Tagging of Item Records for

Medical Heritage Library Digitization

Julie Adamo & Kristen Greenland

January 28, 2011

Project Leader: Michael North

Table of Contents

[Abstract 3](#_Toc468094214)

[Introduction 3](#_Toc468094215)

[Methodology 4](#_Toc468094216)

[Results 5](#_Toc468094217)

[Discussion 6](#_Toc468094218)

[Recommendations 7](#_Toc468094219)

[Appendices 8](#_Toc468094220)

[Appendix A: Survey for Bound Materials Needing Special Handling for Digitization 8](#_Toc468094221)

[Appendix B: Scanning Instruction Codes 9](#_Toc468094222)

[Appendix C: Coding Voyager Records for Digitization Procedures 11](#_Toc468094223)

[Appendix D: Original Project Proposal 16](#_Toc468094224)

# Abstract

**Objective:** The purpose of this project was to aid in the digitization of the National Library of Medicine’s early American books for the Medical Heritage Library by tagging item records with scanning instructions.

**Methods:** Early American books dating up to 1820 were evaluated by conservators to identify characteristics relevant to scanning. These characteristics were then interpreted to determine the appropriate method of scanning. Scanning instructions were added to each item’s records in the Voyager ILS system.

**Results:** Scanning instruction annotations were added to records for 615 items.

**Conclusions:** Tagging these item records with scanning instructions will allow digitization to proceed as efficiently as possible, and will help in providing free online access to a portion of the National Library of Medicine’s collection.

## Introduction

 The Medical Heritage Library (MHL) is a collaborative digital curation project currently underway among a group of leading medical libraries, including the National Library of Medicine (NLM). Residing at the Internet Archive until a custom site is complete, the MHL provides “free and open access to quality historical resources in medicine.” As part of its contribution to the MHL, NLM’s History of Medicine Division (HMD) is digitizing over 1,000 early American books dating up to 1820. To digitize the books, NLM is using a Kirtas scanning machine (Kirtas Technologies, Inc., Victor, NY), an automated scanner capable of handling fragile materials and producing high resolution images with OCR (Optical Character Recognition).

 All books slated for inclusion in the MHL must have their bibliographic and item records tagged with a searchable string in NLM’s Voyager integrated library system. This tagging allows books to be pulled for scanning while also providing a means for keeping statistics on the project. As books have a variety of special conditions for scanning, particularly pertaining to size, boundwith (more than one item bound together), and foldouts, the tagging permits easy grouping of items. Additionally, books that have been reviewed and cannot be scanned need to have their records tagged so they are not selected for scanning in the future. The purpose of the present project was to insert either scanning instructions or “do not scan” notes into bibliographic and item records.

## Methodology

Prior to our involvement in the project, NLM conservators reviewed the items being considered for digitization to determine their condition and need for special handling. Information about the binding/case, size, sewing, textblock, pagination, presence of foldouts and previous repairs was noted on the “Survey for Bound Materials Needing Special Handling for Digitization” form (see Appendix A). The conservator also made a recommendation about the type of scanning appropriate and whether conservation was needed before the item could be scanned. We received batches of these surveys organized by call number and determined the scanning instruction code that matched the survey data. Important considerations were whether the item was under or over the size limits of the scanner, contained foldouts which would require hand scanning, or was otherwise too fragile to scan as-is. A complete list of the scanning instruction codes used can be found in Appendix B. Once we determined the appropriate scanning instructions, we noted them on the survey and could then proceed to tagging the item’s records in Voyager (version 7.1.0, ExLibris, Des Plaines, IL).

In order to tag a record, we followed the procedures outlined in the “Coding Voyager Records for Digitization Procedures” document provided to us by Laura Hartman in HMD (See Appendix C). The first step in the procedure was to look up the bibliographic record in Voyager using the call number or NLM barcode number noted on the survey. We then used the 245 title field to search for the item in Internet Archive (<http://www.archive.org/details/texts>) to determine whether the item had already been scanned at another institution. The Internet Archive is a non-profit organization that offers permanent access to digital historical collections online, and it does not make sense to re-scan items that are already freely available in a high quality format (we ignored scans performed by Google because of the low quality of the images). If an item had been scanned previously, we removed the unnecessary project specific fields from the record (583, 947, 996, and 998) and added a new 583 field with the designation “will not digitize.” If an item had not been scanned previously, we first changed the existing “needs condition review” 583 field to indicate that the review had been completed. We then marked the item for either conservation or scanning based on our prior assessment of the survey data. All bibliographic records that were edited also required an update to their 995|d date field to ensure that the altered record was redistributed to OCLC. After completing the bibliographic record update, we then opened the item record and added the scanning instructions to the notes field.

## Results

We updated 615 records total and completed the tagging project with the help of Krista Stracka, a student worker in HMD. Approximately 8% of the items had been previously scanned by other institutions (does not include items scanned by Google) and were in Internet Archive. Previously scanned items were therefore relatively rare, and a large portion of these seemed to be for specific digitization projects (a large number of items written by one person and all scanned by the same institution). Successful completion of NLM’s portion of the Medical Heritage Library digitization project is now possible and will allow for free access to these unique historical materials.

## Discussion

Now that all of the item records have been tagged, scanning of the materials can proceed as efficiently as possible, with items scanned in batches that have the same scanning instructions. For example, staff in HMD can query the Voyager database to identify all the items that require hand scanning and contain foldouts and then scan those items in a batch. Further tagging of records for items that need conservation will be required because the scanning instructions for these items were not assigned.

There were a few issues that arose during the course of this project and we discovered several conditions that were not readily represented on the surveys. Some items were not labeled as needing conservation, but had notes conveying that conservation would be ideal. While these items were generally interpreted as conservation necessary items, it would be helpful to standardize a process for dealing with them. Additionally, multivolume and boundwith items were not explicitly labeled on the surveys, and scanning instructions therefore had to be interpreted. Through call numbers and/or the way that pagination was recorded, it was relatively easy to identify these items, but explicit labeling would make the process more transparent and further reduce the potential for error. We also found during the course of the project that our tagging instructions did not include information on how to deal with items that were photocopies and would consequently not be scanned. Because this was a new project and the workflow had not been previously tested, there were changes throughout the process and we therefore identified several items that were exceptions to the rules of how to interpret the survey data. The survey form itself change a few times as well, as additional important data fields were added. These issues were minor, and we were able to resolve them and adjust our methodology by consulting with Laura Hartman and Krista Stracka.

This project provided an excellent introduction to using the Voyager ILS, as we had the opportunity to retrieve and update bibliographic and item records. It also helped enhance our knowledge of MARC cataloging fields and the data they contain. Working on a project with the History of Medicine Division was rewarding because we were able to get to know staff members and be exposed to part of the NLM’s historical collections.

## Recommendations

The project was largely straightforward, and with all survey data entered, there is little to recommend in terms of next steps or revisions. Items needing conservation will need to have their records tagged with scanning instructions once conservation is complete. Having established a solid workflow for this digitization project will help facilitate future projects of a similar nature and help them to go more smoothly. It would be helpful for future projects to add checkboxes (or other means) to the surveys for keeping track of when bibliographic and item records have been tagged (this was partially addressed in the final version of the survey form shown in Appendix A). This part of the survey is currently added manually and it would save time to have this step streamlined. Similarly, adding specific sections to mark if items are boundwith or multivolume would eliminate the need for interpretation and standardize the process. When we began the project, a macro existed that allowed for quick searching by barcode number. However, most of the surveys that we entered did not contain barcode numbers and we instead searched by call number. An additional macro for call number searching would save even more time. Overall, once detailed instructions were provided by Laura Hartman, we completed the project without any major roadblocks or difficulties. Scanning for remaining items for the MHL can now proceed.

## Appendices

### Appendix A: Survey for Bound Materials Needing Special Handling for Digitization



### Appendix B: Scanning Instruction Codes

Scanning instructions: Hand scan

Scanning instructions: Hand scan Bdwith

Scanning instructions: Hand scan Foldout

Scanning instructions: Hand scan Foldout Bdwith

Scanning instructions: Hand scan Foldout Multivol

Scanning instructions: Hand scan Multivol

Scanning instructions: Hand scan Overtall

Scanning instructions: Hand scan Overtall Bdwith

Scanning instructions: Hand scan Overtall Multivol

Scanning instructions: Hand scan Overthick

Scanning instructions: Hand scan Overthick Bdwith

Scanning instructions: Hand scan Overthick Multivol

Scanning instructions: Hand scan OVFoldout

Scanning instructions: Hand scan OVFoldout Bdwith

Scanning instructions: Hand scan OVFoldout Multivol

Scanning instructions: Hand scan Undersized

Scanning instructions: Hand scan Undersized Bdwith

Scanning instructions: Hand scan Undersized Foldout

Scanning instructions: Hand scan Undersized Foldout Multivol

Scanning instructions: Hand scan Undersized Multivol

Scanning instructions: HMD scan

Scanning instructions: HMD scan Bdwith

Scanning instructions: HMD scan Flatbed

Scanning instructions: HMD scan Foldout

Scanning instructions: HMD scan Foldout Bdwith

Scanning instructions: HMD scan Foldout Multivol

Scanning instructions: HMD scan Multivol

Scanning instructions: HMD scan Overtall

Scanning instructions: HMD scan Overtall Bdwith

Scanning instructions: HMD scan Overtall Multivol

Scanning instructions: HMD scan Overthick

Scanning instructions: HMD scan Overthick Bdwith

Scanning instructions: HMD scan Overthick Multivol

Scanning instructions: HMD scan OVFoldout

Scanning instructions: HMD scan OVFoldout Bdwith

Scanning instructions: HMD scan OVFoldout Multivol

Scanning instructions: HMD scan Undersized

Scanning instructions: HMD scan Undersized Bdwith

Scanning instructions: HMD scan Undersized Foldout

Scanning instructions: HMD scan Undersized Foldout Multivol

Scanning instructions: HMD scan Undersized Multivol

Scanning instructions: No scan

Scanning instructions: No scan Film scan

Scanning instructions: No scan MHL partner scanned

Scanning instructions: Scan as is

Scanning instructions: Scan as is Bdwith

Scanning instructions: Scan as is Multivol

### Appendix C: Coding Voyager Records for Digitization Procedures

**Coding Voyager Records for Digitization Procedures**

HMDMIA2 books (Americana printed before 1821; all classed in WZ 270)

Associate Project (Fall 2010)

Compiled by Laura Hartman, RBEM, Nov. 22, 2010

1. **Pull up item and/or bibliographic record in Voyager Cataloging Module** that corresponds to the condition report in hand
	1. Log into Voyager Cataloging Module using the separately supplied operator id and password codes
	2. Pull up bibliographic record in Voyager Cataloging Module using identifying information on condition report[[1]](#footnote-1) (barcode number, call number, or bib id)
		1. Prefer to search by barcode number, if given
			1. Use macro F5 to pull up Retrieve an Item Record box
			2. Type in barcode in the Enter barcode box, including the NLM prefix
			3. Do not input a space between NLM and the numeric portion of the barcode
			4. Click Retrieve, to pull up the associated item record
			5. Click Get Bib icon to retrieve the associated bibliographic record
			6. Proceed to step 2.
		2. If the barcode is not given on the condition report, search by the Voyager Bib Id
			1. Use macro F6 to pull up the Retrieve a Record (Bibliographic) box
			2. Type the bib id into the Enter the record ID box.
			3. Click Retrieve, to pull up the bibliographic record
			4. Proceed to step 2.
		3. If Bib Id is not given on the condition report, search by call number
			1. Click on Search icon
			2. Click on Index Selection Tab
			3. Click on Call Number in drop-down window
			4. Enter call number in Search For box, omitting enumeration (volume information)
			5. Click on Do Search box
			6. If Titles Index list appears, double-click on the correct call number to open bibliographic record
			7. Proceed to Step 2.
2. **Search item in Internet Archive using bibliographic information from Voyager**
	1. Open Internet Archive search Texts webpage (<http://www.archive.org/details/texts>)
	2. Copy and paste at least the first 6 words of the title in the Voyager bibliographic record (MARC tag field 245 $a (subfield a)) into the Search box
	3. Click orange Go! Button or just hit enter to perform search
3. Ignoring items marked “digitized by Google” in results list, **determine if scanned book is in Internet Archive**
	1. Click on each hyperlinked title to view fuller bibliographic information
	2. Compare publisher and year information in Internet Archive, with Voyager bibliographic information.
		1. Publisher and year information in Voyager is located in MARC field 260 subfields b and c
	3. If match is found, click on title page image of book
		1. Confirm that title page has been scanned
		2. If so, verify that the title, publisher and year on the scanned title page matches the Internet Archive/Voyager bibliographic data
4. **Annotate condition report IA yes or IA no**
	1. If item is not found in Internet Archive or is found, but the title page is missing, annotate report “IA no”
	2. Otherwise, annotate condition report “IA yes”
5. **Update Voyager bibliographic record**
	1. If condition report is marked “IA yes”
		1. Delete the unnecessary fields from bibliographic record
			1. Highlight fields to be deleted, by clicking on the grey box to the left of each 583, 947, 996, and 998 field
				1. Push delete key
				2. Click yes, when asked if you are sure
			2. Use macro Ctrl + Alt + =, to add new 583 field beginning “will not digitize”
		2. Mark the bibliographic record for redistribution
			1. Place cursor before the date in the 995 $d field
			2. Use macro Ctrl + Alt + G to change the date to “today’s date”
		3. Save the record to the database by clicking on the Save to DB (boat) icon.
			1. Click on “no” if prompted to change “profile”
			2. Click on “continue” not “close” when the Voyager authority validation screen appears.
		4. Close the bibliographic record by clicking on the bottom X in the upper right-hand corner.
		5. If you initially opened this bibliographic record by searching on the barcode, close the item record by clicking on the bottom X in the upper right-hand corner.
		6. Proceed to the next condition report.
	2. If condition report is marked “IA no”
		1. Delete existing 583 field beginning “Needs condition review”, following instructions in step 5.1.1.1
		2. Use macro Ctrl + Alt + Shift + ] to input new 583 beginning “Condition reviewed”
		3. Determine if conservation work is necessary
			1. Condition report may be marked “will conserve”
			2. Box “conservation necessary prior to digitization” will be checked on older condition reports
		4. If conservation work is necessary,
			1. Use macro Ctrl + Alt + Shift + . to input additional 583 field beginning “Will conserve”
			2. Follow steps 5.1.2 through 5.1.6 to mark the record for redistribution, and save and close the record.
		5. If conservation work is not necessary,
			1. Place cursor immediately after the last digit of the date in the 583 field beginning “condition reviewed”
			2. Use macro Ctrl + Alt + Shift + / to insert $l “conservation not necessary” between the text in subfields c and 2
			3. Follow steps 5.1.2 and 5.1.3 above to mark the bibliographic record for redistribution and save the changes to the database.
			4. Do not close the bib record yet
		6. Determine appropriate Scanning Instruction phrase
			1. Determine if item is Scan As Is
				1. Determine if item is multivolume or boundwith
			2. Determine if item is Hand Scan
				1. Determine if item is Undersized, Overthick, etc.
				2. Determine if item has foldouts
				3. Determine if item is multivolume of boundwith
			3. Determine if item is HMD Scan
				1. Determine if item is Undersized, Overthick, etc.
				2. Determine if item has foldouts
				3. Determine if item is multivolume of boundwith
				4. Determine if Flatbed scanner is necessary
6. **Update Voyager item record**
	1. Open item record
		1. If you retrieved the bibliographic record using the item barcode, simply close the bibliographic record (see step 5.1.4)
		2. If you retrieved the bibliographic record using the bib id or call number, click on the get items icon and select the holdings record for location HMD
	2. Open item record note field, by clicking on Quill icon (third from left, at bottom). Icon will have a red X if the note field is empty; or it will have a yellow scroll if there is already text in the note field.
	3. Input appropriate Scanning Instruction phrase into Item record note field (use macro Ctrl + Alt + Keypad 9 or copy and paste from earlier records); if note field already contains text, enter scanning instructions on next line, for clarity
	4. Click on Save to DB (boat) icon to save item record to database.
	5. Close item record by clicking on bottom X in the upper right hand corner.
	6. Close bibliographic record by clicking on bottom X in upper right hand corner.

**Commonly used Voyager shortcut keys:**

**F3 Inserts field above**

**F4 Inserts field below**

**F5 Retrieve by barcode**

**F6 Retrieve by bib id**

**Common macros:**

**Ctrl + Alt + G Inputs today’s date**

**Ctrl + Alt + = Input 583 beginning Will not digitize**

**Ctrl + Alt + Shift + = Input 583 beginning Will conserve**

**Ctrl + Alt + Shift + ] Input 583 beginning Condition reviewed**

**Ctrl + Alt + Shift + / Input $l conservation not necessary**

**Ctrl + Alt + Keypad 9 Retrieve, select, and input appropriate Scanning Instructions phrase**

### Appendix D: Original Project Proposal

Project Title: Tagging Item Records of Early Americana Books for Digitization

Submitted by: Michael North

Description: Over 1,000 early American books dating up to 1820 are ready for scanning for the Medical Heritage Library, however we must tag their bibliographic and item records using the Voyager Cataloging module with a searchable string. This will enable us to pull scannable books and keep statistics on them. Books which cannot be scanned also need to have their records tagged, so that we know not to select them for scanning in the future. Conservation and cataloging reviews have already been performed for most items, and tagging for or against scanning will be based upon our conservation review forms. A macro has been created using MacroExpress so that keyboarding of the tags is not required.

Expected amount of time: about 60 hours.

Expected outputs: Bibliographic and item records tagged for early Americana books to be scanned for the Medical Heritage Library.

Experience gained by Associate Fellow: Experience with Voyager Cataloging module; experience with workflows to prepare materials for digitization.

Project leader: Michael North

1. The content and information supplied on the condition report has evolved over time. More recent forms contain the NLM barcode; older forms may only contain the call number and/or bib id for the item. [↑](#footnote-ref-1)