
Cataloging Manuscripts for HMML: Using the Remote Cataloging Workbook

Copyright© 2008 Hill Museum & Manuscript Library

The Hill Museum & Manuscript Library maintains a large, comprehensive database of the manuscripts it has preserved on microfilm and in digital form. Until now, the entry of information into HMML's electronic catalog had to be done on-site, using one of the networked computers at our library. With more and more material coming from our overseas projects, a method was needed to allow multiple catalogs to work on collections *remotely*, with the data appended to our main database after it arrives at HMML.

HMML's master manuscript catalog is a complex relational database originally designed by a consortium primarily interested in the cataloging of western medieval manuscripts. Attempts have been made to adapt our complex, Access-based data input forms for "portable" use to allow people off-campus to do this work. Something simpler and less dependent on specialized software was called for. The result is this cataloging workbook created in Microsoft Excel, which has been further adapted for use with non-western manuscripts.

The Excel workbook titled "2009 Cataloging Workbook2.xls" is the data-entry tool for manuscript metadata. It presents two Excel spreadsheet forms for entering information about various aspects of a given manuscript.

It is intended that each individual manuscript will have its own workbook. This workbook is analogous to the inventory cards produced by HMML in its early days of cataloging European manuscripts. Data is entered in the workbook pertaining to the manuscript; after it returns to HMML, various steps are taken to put the data in tabular form for appending to the main manuscript database.

Some Cautionary Notes:

- Save the workbook with another name (the manuscript's source code) so that the original, blank workbook can remain unchanged and be used again and again.
- Do not add or remove worksheets from the workbook.
- Do not alter the colors of cells or the position of items in the worksheets.
- Do not add or delete rows or columns in the worksheets.
- Do not change the font used in the workbook.
- The worksheet and field names cannot be altered.
- The cataloging workbook is designed to work well on a computer system having a monitor resolution of at least 1024x768 pixels. It's easier to use on a system with

a monitor resolution in the 1200-pixel range, which is typical of today's laptop computers.

- Except for the “General Info” field, our master database can only handle fields containing fewer than 256 alphanumeric characters. One needs to keep field contents concise.

HMML uses a relational database for its electronic manuscript catalog. Manuscript metadata is organized into two main levels:

- **The manuscript itself**—there is always one and only one of these per workbook.
- **Manuscript Text**—the actual written text(s) within the manuscript. Many manuscripts will have more than one text; some have dozens.

This diagram might clarify the issue a bit:

MANUSCRIPT

TEXT1

TEXT2

TEXT3

And so on.

Each of these metadata levels has its own worksheet in the cataloging workbook. These worksheets are indicated by labeled tabs at the bottom of the workbook window. The tabs are labeled as such:

- **Manuscript:** Top-level data pertaining to the manuscript as a whole is entered in this sheet.
- **Manuscript Text:** Information about the “works” written. There are columns prepared for up to fifty texts or works.

The worksheets have data entry areas, field names and short descriptions of the intended field contents. Not all fields need be filled in; required fields are indicated.

On “native script” field entries:

In the worksheet for texts, there are several fields names indicating “native script.” These are for entries done in the actual native scripts of the language used in the manuscript. These are optional, but very nice to have. The use of the “Arial Unicode MS” font for the worksheet allows the user to enter data in Arabic, Armenian, etc, without having to use

different typefaces. HMML wants its native alphabet entries in Unicode format so that the information is easily transferable and ready for use on the Internet.

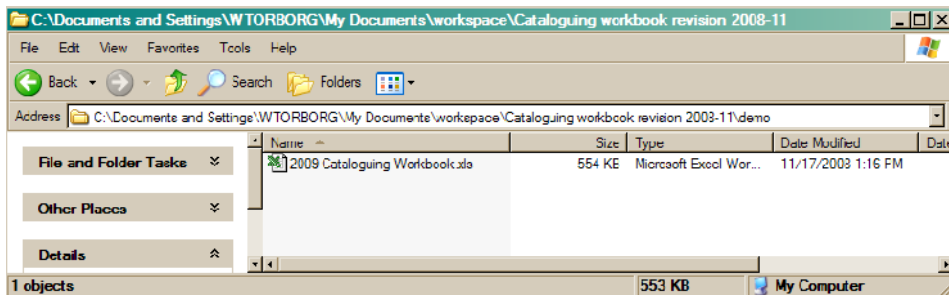
The Arial Unicode MS font is quite comprehensive in the number of font characters it contains, but it still doesn't have some of the more uncommon scripts, such as Syriac and Ge'ez (Ethiopic). For typing in Syriac, Microsoft Windows includes a font titled "Estrangelo Edessa." For creating metadata in Ge'ez, a font called "Titus Cyberbit Basic" (TITUSCBZ.TTF) is very useful. Copy this font file to C:\Windows\Fonts to install it. Consult HMML on ways to obtain any necessary fonts.

Using the cataloging Workbook

The workbook is supplied to the user on CD-ROM or via the internet as a Microsoft Excel 2003 file with the extension ".xls." This allows it to be used in a variety of programs, including:

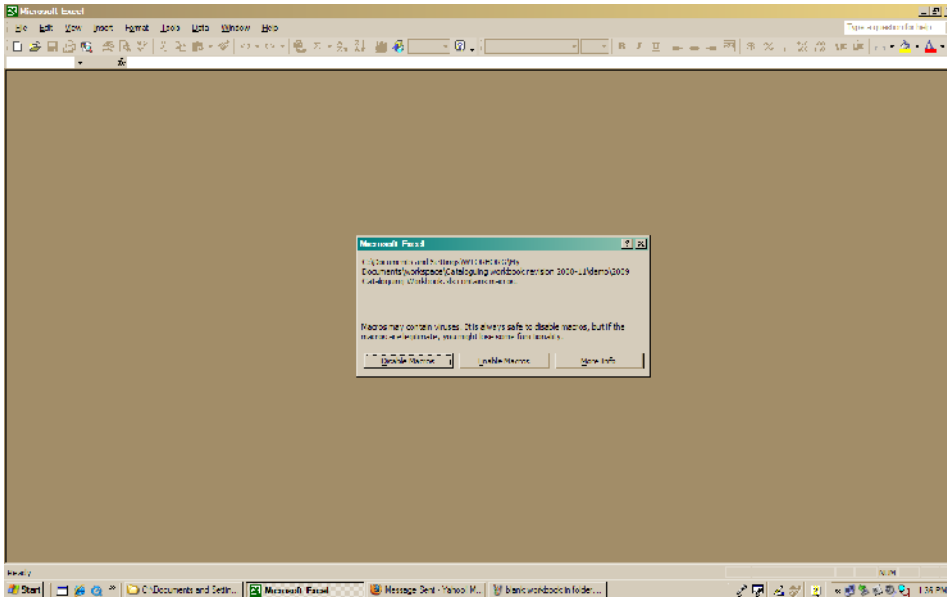
- Microsoft Excel XP/2003 (Windows)
- Microsoft Office Excel 2007 (Windows)
- Microsoft Excel (Macintosh)
- OpenOffice.org Calc (Windows, Macintosh, Linux)

In most cases, the user can simply double-click on the file icon to launch the appropriate software application.



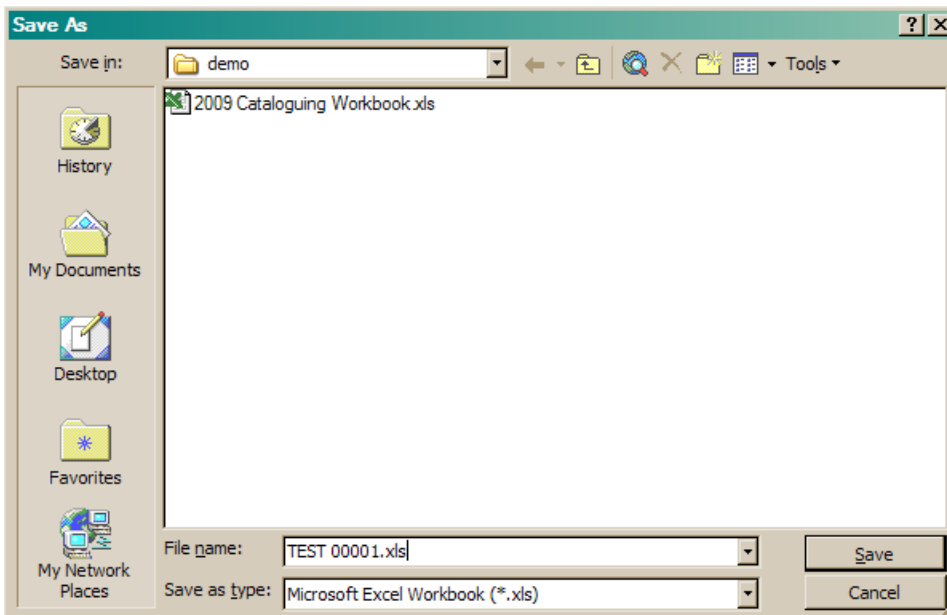
Blank cataloging Workbook Ready to be Opened in Excel. The user will never actually enter information into the master workbook copy; it will be renamed prior to data entry.

Depending on the application the user chooses to open the workbook in, a warning might pop up indicating that the workbook contains *macros* (small bits of programming included in the document). This is normal; the user can disable the macros, as they are only used by HMML personnel after the workbook is returned.



Macro “Warning” When Opening File in Excel 2003. The user can simply click the “Disable Macros” button and start working. The embedded macros in the file are used at HMML to create special “append tables” to add the metadata to its master manuscript database. Other software programs will generally disable the macros by default.

Before the user enters any manuscript metadata into the workbook, it must be immediately re-saved with a different filename to keep the master copy untouched.



Saving the Excel Workbook With a Different Name. Here, the file has been renamed as “TEST 00001.xls.” The actual filename to use is derived from the HMML “source number” of the manuscript being described.

The name given to the cataloging workbook is usually the “HMML source number” of the manuscript itself. Here's a short explanation of the source number from another HMML instructional document:

The projects partnering with HMML are given a simple alphabetical code agreed upon by HMML and the holding institution. This is an important thing; this code is used in a number of ways:

- It's used as the filename prefix for naming the image files
- It is part of the HMML “source” number for a particular manuscript
- Used to label folders to identify specific manuscripts

Some examples of HMML project codes:

Project Code	Code Translation	Location of Project
OLM	Order of Lebanese Maronites	USEK, Kaslik, Lebanon
CFMM	Church of the Forty Martyrs	Mardin, Turkey
MLRI	Muzeul Literaturii Române Iași	Iași, Romania

Each individual manuscript is assigned a “source number” based on this prefix. Examples:

OLM 00256
CFMM 00015
MLRI 00006

The Prefix is written in capital letters, followed by a space. The number following the prefix is *always* written out with five “placeholders,” allowing up to 99,999 manuscripts to be identified for a specific collection. This is the information entered in the “source” field for HMML's metadata systems.

A particular file within a manuscript folder might have a name such as:

OLM_00357_005r_004v.JPG

This file is from the OLM project, number 357. The image is of a double-page spread, folios 5 recto and 4 verso. It is JPEG image file.

Thus, the workbook for this particular manuscript would be named, “OLM 00357.xls” Remember to include the space after the project code prefix.

With the workbook renamed properly, the user can begin the cataloging process. The work doesn't have to be done in any particular order; the user can work on descriptions of the overall manuscript or the texts within the manuscript.

Microsoft Excel 11-51 00001.xls

Cataloguing Workbook for Hill Museum & Manuscript Library

Table: Manuscript

Information pertaining to the complete manuscript and its components

Description of Field Contents	Required	Database Field Name	Cataloguer Input
The name of the city where the manuscript is held. Use the local spelling of the name. Use Latin alphabet.	Yes	City	
The name of the library where the manuscript is held.	No	Library	
Library shelfmark for codes. Use "NY" if no shelfmark is available.	Yes	ShelfMark	
In reference to the bound leaves of the entire codex: the number of front leaves (lowercase roman), plus number of leaves in arabic, plus number of back leaves (lowercase roman), parenthesized as total not as span. Use the same formula for pagination if the manuscript is paginated not numbered. Example: f. 1-100v, f. 101-140p, etc.	Yes	No. of Folios	
General statements on date, type of covering, boards, lumbare (leaves, boards, covers), binder's name if known, references to a comment on early observed aspects of the binding. Note: Focusing is contemporary or early. If none, please to not bind "enter text board". Other examples: parchment, boards, cardboard, etc. Note any.	No	Binding	
Successive owners, agents and individuals. Give names of owners, separated by semicolon, and source of information (such as colophon).	No	Provenance	
Bibliographic citation pertaining to the manuscript as a whole. Cite for publications about the entire manuscript (not). Publications about the text govern the comments for Manuscript text.	No	Ubiography	
A user defined field to be adapted to the needs of a specific manuscript. For example, general information about the location of the manuscript, its date of creation, its genre, etc. Note: Use the user data on ongoing.	No	general info	
USMNL Accession Number: 11-51 0001 0011 The end identification of the text.	Yes	Number	

Manuscript Worksheet in cataloging Workbook. The white cells under the column labeled, "catalog Input" are filled in with the appropriate metadata. Not all fields are required to produce an acceptable record. A column labeled, "Description of Field Contents" guides the user as to the type of information needed for each field.

Microsoft Excel 11-51 00001.xls

Table: Text

Information pertaining to the written texts or works within the manuscript

Description of Field	Required	Database Field Name	Cataloguer Input Text 1	Cataloguer Input Text 2
In reference to the individual text, the number of leaves (p. 1-7v)	Yes	Pages Span of Text		
Name(s) of author(s) (how the name is written in the text or in the title, and the name of the author (the name of the author, etc.) if not stated for the original author(s).	No	Author		
Name of author (redline script)	No	Author (redline script)		
Name of author (search record) of all the text in the text. Use for translation, author, etc. It should be the role in parentheses.	No	Other Authorial Name		
In English: the common descriptive title of the work provided by cataloguer. Use standard terminology and style if possible. Includes generic titles such as "Title", "Introduction", "Index", etc.	Yes	Generic Title		
Title taken from the work entered in script used in work (native script)	No	Title (native script)		
Final description and reason for any missing text (change and/or incomplete, determine fragment, signed text or cover, missing, media of bound, missing)	No	Condition of Text		
Author's initials (the text description of the manuscript) or the name of the author, including the name of the author (the author's initials)	No	Initials (redline script)		

Text Worksheet in cataloging Workbook. Here, one can see data input space for two texts within the manuscript. Scrolling to the right, the user will find space for up to 50 texts to be described.

When data entry is complete, the workbook is saved and closed. The original, blank workbook is then opened and renamed for the next manuscript. As with any computer program, it is recommended that the user save the file at regular intervals to avoid data loss in case of computer crashes, power outages, etc.

Guidelines for Manuscript Metadata Entry

Worksheet: Manuscript

Database Field Name: City

Required: Yes

Contents: The name of the city where the manuscript is held. Use the local spelling of the name. Use Latin alphabet.

Database Field Name: Library

Required: No

Contents: The name of the Library where the manuscript is held.

Database Field Name: ShelfMark

Required: Yes

Contents: Library shelfmark for codex. Use "NA" if no shelfmark is available.

Database Field Name: No. of Folios

Required: Yes

Contents: In reference to the extent of leaves of the entire codex: the number of front flyleaves in lowercase romans, plus number of leaves in arabics, plus number of back flyleaves in lowercase romans, phrased as total not as span. Use the same formula for pagination if the manuscript is paginated, not foliated. Examples: ff. iii+271+ii; pp. I+64. If the manuscript is not physically foliated, place supplied foliation in square brackets: ff. [iii+271+11]. The lack of foliation should be noted in the General Info field below.

Database Field Name: Binding

Required: No

Contents: General statement as to date, type of covering, boards, furniture, binder's name if known. Intended as a comment on easily observed aspects of the binding: note if binding is contemporary, or early. If manuscript is not bound, enter "Not bound." Other examples: parchment, boards, cardboard, etc. Note any pastedowns in field titled, "General info."

Database Field Name: Provenance

Required: No

Contents: Successive owners, private and institutional. Give names of owners, separated by semicolon, and source of information (such as colophon, folio number.)

Database Field Name: Bibliography

Required: No

Contents: Bibliographic citation pertaining to the manuscript as a whole. Only for publications about the entire manuscript itself. Publications about the *text* go into the comments for Manuscript Text.

Database Field Name: General Info

Required: No

Contents: General information about the production, physical characteristics, and history of the manuscript not entered in any other field. This can be written in prose or as a list, following this sequence, with each entry separated by a semicolon:

- Manuscript production: place where the manuscript was copied; faith tradition of scribe; exemplars; location of date in the manuscript.
- Manuscript characteristics: lack of physical foliation (noted as: “unfoliated, foliation supplied by catalogr”); general condition; pastedowns used in the binding, etc.
- Manuscript history: former shelf marks (olim); historical reports of the ms; known readers of the ms.

Any term(s) may be omitted.

Database Field Name: Source

Required: Yes

Contents: HMML source number. Exp: “CFMM 00002.” The actual filenames of the photographic manuscript page images will indicate this.

Database Field Name: Inputter & Date

Required: Yes

Contents: Name of the catalogr and the date when the work was done.

Database Field Name: Country

Required: Yes

Contents: Country where manuscript is presently held. Use the modern political name of the country.

Database Field Name: Colophon

Required: No

Contents: Enter the entire colophon using native script. If absent, indicate “no colophon” or “colophon missing.”

Database Field Name: Outer Dimensions

Required: Yes

Contents: Use centimeters. Exterior dimensions. Height by Width. Measure the support, not the binding. Take this information from the "slate" photo that should be included with page images. This image file should have an “s” prior to the filename extension; an example would be SOAA_0004(q)_s.JPG. This is a photograph of the original inventory card prepared by the photographer, and it should have the dimensions listed. If not, each image of the manuscript should have a centimeter scale included in the photo, allowing the dimensions to be determined.

Database Field Name: Country of Origin

Required: No

Contents: Country of origin of the manuscript. Name, in English, using modern political divisions; if the origin crosses modern political boundaries, choose one name and include a comment with the second country name in the “General Information” field.

Database Field Name: City of Origin

Required: No

Contents: This is the city of origin; enter if stated in colophon or elsewhere in the manuscript, modern political name.

Database Field Name: Century

Required: Yes

Contents: Approximate century of the manuscript's manufacture, expressed in western numerals, using the western calendar. Non-western dates are entered in “Non-Western date.” Examples: 18th century, 19th/20th century.

Database Field Name: YearMonthDay

Required: No

Contents: Use only when the manuscript provides an exact date of manufacture. Enter using western date; enter non-western date in the appropriate field. Examples: 1650 Jan 01; 1752. Enter in General Information the location of the date (the folio or the colophon)

Database Field Name: Layout

Required: No

Contents: Indication of ruling technique and layout. Brief description of columns, number of lines, or any other distinctive aspects of the appearance of the page, separated by semicolons. Examples: 2 cols.; 15 lines.

Database Field Name: Script

Required: No

Contents: Briefly identify the type of script.

Database Field Name: Number of Scribes

Required: No

Contents: Number of readily distinguishable hands with span of leaves copied by each person. Use if more than one scribe wrote the manuscript.

Database Field Name: Scribe

Required: No

Contents: Enter name of the scribe if available.

Database Field Name: Support

Required: Yes

Contents: Choices: Paper; Parchment; Paper and parchment

Database Field Name: Watermark

Required: No

Contents: The identifying mark used by the paper maker. Brief description, with ff #s where the watermark appears.

Database Field Name: Representational Decoration

Required: No

Contents: Painting or drawing in a manuscript that represents humans or animals, usually in miniatures (of any size or shape), historiated initials, or marginal scenes. Use to indicate presence of such illustrations. Give number and brief description of type, for example, “3 miniatures painted across the width of 2 columns of text; numerous pen and ink drawings, many incomplete.” It is not necessary to catalog all illustrations, but indicate their presence.

Database Field Name: Other Decoration

Required: No

Contents: Use to indicate any other significant decoration, i.e., types of initials, borders, rubrication. Type, nature, size, technique. Not necessary to count all!

Database Field Name: Artist

Required: No

Contents: Name of artist/decorator. Only if the name of the artist is identified.

Database Field Name: Acknowledgments

Required: No

Contents: Public recognition of contribution to knowledge about the manuscript.

Database Field Name: Alphabet

Required: No

Contents: Indication of primary writing system of the manuscript.

Database Field Name: Non-Western Date

Required: No

Contents: Non-Western date format, if known. Examples: S= Seleucid Era, known in the Syriac tradition as the “Year of the Greeks”, beginning October 1, 311 BCE; H=Year of the Hijra, beginning July 16, 622 CE; A=Year of Adam (also known as the Byzantine World Era), beginning September 1, 5509 BCE.

Worksheet: Manuscript Text

Database Field Name: Folio Span of Text

Required: Yes

Contents: In reference to the individual text, the number of leaves or pages, e.g. ff. 1-25 or pp. 1-25. If foliation is supplied by catalogr, use square brackets, e.g. ff. [1-25].

Database Field Name: Author

Required: No

Contents: Name(s) of author(s). Give the most common form of the author's name; omit titles as part of the name (for example, patriarch, St.). Leave blank for anonymous authors.

Database Field Name: Author (native script)

Required: No

Contents: Name(s) of author(s) in native script.

Database Field Name: Other Associated Name

Required: No

Contents: Name(s) of other person(s) associated with a given text. Use for translators, editors, etc. Include their role in parenthesis.

Database Field Name: Generic Title

Required: Yes

Contents: Descriptive title of work given by catalogr. Also, common title of work, etc. Includes generic titles such as "Bible"; "antiphonal"; "psalter." In English.

Database Field Name: Generic Title (native script)

Required: No

Contents: Descriptive title of work given by catalogr in native script.

Database Field Name: Condition Of Text

Required: No

Contents: Brief description and reason for any missing text. Choices include: incomplete, defective, fragment, copied out of order, mutilated, misbound, leaves missing.

Database Field Name: Rubric (native script)

Required: No

Contents: Author- and/or title-like information in the manuscript, often set off from the main body of the text by red ink, underlining, display script or spacing. Use native alphabet.

Database Field Name: Incipit (native script)

Required: No

Contents: Beginning words of a text, often used to identify it. Entered in native script

Database Field Name: Explicit (native script)

Required: No

Contents: Ending words of a text, often used to identify it. Entered in native script.

Database Field Name: Language(s)

Required: Yes

Contents: Names, in English, of the language(s) used in the text. Separate languages with semicolon and space

Database Field Name: Further Info. On Text

Required: No

Contents: A non-defined slot to be adapted to the needs of a specific text. Field may include bibliographic information about the text (as opposed to the entire manuscript); authorial misattribution within the rubric; any other observations about the text itself.

Database Field Name: Music

Required: No

Contents: Type of musical notation, when present. Brief description if manuscript contains some musical notation. Leave blank for a manuscript that might be expected to consist entirely of music.

Database Field Name: Subject

Required: Yes

Contents: Keywords, taken from an authority file, that describe in general terms the contents or the genre of the manuscript. Separate keywords with semicolon and space.

Examples: Bible – OT; Gospels

Missal

Medicine; Veterinary

Unknown