EMu Bibliography Module Rework Proposal

Presented by Smithsonian NMNH Informatics

Background

Many users have been frustrated with the EMu Bibliography module as implemented for a very long time. NMNH has mocked up a reworked version of Bibliography. We focused on designing a tool based on defining what our needs were, consulting library and archives staff for current best practices, and building upwards, rather than focusing on current limits and constraints of EMu.

Questions posed to our users: What do you want Bibliography to do?
- What are the pieces of information you need to get your job done and meet the needs of your department?
- In what ways do you need to extract or report on this information? Do you have a need to extract it in specific formats?
- In what ways would you like to be able to search this information?

The responses from NMNH users included focus on time and ease of getting data in and out of Bibliography, both hand entering and structuring imports and reports.

We met with our Libraries staff to review current bibliographical best practices and workflows. The Libraries' Research.si.edu portal contains a relatively simple bibliographic submission page, with a limited number of reference types and fields to correspond to them, in a relatively flat structure.

Based on user feedback and conversations with our Librarians, we are proposing to use the Bibliography module in a very different way. The new focus is designing bibliography to act only as a placeholder reference, so we do not need the same amount of detail as a dedicated bibliography resource. ***EMu would not be maintaining the authority record of the reference, but just needs enough data to act as a pointer and generate the necessary info to reach the authority record.***
We want Bibliography to be easier to use, so that users will be more willing to enter bibliographic data into EMu. In order to achieve that we are looking at alternative data models.

A Fresh Approach

- Moving from a nested to a flat model

    - Simplifies data entry, imports, exports, and reporting.

- Tab switching based on the most common types of publications:

    - Journal Article

    - Book

    - Chapter

    - Citations of any of the above

    - Miscellaneous (all other types)

- Changing the publication types to lookup lists instead of parent/child attachments; journal and article would be two fields in one record.

    - All bibliographic data can be recorded within one record, no need to create nesting parents and children.

    - Record is created for the lowest level of detail needed.

- Fields that align with RIS format fields for streamlined importing and exporting between EMu and external programs such as EndNote or Reference Manager.

New Structure

The tab switching would be controlled by the existing *BibRecordType* field, based on the values "Article", "Book", "Chapter", "Citation" and "Misc".

All fields in these new tabs would map to the same back end columns, even though the displayed labels are different. Tab switching and the registry can display and hide applicable tabs as needed. This flattens the data structure somewhat, and will make it easier to report for values across bibliographic types. For example, reporting on the BibTitle field would return Title data for books, articles, and any Misc type resources, instead of having to report on each of those fields individually as in the current module.

As much as possible, we want these common fields to line up. In the case of titles, the lack of hierarchy may cause confusion with the level of title, so users need to be consistent and aware when querying and reporting:

- Book title and Journal title map to the same location
- Chapter, Article, Resource title and Title (Miscellaneous tab) map to the same location
- In search mode, an Also Search function can be put in place to search across both backend title fields

The new tabs were created based on a list from a survey of NMNH Bibliography users on their important and desired fields:

> Article Title
> Journal Title, full and abbreviated
> Author(s)
> Editor(s)
> Publication Date
> Publication City
> Series
> Volume
> Number
> Range of pages
> Plates and Figures
> Publisher (books)
> DOI linked field
> URL linked field
> ZooBank literature LSID

Most of the fields in bibliography are common across the various reference types and could be combined into unified back end fields (*ArtAuthors_tab, JouAuthors_tab, BooAuthors_tab* combined into *BibAuthors_tab* for exmaple). The remaining reference type-specific fields could then be consolidated on a single tab, eliminating the hierarchical structure.

As mentioned previously, this model treats the data in EMu Bibliography primarily as a finding aid. DOIs serve as an industry standard for quickly navigating to the authority source for a particular bibliographic resource. EMu's GUID grid can store DOIs and other resolvable GUIDs, eliminating the need to maintain high level of bibliographic details in EMu itself.

Field Mapping: EMu to RIS

Below is a preliminary mapping of the redesigned EMu fields to RIS formatting. Note that this is just the first pass, and the mapping will be refined as the designs are finalized.

Label name	EMu Column	RIS field	Comments
Type	BibRecordType	TY	
Book Series	BibSeries	T3	
Journal	BibBookJournalTitle	T2, (JF, JO?)	
Article Title	BibTitle	TI	
Book Title	BibBookJournalTitle	T2, (BT?)	
Chapter Title	BibTitle	TI	
Title	BibTitle	TI	
Resource Title	BibTitle		
Reference Type	BibReferenceType	TY	Need to work how how to have the 5 types for the tab switching, as well as the Other subtypes.
Abbreviation	BibAbbreviation	J2	
Author	BibAuthors_tab	A1, A2, A3,	
Role	BibRole_tab	x	No distinction of roles in RIS, all people can be listed as Authors. Editors will usually be on A2 line, though there also a ED tag.
Volume	BibVolume	VL	
Pages	BibPages	SP, EP	Need two page fields? Start Page and End page separately?
Number	Bib	IS?	Is this Issue
Figures/Plates	BibguresPlates	L4	
Publication Date	BibPublicationDates	DA	
City	BibPublicationCity	CY	
Publisher	BibPublishedBy	PB	
ISBN	BibISBN	SN	
ISSN	BibISSN	SN	
Is Manuscript?	BibIsManuscript		
Source	BibSource	IN	
URL	BibWebIdentifier	UR	
GUIDs	AdmGUIDType_tab+AdmGUIDValue_tab	DOI	Will need a string with AdmGUIDType_tab=DOI + the GUID value.
Keywords	BibKeywords_tab	KW	
		DP	Database Provider

Mapping resources:

http://support.mendeley.com/customer/portal/articles/1006006-what-is-the-mapping-between-ris-files-and-mendeley-

https://en.wikipedia.org/wiki/RIS_%28file_format%29

http://www.risinc.com/support/risformat_tags_01.asp

http://refdb.sourceforge.net/manual-0.9.4/c2166.html

Article Tab

| Journal | | | | Type | Article |
| Article Title | | | | Abbr. | |

Authors/Contributors | Roles | | Manuscript ☐
| | | Et al. ☐

Volume		Number	
Pages		Figures/Plates	
Publication Date		Actual Date	

Keywords		Publisher	
		City	
		ISSN	

Article

A mockup of the redesigned Article tab. The Manuscript, and Keyword fields are new additions, and would also appear on most of the other tab types.

- Instead of linking to a separate Journal record, the appropriate journal can be selected or entered in a Journal lookup list.

- Manuscripts – The Manuscript checkbox can be used with the bibliographic resource is unpublished. This gives the option of flagging records so that Manuscript equals either "yes" or "no", which will make querying for or excluding Manuscripts easier. When or if a particular manuscript is published, the rest of the data can be entered, and the Manuscript checkbox is unchecked, converting the record into a normal Article record.

- Et al. checkbox – when checked would list the first author followed by the et al. in the summary data. When unchecked, all authors will be listed. Can also be checked in lieu of listing all authors. Removal of automatic et al. for three or more authors allows all the names to be accessible via summary line searches if the unit so desires.

- A new Keywords grid would now appear on all the reference type tabs.

Articles Tab Questions
- Does the Articles tab need to have page numbers? Many articles are retrieved digitally on their own without the rest of the journal content.

Book Tab

Book Title		Type	Book

Manuscript ☐

Et al. ☐

Authors/Contributors	Roles

Book Series

Volume		Number	

Publication Date		Actual Date	

City

Keywords

Publisher

ISBN

Book

Similar to the Articles tab above, but including Book specific fields.

- Book series is now just a lookup field. The *Series* name will be added upon the first recorded instance. Users wishing to find all books in the series would simply query on the Series field; to find all book series the user would search for any Book record where *Series* had a value. They could then determine existing series using a sort report. This and the elimination of the linking of Journals to separate Articles will eliminate the need for the References tab in Bibliography.

Chapter Tab

Chapter Title		Type	Chapter

Book Title

Manuscript ☐

Et al. ☐

Authors/Contributors	Roles

Volume		Number	

Pages		Figures/Plates	

Publication Date		Actual Date	

Keywords		City	
		Publisher	
		ISBN	

Chapter

Again similar to the Articles tab above, continuing the flat structure. The Chapters tab expands on the Books tab by providing further detail when a specific chapter of a book is to be referenced.

- Additional fields for Chapter Title, Pages and Plates

Citation Tab

Citation Type	*"Book" "Chapter" or "Article"* 🔖		Type	Citation 🔖

Resource Title	*Article or Chapter title*

Journal/Book	*Journal or book title*		Abbr.	🔖

Authors/Contributors	Roles	🔖	Manuscript	☐
			Et al.	☐

Volume		Number	

Pages		Figures/Plates	

Publication Date		Actual Date	

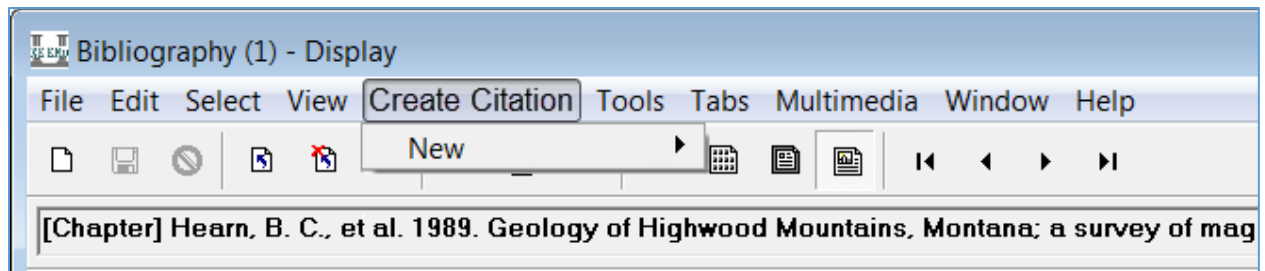| Keywords	🔖	Publisher	
		City	🔖
		ISSN/ISSB	

Citation

The Citations Tab would provide the lowest level of granularity for bibliographic references.

- Single Citations tab for all Bibliographic record types.

- The *Reference Type* field seen here will be the same one used by the Misc tab. This list will need to be hierarchal with Type, so that Type=Citation> Ref Type=Book/Article/Chapter, while Type=Misc> Ref Type = Thesis/newspaper/etc.

- Citations would not be required when linking references. If a unit or institution does not reference to the Citation level of detail, this tab could simply be hidden, and they would link Articles/Books etc. as their references.

Citations Menu

The plan for Citation is to make a menu item that creates a new record from a Book/Chapter/Article record (and miscellaneous) and dittos all the fields. First the user would query for the resource to cite. From the menu, there would be a "Create Citation" button. This would create a new Bibliography record with a *Type* of "Citation" (potentially opening in a new EMu window), and ditto most of the data from the previous record. This would reduce redundant data entry for the Citation, and prevent accidental modification of existing Article, Book or Chapter records if a user was trying to cite them.

We want to emphasize that this would *just* be a ditto. In order to keep the flat structure within Bibliography, there will be no parent-child linking between Article, Book and Chapter records and their Citations. Citations from a particular reference can be found using the Title field, and when querying users can choose to include or exclude citations using the *Type* field.

- Since most fields will be dittoed, only *Pages* and *Figures/Plates* would need to be updated individually.

- If Citation is from a book, *Resource Title* will be null

- Similar references can be found by searching *Title.* Searching for citations or non citations controlled by the *Type* field.

Miscellaneous Tab

Ref Type	Website/Newspaper/Magazine/Thesis etc		**Type**	Misc
Title			**Abbr.**	
Source	*Institution, Publisher or periodical title*			

Authors/Contributors	Roles

Et al. ☐

Publication Date [   ]

Keywords [icon]   **URL** [   ]

[   ]
[   ]

Miscellaneous

The Miscellaneous tab can be used as a catch-all for the less frequently used bibliography types, including website, newspaper, thesis, etc.

It has two title fields mapping to T1 and T2 of the RIS Format (As an example, Source/T2 for Newspaper/blog name and Title/T1 name of article, name of thesis, etc).
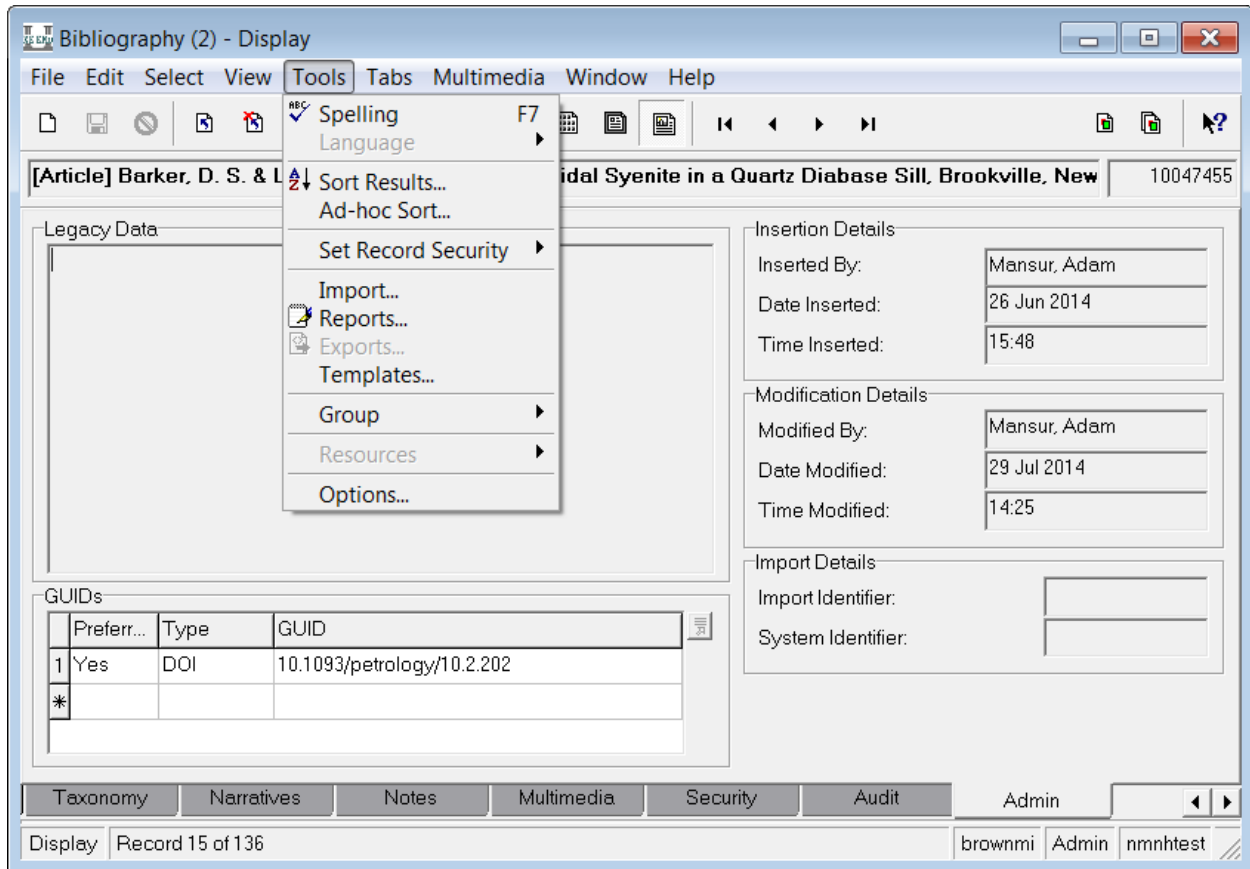
Questions for Miscellaneous Tab
For Misc. record, do we need volume, number, pages, figures? Institution (only for thesis)? Or Publisher/Institution instead of institution?

We need for the larger EMu community to decide standards for the Miscellaneous type and what it's used for, and the correct RIS mapping , particularly if we are trying to make the new Bibliography work for other institutions.

Special cases
The combination of Book, Chapter, Article, Citation and Miscellaneous tabs are designed to handle the majority of bibliographic references encountered by museums and institutions. In many cases, special exceptions can still be handled by the aforementioned tabs. For example, a thesis can be treated as a book, and an article in a newspaper can be treated as any other article, with title of the newspaper acting as the Journal.   Again, we are trying to look at this redesign from the perspective of "How much bibliography do we really need", especially if bibliography in EMu is to act more as a finding aid to the repository of record for the references, and not act as an authority in and of itself.

Existing Tab additions or modifications

DOIs and other GUIDS on Admin tab

DOIs and LSIDs can be stored in the GUID grid on the Admin tab, according to their type. We can write the necessary xml to enable the *Resources* button for each type of ID so that they can be linked directly to the resolver.

NMNH Informatics proposes creating an import tool based on DOI that would retrieve data from DOI link in RIS format, which could then be imported into EMu. This may be something that can be done using the Operations module, further investigation is needed.

Notes tab
Convert the text box notes tab to the Notes Grid found in Catalog and some other modules.

References Tab
With the flat structure, the need for the references tab can be eliminated.

Data Formatting Questions

Authors

How should we handle names?

- Do we want to continue using the Parties module to link names, or start entering names natively into a Bibliography text field or lookup list? Text is easiest to enter but we lose standardization, lookups could become unwieldy as they grow, and attachments are cumbersome to enter, import and export. *The SI librarians have recommended keeping Parties for the standardization it can bring.*

- Is there a standardized format for names that should be used for each/all departments? This will become a larger issue if Parties module is not used for name, we can mask names in Summary Data if linked to Parties, but if names are a local Lookup, names must be formatted specifically and correctly during data entry.

- Will having all Authors and Editors in a single grid cause issues with separating them out into the Authors RIS fields (AU, A1…); namely, does the A* RIS field tag all people associated with the reference regardless of role?

- Is there ever a need to search for Parties that are attached in Bibliography along with other modules?   Is it ok if users can't use Show attached in Parties to view places where that name occurs in Bibliography?

There was consensus at the 2015 EMu conference discussion group that there was greater utility in maintaining a link to Parties for names, for standardization and consistency.

Citations

- How important is citation level data? Does anyone stop at Article/Chapter?

- Will there ever be a need to cite non book/journal articles? How would one record a citation for a website?

- Are only plates/figures and page numbers needed for citations?

- Can manuscripts be cited, if so, should Citations tab have a checkbox?

These points were discussed at the 2015 EMu conference. Plate, figure, and page number fields can easily be hidden if not needed. For units or institutions that do not use citations, they would just not utilize that tab, creating a citation would not be required to create a bibliographic reference.

Summary Data

What would the summary data be for the reformatted module?

- Does format change based on publication type?

- Will using or not using Parties affect how names should be displayed or entered?

RIS and Formatting for other exports

- RIS fields can be used with BibTex markup

- Would like to see data structure for some major collaborates, such as AAA, to see if just the RIS fields will be sufficient.

RIS is used by: Endnote, RefWorks, Zetero, Mandalay, SRO/SIL