R:\DigitalCollections\documentation\efacs_pageturner\Version 4\PageturnerMasterDoc_v4.doc

Pageturner Model Documentation

UW Madison Libraries' Local Usage Guide and Interpretations [1]
Version 4.0

Authored by Kirstin Dougan, Amy Rudersdorf, and Jessica Williams

---

[1] *This document is based on the definitions and specifications detailed in the UW Libraries' Digital Library Data Dictionary for Electronic Facsimile Collections (http://www.library.wisc.edu:4000/dept/ltg/DigiLib/EFacs/EFacsDataDictionary.html), and attempts to broaden the narrative in language that is more accessible. It is intended for use in conjunction with the data dictionary, not to replace it. This document outlines what should be placed in each metadata field and how the data should be formatted.*

Log of Revisions

4/1/2005
Title pages will be automatically transcribed for books that cannot be OCR'd. (See *What is OCR?,* Section IV, page 9.)

6/16/2004
This version (4.0) of the Pageturner documentation marks a significant change to every existing level (or "object") within the data model, as well as the addition of the new level ("object"), "Subcollection." Please pay close attention to the layout of each level (note new fields and deleted fields) as well as instructions that have been updated to reflect the changes.

11/7/2004
Author titles - - In the Issue_Author, Issue_Editor, and Item_Author field, only include titles and abbreviations of titles of nobility, address, honour, and distinction, initials of societies, qualifications, etc. if such data is necessary grammatically, the omission would leave only a person's given name or surname, the title is necessary to identify a person, the title is a title of nobility, or is a British term of honour (AACR2r, 1-15). **See update to documentation on p. 8.**

Multiple Authors - - Multiple authors in the Issue_Author, Issue_Editor, and Item_Author fields will no longer be delimited by a semi-colon. Rather, use the pipe (i.e. |) when delimiting multiple authors. **See update to documentation on p. 8.**

I. Introduction

 The Pageturner model is used for materials with sequentially structured contents (for example, a monograph, serial, or report) and where the structure must be reflected in the online environment in order to provide a natural interface for use. With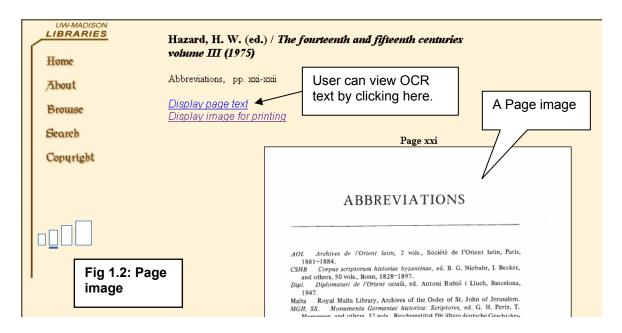 the Pageturner model, the Resource is presented as a series of Items (such as chapters or articles) and may also be viewed as a series of Pages. This allows the user to move through the Resource page by page, as well as navigate through the higher-level structure (such as chapters or articles).

Fig 1.1: Pageturner model

As preparation for using the Pageturner model, the Source is imaged (scanned) and may also be run through an optical character recognition (OCR) program. In the interface, raw text from the OCR process is "hidden" and used for keyword searching. However, the user views an image of the actual Page. Alternatively, the user may choose to view the resulting text from the OCR process instead, clicking on the "Display page text" option in the interface (not shown in the example above).

Fig 1.2: Page image

II. Definitions

Throughout this document "**Source**" refers to the original object being described. "**Resource**" refers to a set of information, which was encoded from the Source and is now being converted to electronic form.

What is Metadata?

Metadata can be defined as "data about data," or any information associated with or about a particular Resource. Examples of metadata include the author of a book, the title of an article, or the number of pages in a letter. In this document, metadata refers to *structural, descriptive, and administrative information about Resources that make up our digitized collections*. These Resources may be 'analog' items (journals, books, etc.) that have been turned into a digital collection, or may be items that were 'born digital' (email, MSWord documents, etc.) and that only exist online.

Metadata is created for and associated with the digital Resource to support its cataloging, discovery, use, storage, and migration. It is most often divided into three conceptual types (there is some overlap between the three).

> **Descriptive metadata:** used for the indexing, discovery, and identification of a digital Resource. Examples of descriptive metadata include title, author, publisher, and physical format.

> **Structural Metadata:** information used to display and navigate through digital Resources; also includes information about the internal organization of the digital Resource. Structural metadata indicates structural divisions of a Resource (i.e., chapters in a book) or sub-relationships (such as distinct parts of a letter; e.g., salutation, body, closing).

> **Administrative metadata:** represents the management information for the object, and includes information the user needs to access and display the Resource, as well as rights management and long-term preservation and archiving information. Administrative metadata includes the

resolution an image was scanned at, the hardware and software used to produce an image, compression information, pixel dimensions, etc.

III. General Principles

1.) The default level of capture is cover-to-cover, including blank pages, and tipped-in objects, but not necessarily things like tissue overlays.
2.) When entering a **title** at the Collection, Aggregate, Issue, or Item level:
   a. At the aggregate level, if a uniform title is present, enter the uniform title, which will directly correspond with the Machine Readable Cataloging Record (MARC)[2] 130 field. A uniform title is the title used for cataloging purposes when a work has appeared under more than one title (such as translations into several languages), or when the work being cataloging is of a collective nature such as "Complete Works." If there is no uniform title or if the item is not cataloged, enter the title exactly as it appears until you reach the first period or logical breaking point. At the issue level, enter the issue title that directly corresponds with the title in the MARC 245 field. At the item level, enter the title for the item exactly as it appears on the item title page until you reach the first period or logical breaking point.
   b. Transcribe the title from the beginning of the title to the first period (include the period). If there is no period, use common sense to discern where the title ends and the rest of the descriptive information begins. Do **not** include the period at the end of the title.
   c. Follow capitalization rules as outlined in AACR2. In English this means capitalizing only the first word of a title and any proper nouns.
   d. Do not add brackets to inserted punctuation.
   e. Do not change the spelling of a word that is misspelled, uses archaic spelling, or is spelled in an unfamiliar way. If a word is misspelled, enter the misspelled word as is followed by the correct spelling of the word in brackets. (i.e. the elixer [elixir] of life)
   f. Currently there is no way to support styles such as bold or italics in the text. However, most basic and some expanded Latin characters are supported. For example, "ë," "ă," " ▪ ," and "Lj," are supported. Additionally, the use of a few common characters are restricted, since they are used in the mark-up language (SGML) we use both to store and distribute our data. These characters include "&," "<," and ">." To enter any of these characters, first, refer to the chart of diacritics at http://www.ramsch.org/martin/uni/fmi-hp/iso8859-1.html.[3] Using the "description" and "char" column, identify the character you would like to input. Then, determine the corresponding number in the "code" column. To enter the number, hold down the "alt" key on your keyboard, type "0" in replace of the # sign in the character code, and then type in the numbers for the character code on the number keypad on the right of the keyboard. (If you try to use the numbers at the top of your keyboard, the character codes will not work.) It is important to note that the character code will always be preceded by a "0." (i.e. for the ampersand, hold down the "alt" key, then key in the code, "0, 3, 8" on the numbers keypad).
   g. If it is necessary to use ellipses, first turn off the Microsoft ellipses function on your computer. Microsoft adds specific features to certain characters in order to condense space. Basically, when a user types ellipses, Microsoft changes the ellipses from three

---

[2] "Machine-readable" means that one particular type of machine, a computer, can read and interpret the data in the cataloging record. "Cataloging record" means a bibliographic record, or the information traditionally shown on a catalog card.
[3] Also see http://www.bbsinc.com/iso8859.html.

keys to one key. The code used to condense the ellipses from three keys to one key will not export from Excel into Site Search. Therefore, if you are entering the metadata while at the DCG, this will not be an issue as this function is already turned off on all DCG computers. To turn off the Microsoft ellipses function at your own PC:

   a. Open Excel
   b. From the Tools menu, select Autocorrect
   c. Select the row that displays the ellipses (i.e. ...)
   d. Press the delete key.
   e. This will change the ellipses function for all Microsoft Office products.

 h. Avoid line breaks or hard returns.
 i. Please see Appendix A for examples.

3.) When entering **author** information at the Aggregate, Issue, or Item level:

   a. When attributing author at the Aggregate, Issue, or Item level, always use the personal name entry that corresponds with the Machine Readable Cataloging Record (MARC)[4] 100 field. (This is the "Author field in your local, online catalog.) The personal name is the name of the person chiefly responsible for the creation of the artistic or intellectual content of an item. If there is no personal name ("100" or "Author" field), check to see if there is a corporate name (MARC 110 field [this information should also be listed in a field titled "Author" or "Corporate Author" or an equivalent]). A corporate name is the name of an agency, association, business, firm, government, institution, nonprofit enterprise, performing group, etc. who is responsible for the creation of the artistic or intellectual content of an item. If there is not a MARC record for the Aggregate, Issue, or Item level, refer to the title page of the original object to ascertain whether there is an author. If there is no author, leave the author field in the Pageturner model empty.

   b. If there are multiple, unique values to be entered in a field, they should be delimited with a pipe and a space, "; ". E.g., Multiple authors: `Smith, Joe| Brown, Tom`

   c. If there are 3 or less authors, enter all three authors. If there are 4 or more authors, enter the first author followed by et al. See page 21 for more detailed instructions and examples.

   d. Do **not** include the period after the date range that follows the authors name according to the LC Name Authority file.

   e. Only include titles and abbreviations of titles of nobility, address, honour, and distinction, initials of societies, qualifications, etc. if such data is necessary grammatically, the omission would leave only a person's given name or surname, the title is necessary to identify a person, the title is a title of nobility, or is a British term of honour (AACR2r, 1-15).

IV. OCR

What is OCR?

Page images for electronic facsimile Collections may be scanned through an OCR program. OCR (Optical Character Recognition) is the method for the machine-reading of typeset and typed letters, numbers, and symbols using optical sensing (usually a scanner) and a computer. A computer program analyzes the patterns and identifies the characters they represent, with some tolerance for less than

---

[4] "Machine-readable" means that one particular type of machine, a computer, can read and interpret the data in the cataloging record. "Cataloging record" means a bibliographic record, or the information traditionally shown on a catalog card.

perfect/uniform text. OCR is also used to produce text files from computer files that contain images of alphanumeric characters, such as books, journals, and typewritten letters. The clearer the original text page, the better the image-scan, and thus the more accurate the OCR'd text. Even the cleanest scan will not result in perfectly OCR'd text, however. In some cases, the text will need to be corrected manually. The title page is automatically transcribed for all books that cannot be OCR'd.

V. Metadata

To enter the metadata, you will need to have either digitized Page images or the actual Source(s) on hand, and the pageturner spreadsheet or database template.

The Template

Adapting the Access Database Template

Adapting the template is an individual decision, and is best avoided whenever possible. However, it may be necessary according to your needs. Some general things to consider:

- DO NOT ADD TABLES OR FIELDS TO THE DATABASE TEMPLATE. Many of the scripts and programs used to manipulate the final data were created to work with the tables exactly as they appear in the database.
- DO NOT CHANGE FIELD NAMES. See above. Field names should have each word capitalized (i.e., *Item_Author, Item_Page_Sequence_No_List*). Do not change this.
- Entering **default values**. A field's default value can be automatically entered into that field every time a new record is created. Examples of default values include Collection name and Issue IDs. You can change the default values after each batch if necessary. However, be sure that the default value is correct or empty before you begin data entry, or all subsequent records will have the wrong values in that field.
- To **autonumber** or not to autonumber? You may find that this feature is more problematic than it is worth, but you may choose to use it for fields such as the sequence numbers. Only one autonumber field is allowed in an Access table.

Adapting the Excel Spreadsheet Template

- DO NOT ADD TABLES OR FIELDS TO THE WORKSHEET TEMPLATE. Many of the scripts and programs used to manipulate the final data were created to work with the worksheets exactly as they appear in the spreadsheet. There are a few things you can do to alter the way in which data is input.
- DO NOT CHANGE FIELD NAMES. See above. Field names should have each word capitalized (i.e., *Item_Author, Item_Page_Sequence_No_List*). Do not change this.
- DO NOT ENTER ANYTHING IN THE "UNUSED" COLUMNS OR ROWS OF THE WORKSHEETS.

Levels

There are six main levels into which you will enter the metadata for your Collection. If using MSAccess for metadata entry, these levels will be represented by individual tables. If using MSExcel, these levels will be represented by individual worksheets.

1.) Collection Level

This level describes your digital Collection as a whole. One of our digital Collections is called *Illustrated Shakespeare*. It consists of fifteen monographs all related to the topic of illustrations of Shakespeare's works. One of the titles is *The spirit of the plays of Shakespeare*. This title is a smaller part of the whole *Illustrated Shakespeare* Collection.

2.) Subcollection Level

This level groups together subject related issues and aggregates within a collection in order to allow cross searching. **Collections may or may not have subcollections. Subcollections are assigned by project owners or content providers.**

3.) Aggregate Level

A logical level of organization lower than Collection and Subcollection but higher than that of the individual Issue (see below). For most serials, the Aggregate level will contain information about the volumes. In a multi-volume set, such as *The spirit of the plays of Shakespeare* (which consists of five volumes), the set is an Aggregate. The Aggregate may be either a physical or a purely intellectual division. **NB: Not all Collections will have Aggregates. For (single-volume) monographs, the Aggregate level will not be used.**

4.) Issue Level

This is described as the basic unit of distribution. The Issue refers to a single-volume monograph, an individual journal Issue, or to a folder of unbound—but related—documents. For multipart-monographs and some serials the Issue may correspond to a single volume of the multi-volume set.

5.) Item Level

This is the only unit of organization recognized within an Issue. Items are articles, chapters, letters (in the folder example above), or similar designations.

6.) Page Level

A single Page within a Resource. Occasionally this may be a single image of two facing Pages.

Step-by-Step Metadata Entry

1. Collection

The Collection table or worksheet will contain a single entry describing the entire Collection.

The fields and how to fill them

*Required fields are in **bold**, some fields are "required when applicable" and are marked in **bold** and with an asterisk (*).*

Field	Description	Notes	Repeatable?
Collection_ID	Identifier for Collection; unique within scope of all Digital Collections	Assigned by DCG; name will be used to create Collection environment on server e.g., "IllusShake"	NO
Collection_Title	Long form of Collection title	Assigned by Collection owner; include the entire title e.g., "Illustrated Shakespeare"	NO
Collection_Title_NFC	Number of non-filing characters at start of Collection_Title	Positive Integer that tells the script not to index stopwords or articles. Include the space after the article e.g., "a =2, an =3, the =4, der =4, das = 4, le =3, la =3," etc.	NO
Collection_Availability	Information about copyright, access rights, etc.	Copyright statement for *entire Collection* e.g., Copyright © 2004 Board of Regents of the University of Wisconsin System	NO

*Required fields are in **bold**, some fields are "required when applicable" and are marked in **bold** and with an asterisk (*).*

NB: If a field is marked "Repeatable," multiple values can be placed in that field. Each value must be separated by a semicolon and a space, "; ". This is true in all tables.

Examples

Fig 3.1: Collection

	A	B	C
1	Collection_ID	Collection_Title	Collection_Title_NFC
2	UWAlumni	University of Wisconsin Alumni Collection	

Cont'd

Fig 3.1a: Collection

	D
1	Collection_Availability
2	Copyright 2002 Board of Regents of the University of Wisconsin System

2. *Subcollection*

The level between Collection and Aggregate, Subcollection, is loosely defined as an arbitrary grouping of elements in a collection. More specifically, the Subcollection level serves as a way to group Issues and/or Aggregates of a collection that are of similar thematic content. In turn, this grouping allows for controlled searching across related subject matter.

Subcollection Table			
Field	**Description**	**Notes**	**Repeatable?**
Collection_ID		Carried over from Collection table	NO
Subcoll_ID	Identifier for Subcollection; unique within scope of all Digital Collections	Assigned by DCG; name will be used to create Collection environment on server e.g., "GerRecon"	NO
Subcoll_Title	Long form of Collection title	Assigned by DCG owner; include the entire title e.g., "Germany under Reconstruction"	NO
Subcoll_Title_NFC		Carried over from Collection table	NO

Example

Fig 3.2: Subcollection table

3. *Aggregate*

The level between Collection and Issue, Aggregate, may be an intellectual or physical subdivision. For serials, the Aggregate usually corresponds to the overall work that the combined volumes represent. An example of the Aggregate is Catesby's *Natural History*, which has two volumes. While each volume is considered an Issue, the work as a whole is the Aggregate, as it contains the individual volumes. Not all Collections will have Aggregates. For example, conference proceedings would not have an Aggregate. There would only be Issues (each year of proceedings) and Items (papers within the proceedings), unless one year of proceedings were issued in several physical volumes.
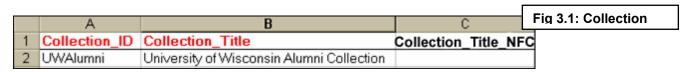
The fields and how to fill them

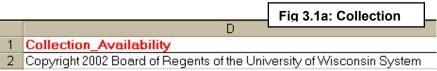*Required fields are in **bold**, some fields are "required when applicable" and are marked in **bold** and with an asterisk (*).*

Aggregate Table			
Field	**Description**	**Notes**	**Repeatable?**
Collection_ID		Carried over from Collection table	NO
Aggregate_Sequence_No	Sequence number of the Aggregate	Unbroken sequence of 4-digit #s beginning with 0001 for the first Aggregate in the Collection and continuing in unbroken sequence through	NO

		the last	
Aggregate_ID*	Unique (within Collection) identifier for Aggregate	Once assigned should not be changed; usually a textual identifier, e.g., `JCEV23`	NO
Aggregate_Author*	Author of Aggregate	Use form of name from LC authority file if extant; `LN, FN, DOB-DOD`	YES
Aggregate_Editor	Editor of Aggregate	Use form of name from LC Authority File if extant; `Last Name, First Name, Date of Birth-Date of Death`	YES
Aggregate_Title*	Title of Aggregate	Follow AACR2 capitalization	YES
Aggregate_Title_NFC		Carried over from Collection table	NO
Aggregate_Title_Level*	Type of Title of the Aggregate; see notes below	`m`[onographic] `j`[ournal] `s`[eries] `u`[npublished]	NO
Aggregate_Issue_Sequence_No_List	Range of *issue_sequence_no's* included within this aggregate	e.g., if this is a five (5) volume set, the entry here will read `0001-0005`	NO

*Required fields are in **bold**, some fields are "required when applicable" and are marked in **bold** and with an asterisk (*).*

When entering *Aggregate_Title*

1.) At the aggregate level, if a uniform title is present, enter the uniform title in the Aggregate_Title field, which will directly correspond with the MARC 130 field. A uniform title is the title used for cataloging purposes when a work has appeared under more than one title (such as translations into several languages), or when the work being cataloging is of a collective nature such as "Complete Works." If there is no uniform title or if the item is not cataloged, enter the title exactly as it appears until you reach the first period or logical breaking point. Do **not** include the period.

Aggregate_Title_Level:

- ▪ *m[onographic]*=Aggregate is multipart monograph (monograph with more than one volume)
- ▪ *j[ournal]*=journal/magazine/serial publications (e.g., Aggregate is a journal volume)
- ▪ *s[eries]*=Aggregate contains volumes in a monographic series
- ▪ *u[npublished]*=unpublished material (letters, manuscripts, etc.)

Fig 3.3: Aggregate table

Examples

	A	B	C	D	E
1	Collection_ID	Aggregate_Sequence_No	Aggregate_ID	Aggregate_Author	Aggregate_Editor
2	History	0001	Hanotaux	Hanotaux, Gabriel, 1853-1944	
3	History	0002	Verdenskrigen		
4	History	0003	Fragments	Bairnsfather, Bruce. 1887-	
5	History	0004	HistCrusades		Setton, Kenneth M.
6	History	0005	VisitingSeries		
7					

Cont'd

<div style="text-align: right;">

Fig 3.3a: Aggregate table

</div>

	F	G	H	I
1	Aggregate_Title	Aggregate_Title_NFC	Aggregate_Title_Level	Aggregate_Issue_Sequence_No_List
2	Histoire illustrée de la guerre de 1914		m	0001-0017
3	Verdenskrigen i Billeder.		j	0001-0014
4	Fragments from France.		j	0001-0006
5	A History of the Crusades	2	m	0001-0006
6	Visiting expert series		s	0001-0018
7				

4. Issue

Describes an individual journal issue, a monograph, or a folder (or other collection of unbound—but intellectually related—materials).

The fields and how to fill them

*Required fields are in **bold**, some fields are "required when applicable" and are marked in **bold** and with an asterisk (*).*

Issue Table			
Fields	**Description**	**Notes**	**Repeatable?**
Collection_ID		Carried over from Collection table	NO
Aggregate_ID*		Carried over from Aggregate table; if used	NO
Subcoll_ID*		Carried over from Subcollection table; if used	NO
Issue_Sequence_No	Sequencer for Issue; unique within scope of Aggregate	**Used even if there is no aggregate**; unbroken sequence of 4-digit #s beginning with `0001`; **all un-aggregated issues will be seq. # `0001`—so there may be more than one `0001` in a Collection**	NO
Issue_ID	Identifier for issue; unique within scope of Collection	Assigned by DCG; once assigned should not be changed; can be any unique combination of letters and numbers e.g., `WT1930`	NO
Issue_Std_No	Standard number or identifier such as ISSN or ISBN	Associated with Issue; include type and value, separated by a semicolon e.g., "`ISBN; 0-674-79002-2`"	YES

*Issue_Printed_No**	**Includes** Sequential Issue (in some cases, Volume or Part) number as printed on source's title page or cover **AND** Type of Issue enumerated by *Issue_Printed_No*	`e.g. Volume 43; Volume 43, Number 2; Volume 43, Number 2, Section 1`	NO
*Issue_Author**	Author of Issue	Use form of name from LC Authority File if extant; `Last Name, First Name, Date of Birth-Date of Death`	YES
Issue_Editor	Editor of Issue	Use form of name from LC Authority File if extant; `Last Name, First Name, Date of Birth-Date of Death`	YES
Issue_Submitter	Person submitting Issue for inclusion in the Collection	Follows LCNAF. Last name, First name: Institution. Department (E.g., `Laudati, Geri: University of Wisconsin--Madison. Libraries. Mills Music Library`); for non-UW Madison projects, the default is to omit personal names	YES
*Issue_Title**	Title of Issue	Follow AACR2 capitalization; **don't include information such as "Part I," that is captured elsewhere**	YES
*Issue_Title_NFC**		Carried over from Collection table	NO
Issue_Title_Level	Type of Title of the Issue; see notes below	`m`[onographic] `j`[ournal] `a`[nalytic] `u`[npublished]	NO
*Issue_PubPlace**	Place of publication	Does not need to follow AACR2 conventions. Not repeatable, so use first city listed.	NO
*Issue_Publisher**	Publisher of Issue		NO
Issue_Chron	Date of publication of source Issue as printed in source or period of time represented by source Issue	e.g., `March 1932`	NO
Issue_Extent	The physical characteristics of the Issue.	Equivalent to MARC300; can include things like the # of numbered pages, e.g., `296p.`	NO
Issue_Page_Sequence_No_List	Range of *Page_Sequence_No's* included within this issue	4-digit sequence numbers separated by hyphen, e.g., `0001-0211`	NO

Issue_Text	Is OCRd text available for this Issue	`y` or `n`	NO
Issue_Abstract	A textual summary of the content and significance of the Issue.	No line breaks or markup may be included in the value.	NO
Issue_Availability	Information about copyright, access rights, etc.	**e.g.,** `Copyright 2002 Board of Regents of the University of Wisconsin System`	YES

Required fields are in **bold***, some fields are "required when applicable" and are marked in* **bold** *and with an asterisk (*).*

When entering *Issue_Title*

1.) At the issue level, enter the issue title that directly corresponds with the title in the (Machine Readable Cataloging Record (MARC) record (245 field).

Issue_Title_Level:
- *m[onographic]*=for monographic titles 1) when there is no Aggregate, or the *Aggregate_Title_Level* is "s," 2) when the *Aggregate_Title_Level* is "m," and the *Issue_Title* is substantially the same as the *Aggregate_Title*
- *j[ournal]*=journal/magazine/serial title substantially identical to the *Aggregate_Title*
- *a[nalytic]*=for analytic titles when the *Aggregate_Title_Level* is "j" or"m," and the Issue is to be cataloged separately (such as a special issue of a journal bearing its own title)
- *u[npublished]*=unpublished material; letters, manuscripts, etc. (may be a supplied title)

Examples

Fig 3.4: Issue table

	Collection_ID	Aggregate_ID	Subcoll_ID	Issue_Sequence_No	Issue_ID	Issue_Std_No	Issue_Printed_No
62	History	HistCrusades		0001	CrusOne		volume, I
63	History	HistCrusades		0002	CrusTwo		volume, II
64	History	HistCrusades		0003	CrusThree	ISBN; 0-299-06670-3	volume, III
65	History	HistCrusades		0004	CrusFour	ISBN; 0-299-06820-X	volume, IV

Cont'd

Fig 3.4a: Issue table

	Issue_Author	Issue_Editor	Issue_Submitter
80	Kuenzli, Irvin R., 1897-		Hill, Vicki: University of Wisconsin--Madison. Libraries
81	Bird, Dillard E.		Hill, Vicki: University of Wisconsin--Madison. Libraries
82	Meyer, Henry Joseph, 1913-		Hill, Vicki: University of Wisconsin--Madison. Libraries
83	Southall, Sara		Hill, Vicki: University of Wisconsin--Madison. Libraries

Cont'd

Fig 3.4b: Issue table

	K	L	M	N	O	P	Q
1	**Issue_Title**	**Issue_Title_NFC**	**Issue_Title_Level**	**Issue_PubPlace**	**Issue_Publisher**	**Issue_Chron**	**Issue_Extent**
2	Histoire illustrée de la Guerre de 1914		m	Paris	Gounouilhou	1915	
3	Histoire illustrée de la Guerre de 1914		m	Paris	Gounouilhou	1915	
4	Histoire illustrée de la Guerre de 1914		m	Paris	Gounouilhou	1915	
5	Histoire illustrée de la Guerre de 1914		m	Paris	Gounouilhou	1915	

Cont'd

Fig 3.3c: Issue table

	R	S	T	U
1	**Issue_Page_Sequence_No_List**	**Issue_Text**	**Issue_Abstract**	**Issue_Availability**
2	0001-0311	Y		Copyright 2003 Board of Regents of the University of Wisconsin System
3	0001-0290	Y		Copyright 2003 Board of Regents of the University of Wisconsin System
4	0001-0314	Y		Copyright 2003 Board of Regents of the University of Wisconsin System
5	0001-0314	Y		Copyright 2003 Board of Regents of the University of Wisconsin System

5. Item

Describes an individual chapter, section, letter, etc. within an Issue.

The fields and how to fill them

*Required fields are in **bold**, some fields are "required when applicable" and are marked in **bold** and with an asterisk (*).*

Item Table			
Field	**Description**	**Notes**	**Repeatable?**
Collection_ID		Carried over from Collection table	NO
Issue_ID		Carried over from Issue table	NO
Item_ID*	Identifier for Item; unique within scope of Issue identified by Issue_ID.	An Item is assigned a unique identifier if the Item_Type is "article" or "work" if that "article" or work" has one or more authors.	NO
Item_Sequence_No	Identifier for Item; unique within scope of Issue	Unbroken sequence of 4-digit #s beginning with `0001`	NO
Item_Std_No	Standard number or identifier such as ISSN or ISBN	Include type and value separated with a semicolon e.g., "`ISBN; 0-674-79002-2`"	YES
Item_Type	Type of Item	Must come from TEI list (see below)	NO
Item_Author	Author of Item	Use form of name from LC authority file if extant; `Last Name, First Name, Date of Birth-Date of Death.` Use this if Items within Issues have separate authors (e.g., a journal issue with separate authors for	YES

		each article, or a book in which each chapter has a separate author) **Do not use if all *Item_Authors* are the same in all items or at the Issue level.**	
Item_Title	Title of Item	Transcribe ***up to and including first period.*** Follow AACR2 capitalization. See notes below.	YES
*Item_Title_NFC**		Carried over from Collection table	NO
Item_Abstract	A textual summary of the content and significance of the Item.	No line breaks or markup may be included in the value.	NO
Item_First_Printed_Page_No	Page number printed on first Page of this Item	If written in roman numerals, enter roman numerals This can include a plate # (e.g., `Plate VI`)	NO
Item_Page_Sequence_No_List	Range of *Page_Sequence_No's* included within this Item	4-digit sequence numbers separated by hyphen, e.g., `0001–0035`; used even if there is only one page, e.g., `0036-0036`	NO

*Required fields are in **bold**, some fields are "required when applicable" and are marked in **bold** and with an asterisk (*).*

Item types to be used in the Item_Type field

Item Type Table		
`Section [default]`	`Acknowledgements`	`Notes`
`Frontispiece`	`Errata`	`Index`
`Contents`	`Chapter`	`Appendix`
`Masthead`	`Article`	`Glossary`
`Foreword`	`Editorial`	`Bibliography`
`Preface`	`Work`	`Colophon`
`Dedication`	`Act`	`Cover`
`Abstract`	`Scene`	`Title page` (*added from*
`Introduction`	`Letter`	*UW Data Dictionary*)

For definitions and examples of these terms see Appendix B.

When entering *Item_Title*

1.) Only the text that actually appears on/in the Source should be captured.
2.) Transcribe the title from the beginning of the title to the first period (do not include the period). If there is no period, use common sense to discern where the title ends and the rest of the descriptive information begins.
3.) Be sure to check for subtitles. For instance, if the title appears on the book as
   "Part 3
    Afternoon Invigorations
    A nice way to spend the day"
Transcribe the title as follows:
    Part 3: Afternoon invigorations: a nice way to spend the day
4.) Follow capitalization rules as outlined in AACR2. In English this means capitalizing only the first word of a title and any proper nouns.
5.) Do not change the spelling of a word that is misspelled, uses archaic spelling, or is spelled in an unfamiliar way. If a word is misspelled, enter the misspelled word as is followed by the correct spelling of the word in brackets. (i.e. the elixer [elixir] of life).

6.) Do not add brackets to inserted punctuation, such as colons for subtitles.

7.) If there is no Item_Title on the printed page, you may leave that field blank if the Item_Type is one of the following Item_Types:

```
Cover
Introduction
Foreword
Contents
Masthead
Frontispiece
```

8.) If an Item_Title is available on the printed page, enter the title into the Item_Title field. Include in brackets preceding the title the following Item_Types:

```
Cover
Introduction
Foreword
Title Page
Contents
Masthead
Frontispiece
```

Example:

       Item type = Introduction
       Item title = To the alumni

       Then, type into the fields:
       Item_Type = Introduction
       Item_Title = [Introduction] To the alumni

9.) Any words or phrases that do not actually appear on the original object must be enclosed in square brackets "[ ]." E.g., `[Cover]`.

10.) For the Item_Types listed below, include the bracketed name of the Item type and the title or caption listed on the Item (if there is one) in the Item_Title field. (**Note** - - the word, "Cover," in brackets will suffice for both front and back covers.)

     i.   [Cover] Wisconsin alumnus
     ii.   [Half-title] Wisconsin alumnus
     iii.   [Title page] Wisconsin alumnus
     iv.   [Frontispiece] Campus of the future
     v.   [Masthead]

11.) For Item_ID, use the combination of author's last name and an original word from the title.
`(e.g., Title = All the king's men; Author = J. Robinson; Item_ID = RobinsonKing.)`
You may also use numbers to distinguish between similar Item_IDs within a single Issue. `(e.g., RobinsonKing1, RobinsonKing2, etc.)`

Blank and Marbled Pages

When considering the ***Item_Page_Sequence_No_List***:

1.) If there are blank pages in between Items (sections), they should be considered part of the preceding section. Example:

    Section 1: Page 1
              Page 2
              Page 3 [blank page]

Section 2:   Page 4
             Page 5

2.) Also, if there is more than one item on a page (e.g, the first half of the page is titled "Foreword" and the second half of the page is titled "Introduction"), the ***Item_Page_Sequence_No_List*** will reflect this by repeating that page sequence number. Example:

Item_Sequence_No	*Item_Type*	*Item_Page_Sequence_No_List*
0001	Foreword	0001-0002
0002	Introduction	0002-0003
0003	Section	0004-0007

3.) Marbled pages get their own section. Example:

Item_Sequence_No	*Item_Type*	*Item_Page_Sequence_No_List*
0001	Front Cover	0001-0001
0002	Section*	0002-0004
0003	Title page	0005-0006

* [Marbled pages] would go in the *Item_Title* and this section would include any blank pages that followed the marbled pages.

Examples

Fig 3.5: Item table

	A	B	C	D	E	F	G
1	Collection_ID	Issue_ID	Item_ID	Item_Sequence_No	Item_Std_No	Item_Type	Item_Author
1659	History	Konnte		0016		Cover	
1660	History	Symposium		0001		Cover	
1661	History	Symposium		0002		Title Page	
1662	History	Symposium		0003		Title page	
1663	History	Symposium		0004		Preface	Hiller, Kurt
1664	History	Symposium		0005		Contents	

Cont'd

Fig3.5a: Item table

	H	I	J	K	L
1	Item_Title	Item_Title_NFC	Item_Abstract	Item_First_Printed_Page_No	Item_Page_Sequence_No_List
1659					0146-0146
1660	[Cover] After nazism -- democracy?				0001-0004
1661	[Half title] After nazism -- democracy?				0005-0006
1662	[Title page] After nazism--democracy? A symposium by four Germans				0007-0008
1663	Editor's preface			5	0009-0010

There should be no numbering breaks here, although page sequence numbers may be repeated to indicate more than one item on a page.

6. Page

Describes the individual Page image. One entry per Page.

The fields and how to fill them

Required fields are in **bold***, some fields are "required when applicable" and are marked in* **bold** *and with an asterisk (*).*

Page Table			
Field	**Description**	**Notes**	**Repeatable?**
Collection_ID		Carried over from Collection table	NO
Issue_ID		Carried over from Issue table	NO
Page_Sequence_No	Identifier for Page; unique within scope of Issue	Unbroken sequence of 4-digit #s beginning with `0001`	NO
Page_Description	Textual description of Page content	Subject headings/terms that will be provided by the project owner	YES
Page_Printed_No*	Sequential Page number as printed on source page	There may be none; if written in roman numerals, enter roman numerals. **This can include a plate# such as "`Plate VI`."****	NO
Page_Text	ASCII or Unicode encoding of text on Page	Whatever is in this field will override the presence of an OCR .txt file.	NO
Page_Location	System subpath (within Collection) to image file	The filepath will always begin with "EFacs" and end with a "/" e.g., `EFacs/ArtsSoc01n02/`	NO
Page_Filename	Base filename for image of this Page	Alphanumeric; **do not include file extension**	NO
Page_Format	Type of file served	Valid MIME type	NO
Page_Notes	Additional information about the source, which might impact scanning quality. Also used to indicate context for pages without printed page numbers.	If Page has no printed number include 2-3 words of text from page here.	NO

Required fields are in **bold***, some fields are "required when applicable" and are marked in* **bold** *and with an asterisk (*).*

**Note: If the page number printed on the page is incorrect, enter the *Page_printed_no* as follows:

- ▪ `Incorrect number that appears on the page [i.e. correct page number]`

    Example: `330 [i.e. 223]` *Note that there is no comma after i.e. as one would expect (AACR2 1.4B6 or 1.4F1)*

Unless the project owner specifies and provides a guide to page numbering, do not add page numbers if none are printed on the page.

Examples

	A	B	C	D	E	F
1	Collection_ID	Issue_ID	Page_Sequence_No	Page_Printed_No	Page_Description	Page_Text
17610	History	Vertriebene	0027	25		
17611	History	Vertriebene	0028	26		
17612	History	Vertriebene	0029	27		
17613	History	Vertriebene	0030	28		
17614	History	Vertriebene	0031	29		
17615	History	Vertriebene	0032	30		
17616	History	Vertriebene	0033	31		
17617	History	Vertriebene	0034	32		

Cont'd

Fig 3.5a: Page table

G	H	I	J
Page_Location	Page_Filename	Page_Format	Page_Notes
EFacs/GerRecon/Vertriebene/	0027	image/tiff	
EFacs/GerRecon/Vertriebene/	0028	image/tiff	
EFacs/GerRecon/Vertriebene/	0029	image/tiff	
EFacs/GerRecon/Vertriebene/	0030	image/tiff	
EFacs/GerRecon/Vertriebene/	0031	image/tiff	
EFacs/GerRecon/Vertriebene/	0032	image/tiff	rotate
EFacs/GerRecon/Vertriebene/	0033	image/tiff	rotate

VI. Metadata Quality Control

Once all of the structural metadata has been entered, quality control should be performed on the data. This is an important step in the metadata entry process as it will ensure that data and its accompanying digital images match up and properly replicate the book structure online.

Before giving the Imaging Unit access to the metadata, and before exporting data, check the following:

Overall:
1. For each field check documentation for input guidelines (such as capitalization and name order) and make sure they are being met.
2. Check for typos or obvious misspellings, especially in ID fields, etc.
3. Sequence numbers should always be an **unbroken sequence** of four-digit, zero-padded numbers, without gaps or duplications.
4. If there are multiple authors, treat them in the following way:
   a. For up to three authors, write out each author's name in the appropriate author field (aggregate, issue, item). (i.e., Wilkens, John H.| Cooper, Heidi M.| Clinton, Hillary B.)
   b. For four or more authors, only write out the first author followed by a comma and et al. (i.e., Wilkens, John H., et al.) Do NOT put a pipe after the first author or the "et al." will appear on a separate line online, which is incorrect.
5. If an author's last name is followed by "Jr.," enter the author's name in reverse order followed by a comma and "Jr." (i.e., Wilkens, John H, Jr.)

At the Collection Level:
1. The following required fields are completed:
   a. *Collection_ID*,

       b. *Collection_Title* (which should only have first word and proper nouns capitalized)
       c. *Collection_Title_NFC*
       d. *Collection_Availability*.
2. Spelling and grammar match original object.

At the Aggregate Level (if used):

3. The following required fields are completed:
       a. *Collection_ID*,
       b. *Aggregate_Sequence _No*,
       c. *Aggregate_ID*, and
       d. *Aggregate_Issue_Sequence_No_List*.
4. If applicable, the following fields should be completed:
       a. *Aggregate_Author*
       b. *Aggregate_Title*
       c. *Aggregate_Title_NFC*
       d. *Aggregate_Title_Level*
5. *Aggregate_Issue_Sequence_No_List* matches the *Issue_Sequence_No*(s) for that Aggregate in the Issue table.
6. Spelling and grammar match original object.
7. Names are pulled from LC Name Authority File and follow `LN, FN, DOB-DOD` (if dates are available).

At the Issue Level:

8. The following required fields are completed:
       a. *Collection_ID*,
       b. *Issue_Sequence_No* (any issues without an aggregate get an *Issue_Sequence_No* of `0001`),
       c. *Issue_ID*, and
       d. *Issue_Page_Sequence_No_List* (make sure this corresponds to the last entry in the *Item_Page_Sequence_No_List* for this Issue at the Item level).
       e. *Issue_Text*
       f. *Issue_Availability*
9. Spelling and grammar match original object.
10. If applicable, the following fields should be completed:
       a. *Aggregate_ID*
       b. *Issue_Printed_No*
       c. *Issue_Author*
       d. *Issue_Title*
       e. *Issue_Title_NFC*
       f. *Issue_Title_Level*
       g. *Issue_PubPlace*
       h. *Issue_Publisher*
11. Names are pulled from LC Name Authority File and follow `LN, FN, DOB-DOD` (if dates are available). Do **not** include the period after the `LN, DOB` or `DOD`

At the Item Level:

12. The following required fields are completed:
       a. *Collection_ID*,

    b. *Issue_ID*,

    c. *Item_Sequence_No*,

    d. *Item_Type*, and

    e. *Item_Page_Sequence_No_List*. The last item in this list should correspond to the range for this issue in the *Issue_Page_Sequence_No_List* at the Issue level.

13. Spelling and grammar match original object.

14. If Item_Type is Cover, Title Page, or Frontispiece, the Item Title should be formatted as such:

```
[Cover] Wisconsin alumnus
[Half-title] Wisconsin alumnus
[Title page] Wisconsin alumnus
[Frontispiece] Campus of the future
[Cover]
```
Note - - [Cover] will suffice for both describing front and back covers. If there is nothing on the back cover, just write "[Cover]."

15.   Names are pulled from LC Authority File and follow `LN`, `FN`, `DOB-DOD` (if dates are available). Do not include the period after `LN`, `DOB`, or `DOD`.

<u>In the Page Table:</u>

16. The following required fields are completed:

    a. *Collection_ID*,

    b. *Issue_ID*,

    c. *Page_Sequence_No*,

    d. *Page_Location*, (make sure this has a '/' after it)

    e. *Page_Filename*, and

    f. *Page_Format*.

    g. Ensure that *Page_Sequence_No* numbers in the Page table correctly correspond with the *Item_Page Sequence_No_List* in the Item table.

17. Spelling and grammar match original object.

18. As many other fields as possible are completed.

19. Names are pulled from LC Name Authority File.

Appendix A: Capitalization Examples based on AACR2

Basic title
- *On item:* The Materials of Architecture
- *AACR2:* The materials of architecture
- *On item:* Still Life with Bottle and Grapes
- *AACR2:* Still life with bottle and grapes

Title with proper nouns
- *On item:* The 1919/1920 Breasted Expedition to the Far East
- *AACR2:* The 1919/1920 Breasted Expedition to the Far East

Title with alternate title
- *On item:* The Edinburgh World Atlas, or, Advanced Atlas of Modern Geography
- *AACR2:* The Edinburgh world atlas, or, Advanced atlas of modern geography
- *On item:* A Dictionary of American English on Historical Principles
- *AACR2:* A dictionary of American English on historical principles

Title with subtitle and proper nouns
- *On item:* The Greenwood Tree: Newsletter of the Somerset and Dorset Family History Society
- *AACR2:* The greenwood tree: newsletter of the Somerset and Dorset Family History Society
- *On item:* Quo Vadis?: A Narrative from the Time of Nero
- *AACR2:* Quo vadis?: a narrative from the time of Nero

Title within a title
- *On item:* Selections from the Idylls of the King
- *AACR2:* Selections from the Idylls of the king

Multiple work titles in title and proper nouns
- *On item*: King Henry the Eighth; and, The Tempest
- *AACR2:* King Henry the Eighth; and, The tempest

Various foreign language examples
- Les misérables (basic title)
- Les cahiers du cinema (basic title)
- Coppélia, ou, La fille aux yeux d'émail (title with alternate title)
- Strassenkarte der Schweiz = Carte routière de la Suisse = Carta stradale della Svizzera = Road map of Switzerland (Title with alternate titles and proper nouns)
- Sechs Partiten für Flöte (German title)

In German, all nouns are capitalized. In French, capitalization is similar to English.

Appendix B: Item Type Definitions

Abstract—a brief, objective summarization of the essential content of a **work**, presenting the main points in the same order as the original. In a journal article, the abstract follows the title and name(s) of the author(s), and comes before the text of the article. Different from a **summary,** which comes at the end of the work.

Acknowledgements—the section in the front of a work in which the author recognizes the scholarly or work-related contributions of others. In contemporary monographs, this usually follows the preface or foreword and comes before the introduction. Not the same as the **dedication,** which is a shorter, less formal note written by the author addressing the book to someone in order to honor or memorialize them.

Act—one of the major divisions in the action of a play, marked onstage by the dropping of the curtain and an intermission. Acts are usually made up of scenes.

Appendix—appears at the back of the work, after the main text but before the notes, glossary, bibliography, and index. Appendices contain supplementary material such as statistical tables. Not the same as **notes,** which give the source of a quotation or idea.

Article—a nonfiction, prose composition that is published under its own title in a collection or periodical containing other works of that same form. May be written by a single author or several.

Bibliography—a list of references to sources cited in the text of an article or book. Usually appears at the end of the work, after any appendices but before the **index**.

Chapter—one of several major divisions of a book. Each chapter is complete in itself but relates to those before and after it. May be given a title or simply numbered. In modern publications, chapters are created and specified by the author, and will appear by name or number in the table of contents. See **section.**

Colophon—a statement given at the end of a work, or sometimes on the back of the title page, giving detailed information about the printing of the work. This may include the name of the printer, typeface, grade of paper, and the type of binding used.

Contents—a list of all of the divisions, chapters, articles, or individual works contained in a publication. Often listed near the front of the publication as a "table of contents," in order of appearance and with page numbers.

Cover—the outer protective covering of a book, periodical or manuscript.

Dedication—a brief note in the front of a work, written by the author and addressing the work to one or more people in order to honor or memorialize them. In modern publications, the dedication is usually printed on the right-hand page following the title page. Compare to **acknowledgment.**

Editorial—a short essay expressing the opinion or position of the chief editors of a newspaper or magazine. Usually addresses a current political, social, or professional issue.

Errata—errors discovered after a publication has gone to press and thus not able to be corrected in the main text. They are printed as a separate list and are either glued in after the binding has been done or printed in a later volume of a multi-volume set.

Foreword—introductory remarks preceding the text of a work, usually written by someone other than the author of the work. Generally comes after the dedication and before the introduction, if there is one.

Frontispiece—an illustration immediately preceding the title page or first page of a book. The frontispiece is not given a page number.

Glossary—an alphabetical list of the specialized terms used in a particular **work**, giving short definitions. Usually appears at the end of the **work**, after any notes but before the **bibliography** and **index**.

Half-title—the title of a book as printed on the right-hand page which precedes the full title page. It originated as a blank page that was included to protect the title page, and is usually printed in a smaller font. [NB: this is not an *Item_Type*, but can be used as part of the *Item_Title* if the *Item_Type* has been specified as "Title Page."]

Index—an alphabetically-arranged list of headings consisting of the important names, places, and subjects mentioned in a written work, and referring the reader to page numbers where these may be found. The index is usually the last thing in a book. In a multi-volume set, the index may be a separate volume. Usually only nonfiction works are indexed.

Introduction—the part of a book stating the purpose and subject of the work. Usually written by the author. Normally follows the **preface** or **foreword**.

Letter—a handwritten, typewritten, or printed personal or business message. Usually enclosed in an envelope and physically delivered to the recipient.

Masthead—in a periodical, a box or column printed in each issue stating the title of the publication, its publisher, owner and editors, frequency, ISSN, subscription rates, copyright information, and contact information. In magazines, the masthead is usually found on or near the page with the table of contents; in newspapers, it is generally on the editorial page or the front page.

Notes—a statement explaining something in the text of a work, or giving the source of a quotation or idea that is not the author's own. Notes are usually numbered and may be listed as footnotes on the bottom of the applicable page; at the end of a chapter; or at the end of the work as a whole.

Preface—a preliminary statement at the beginning of a book, usually *written by the author*, stating the origin, scope, purpose, etc., of the work. Compare to **foreword,** which is usually written by someone other than the author. In modern publications, either the preface or the **foreword** will appear after the **dedication** but before the **introduction**.

Scene—in a play, one of the subdivisions of an act. Each scene presents continuous action in one place.

Section—an intellectual division of a work, similar to a chapter but used when the author has not created chapters. This should be used when no other type definition can be logically applied to the data.

Title page—the page at the beginning of a publication that gives the official title of the **work**, and usually the name(s) of the author(s), editor(s), etc. Usually appears following the **half-title**. Volume number, date and place of publication may also be indicated.

Work—an expression of human thought or feeling in language, symbols or images, offered for purposes of communication and record. Not the physical document, but the intellectual content behind it.

Appendix C: Where Does the Information Go?

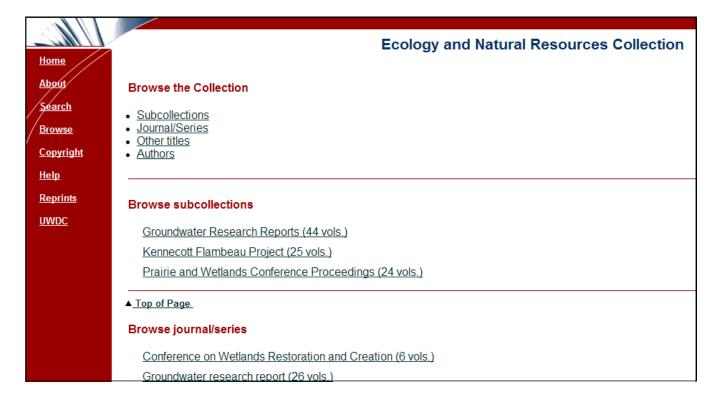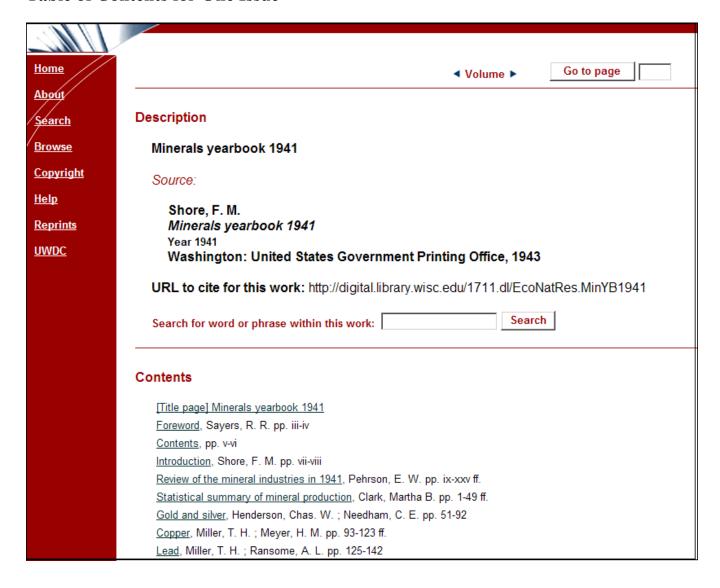**Browse Screen Listing Issues in a Collection**

Fig 3.6: Browse Page

Fig 3.6a: Browse Page

Table of Contents for One Issue

Home
About
Search
Browse
Copyright
Help
Reprints
UWDC

◄ Volume ►   Go to page

Description

Minerals yearbook 1941

Source:

Shore, F. M.
Minerals yearbook 1941
Year 1941
Washington: United States Government Printing Office, 1943

URL to cite for this work: http://digital.library.wisc.edu/1711.dl/EcoNatRes.MinYB1941

Search for word or phrase within this work: **Search**

Contents

[Title page] Minerals yearbook 1941
Foreword, Sayers, R. R. pp. iii-iv
Contents, pp. v-vi
Introduction, Shore, F. M. pp. vii-viii
Review of the mineral industries in 1941, Pehrson, E. W. pp. ix-xxv ff.
Statistical summary of mineral production, Clark, Martha B. pp. 1-49 ff.
Gold and silver, Henderson, Chas. W. ; Needham, C. E. pp. 51-92
Copper, Miller, T. H. ; Meyer, H. M. pp. 93-123 ff.
Lead, Miller, T. H. ; Ransome, A. L. pp. 125-142

Page View for One Page in an Item