioData_Year-Workform, and InterestAreas_Year from the MySQL database and export new versions.
- *Export Summary Information to Excel* – this process updates an Excel table (V:\BIBLIO\tempfiles2\test.xls) containing workform-year summary information. This information is used with Visio to create a bar chart showing a cumulative count of HRS publications. See Step 5, below.

⁶ The query *Summary 2b: Export (No Drop) Bibliography Information* exports the summary information directly to the MySQL database without dropping the corresponding tables.

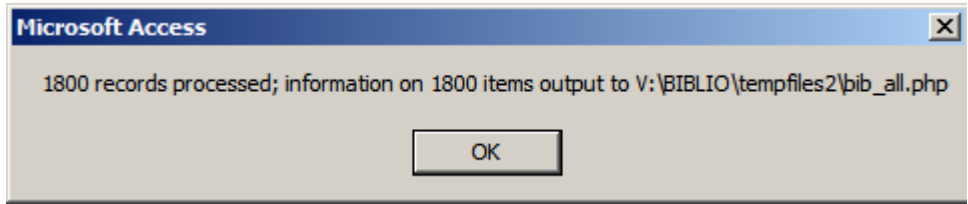


Step 5. Generate a Static Web Page Listing.

The final step in the bibliography process is to create a subject bibliography in static HTML format to be linked from the [Publications Page](#) of the HRS Web site.



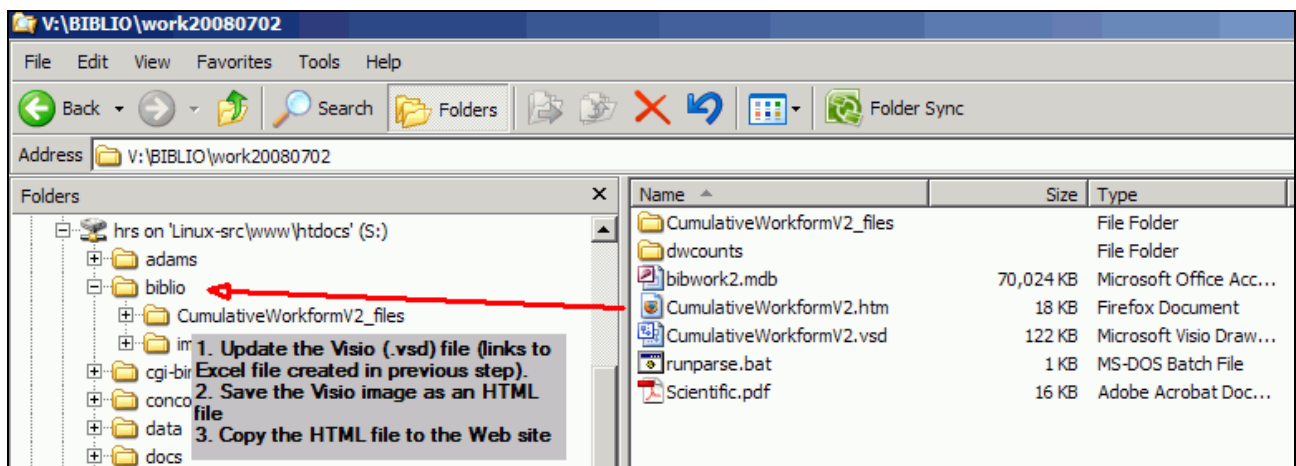
Click the *Create Bibliographic Listing* button; when the program terminates, click *OK*.



Now you can copy V:\BIBLIO\tempfiles2\bib_all.php to the **pubs** folder on the HRS Web site:

```
\\linux-src\www\vhosts\hrsonline.dir.umich.edu\html\modules\pubs\bib_all.php
```

Step 6. Update Bar Chart



The Excel file (V:\BIBLIO\tempfiles2\test.xls) mentioned above is designed to automatically produce a bar chart that displays a cumulative count of journal publications, books and book chapters, dissertations, and working papers produced by researchers using HRS data products. This chart can be manipulated via an OLE link from Visio. Assuming that the chart is correct, all that is needed is for you to annotate the Visio version of chart as appropriate. Once that task is complete, the Visio file can be saved in HTML format for use on the HRS web site. Just copy CumulativeWorkformV2.htm to:

```
\\linux-src\www\vhosts\hrsonline.sir.umich.edu\html\modules\biblio
```

V. Final Steps.

Step 1: Check summary tables and graphs.

These tables are displayed on the Public Web site ([Site Statistics](#)) and Staff Web site ([Usage Statistics](#)). The numbers reflected in these tables should match the information in the Endnote and Access databases.

- Count by Workform

- Count by Interest Area

Step 2: Check current status of local copies of researcher papers.⁷

Under the terms of use for HRS data, researchers agree to provide copies of all papers and other publications. These documents are maintained on line for use by HRS staff members.

- *Full Listing* - Display all records in table *papers*.
- *Storage Summary* - Publication/Working Paper count.
- *Problem Listing* - Display records with a null value for the PDF Blob field.

Note: Status reports may take 10 or more seconds to run

Step 3: Make Additions/Changes to the document database⁸

This process requires that you know the EndNote ID linking number (*ProciteID*) that corresponds to the paper to be added. Before inserting papers for Web use, they should be collected in folder <\\src-hrs\hrs\BIBLIO\PAPERS\newpubs>.

- *Add* – Insert a PDF version of a research paper or other document into the database.
- *Drop* – Remove a document from the database.
- *View* – View the document associated with a given citation identifier
- *Move* – Move or copy the document associated with a given citation identifier to an archival table. You should use this feature for documents that are being replaced by a newer version; for example, a working paper replaced by a journal article.

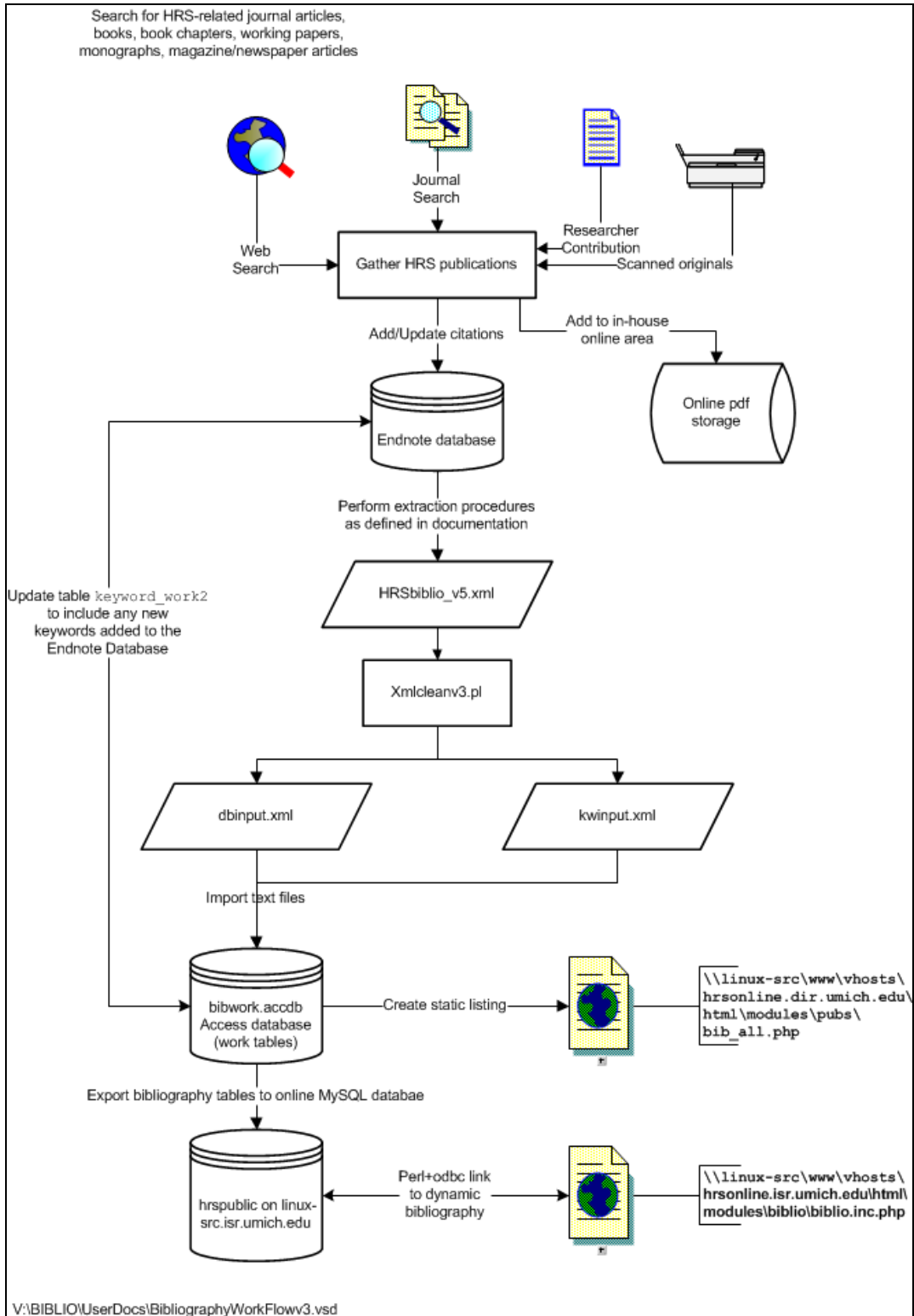
⁷ See [Bibliography Maintenance](https://ssl.isr.umich.edu/hrs-extern/gri45/index.html) (https://ssl.isr.umich.edu/hrs-extern/gri45/index.html)

⁸ Documents to be displayed online are stored in table *papers* in MySQL database HRSONline2; archived documents are stored in table *papers_old*.

Appendix A: Medical Keywords To Eliminate

8, 24, animal, mouse, rat, rabbit, rickets, proteins, lipoproteins, lipids, marker, markers, chemical, acute, water, reproduction, tumor, alpha, ascites, cirrhosis, ventilation, cell, test, case, patient, postoperative, bispectral, creatinine, ventricular, vitro, utero, ultrasound, hydronephrosis, unilateral, bilateral, nitric, IV, intravenous, lesion, hypertensive, amplitude, perfusion, metabolism, dose, doses, sodium, receptor, transplant, transplantation, adolescent, pregnancy, fertility, radiation, etiology, gene, genes, genetic, acute, post-natal, potassium, administered, administer, post-operative, thrombin, laparoscopy

Appendix B: Overview of the Bibliography Update Process



Appendix C: Content Area List

ADL/IADL
Consumption and Savings
Cross-National
Demographics
Disabilities
Employment and Labor Force
End of Life Decisions
Event History/Life Cycle
Expectations
Family and Transfers
Genetics
Health Care
Health Conditions and Status
Housing
Income
Insurance (General)
Medicare/Medicaid/Health Insurance
Methodology
Net Worth and Assets (Capital Gains)
Pensions
Public Policy
Retirement Planning and Satisfaction
Risk Taking
Social Security
Time Use
Women and Minorities

Appendix D: Setting Up a ProQuest Email Alert

1. Initiate your search

ProQuest

Basic Advanced Publications My Research
0 marked items

Databases selected: Dissertations & Theses

Advanced Search Tools: [Search Tips](#)

search term

"Panel Study of Income Dynamics" Citation and abstract

AND Citation and abstract

AND Citation and abstract

[Add a row](#) | [Remove a row](#) **Search** **Clear**

Database: Interdisciplinary - Dissertations & Theses [Select multiple databases](#)

Date range: All dates

Limit results to: Full text documents only

[More Search Options](#)

Copyright © 2007 ProQuest Information and Learning Company.

2. Review results; if satisfactory, select Set Up Alert

ProQuest

Basic Advanced Publications My Research
0 marked items

Databases selected: Dissertations & Theses

Results

346 documents found for: ("Panel Study of Income Dynamics") » [Refine Search](#) | [Set Up Alert](#) ✕

Dissertations

Mark all [0](#) marked items: [Email](#) / [Cite](#) / [Export](#)

1. [Asset pricing and asset allocation in the presence of durable consumption goods](#)
by Siegel, Stephan, Ph.D., **Columbia University**, 2006, 100 pages; AAT 3213594

[Abstract](#) [24 Page Preview](#)

3. Answer the pop-up window questions

The screenshot shows the ProQuest website interface for setting up a search alert. At the top left is the ProQuest logo, and at the top right are links for Help, Privacy Policy, and Close window. The main heading is "Set up Your Search Alert" in red. Below this is a paragraph explaining that new documents matching the search will be sent to the provided email address. The form includes several sections: "Search" with pre-filled text: ("Panel Study of Income Dynamics"), Database: Dissertations & Theses, Look for terms in: Citation and abstract, and Publication type: All publication types; "Schedule" with a dropdown menu set to "Monthly" and a checkbox for "Send message when there are no new results"; "Stop after" with a dropdown menu set to "2 months" and a note "(alerts can be renewed.)"; "Email address" with an empty text box and a note "(Enter a single email address)"; "Subject" with an empty text box starting with "("; "Message" with a large empty text area and a note "(optional - enter a note to be included with your results)"; and finally, "Save" and "Cancel" buttons at the bottom.

ProQuest [Help](#) | [Privacy Policy](#) | [Close window](#)

Set up Your Search Alert

As new documents matching your search are added to our database, ProQuest will automatically send new results to the email address you provide below. Choose your alert options and click the Save button.

Search: ("Panel Study of Income Dynamics")
Database: Dissertations & Theses
Look for terms in: Citation and abstract
Publication type: All publication types

Schedule: send this alert : Monthly
 Send message when there are no new results.

Stop after: 2 months (alerts can be renewed.)

Email address:
(Enter a single email address)

Subject: ()

Message:
(optional - enter a note to be included with your results)

That's all there is to it!