

## **The Deceased Preaches His Own Eulogy: *Training Students to Provide Access Points on Discovery Level Records***

Felicia Piscitelli<sup>1</sup>, Lisa Furubotten<sup>2</sup>, Anton duPlessis<sup>3</sup>, Alma Beatriz Rivera Aguilera<sup>4</sup>, Ángel Villaba Roldán<sup>5</sup>

### **Abstract**

This paper is an outgrowth of *Discovering a New World: Cataloging Old and Rare Imprints from Colonial and Early Independent Mexico*, a project funded by the Council on Library and Information Resources (CLIR). The grant succeeded in developing a methodology by which Hispanophone student workers, with minimal training, can create discovery level records for collections by inputting bibliographic metadata into a web-based template.

In addition to descriptive elements, access points for corporate and personal names and some type of subject access are needed to truly make this collection discoverable by users. Can students with rudimentary training provide such access points? If so, are they limited to inputting uncontrolled terms, or can they be trained to utilize controlled vocabularies? We experimented with four input methods to evaluate which ones effectively enable students to populate the access point elements:

- Uncontrolled vocabulary. For example, name headings entered in a surname(s), forename(s) format; keywords; summaries.
- Controlled vocabulary embedded in the metadata collection tool. For example, genre terms taken from the Rare Books and Manuscripts Section (RBMS) Controlled Vocabularies, and geographic subject headings taken from Library of Congress Subject Headings (LCSH).
- Dynamic feedback list. For example, a list of controlled name headings for printers from the National Authority File (NAF), with feedback whereby a cataloger updates the list
- Use of authorized thesaurus. For example, students captured authorized name headings from the Virtual International Authority File (VIAF).

In this paper, we discuss the pros and cons of each method. We also give consideration to the determination that, while providing lists of controlled headings might be useful for small and/or homogeneous collections, such a method proves unwieldy for a large heterogeneous collection, and it's a useful investment to train inputters to use authority files.

### **I. THE MEXICAN COLONIAL COLLECTION AT CUSHING MEMORIAL LIBRARY & ARCHIVES**

The Mexican Colonial Collection at the Cushing Memorial Library and Archives at Texas A&M University consists of over 4,000 books, manuscripts, broadsides, newspapers, pamphlets, and religious works covering practically all aspects of life in New Spain and Mexico from the 16th to the mid-19th centuries (i.e., the end of the Maximilian era). This collection encompasses the fields of history, anthropology, linguistics, religion, political science, philosophy, civil and canon law, military science, the history of the book, the sciences, and arts. There are legal documents, including the tribunal of the Inquisition; texts in native languages; theological and religious

---

<sup>1</sup> Associate Professor, Special Collections Cataloger and Italian Resources Librarian, Texas A&M University Libraries.

<sup>2</sup> Associate Professor, Serials Cataloging. Texas A&M University Libraries.

<sup>3</sup> Clinical Assistant Professor, Curator of the Mexican Colonial Collection; Director, Primeros Libros de las Américas, Texas A&M University Libraries.

<sup>4</sup> Coordinadora de Automatización, Biblioteca Francisco Xavier Clavigero, Universidad Iberoamericana.

<sup>5</sup> Jefe de Catalogación, Hemeroteca Nacional de México.

works (many of which contain beautiful woodcuts and engravings); newspapers and speeches concerning Mexico's independence movement, and even cookbooks. While most of these items are in Spanish, a number are in Latin, Nahuatl, other indigenous languages, French, English, or Italian. A substantial number of these items are unique, Texas A&M being the only U.S. institution owning copies.

## II. PROJECT OVERVIEW

Several hundred bound volumes containing multiple works, and over 600 single-sheet items required cataloging to make them discoverable. In a 2012-2014 CLIR project the investigators - three from Texas A&M, and two from libraries in Mexico - developed a methodology whereby Hispanophone students<sup>6</sup> input bibliographic metadata, in stages, into an easy-entry web-based template developed in SharePoint<sup>7</sup>. The challenge was to discover strategies to enable students to produce records similar to those produced by a professional cataloger, and which would amalgamate well with records already in the bibliographic catalog.

We divided the work into stages to determine how students would handle increasing degrees of difficulty, and the associated costs of generating more complex records. An instruction manual with illustrations and examples was prepared and students underwent a brief period of training.

These stages were as follows:

- *Basic Descriptive Elements:* The students were asked to input into the template values for simple descriptive elements .
- *Keywords:* Students “tagged” the items, using terms they thought described the items.
- *Additional Descriptive Elements for Rare Books:* In the 3<sup>rd</sup> stage students input content into additional fields representing descriptive elements particular to the features of rare books.
- *Single Sheet Items:* As they required some distinct elements, a second template was designed for the single sheet and broadside items.

| Basic Descriptive Elements  |                          | Additional Elements     |
|-----------------------------|--------------------------|-------------------------|
| Title                       | Author                   | Illustrations           |
| Variant Title               | Author 2                 | Illustration note       |
| Subtitle                    | Author: Corporate Body   | Ornaments               |
| Variant Subtitle            | Place: Country           | Font Color              |
| Statement of Responsibility | Place City               | <u>Printer's Device</u> |
| Publication: Place          | Printer: Authorized Form | <u>Binding</u>          |
| Printer                     | Printer: Corporate Body  | Marca de Fuego          |
| Date                        | Date: Data Form          | Handwriting             |
| Language                    |                          | Bookplate/Ex Libris     |
| Pages                       |                          | <u>Bookstamp</u>        |

Figure1: Sample Elements

<sup>6</sup> A total of thirteen student technicians were hired to perform the metadata entry. Strong skills in reading and writing Spanish were mandatory; attention to detail and facility with data entry were other requisites. All of the students were native or near-native Hispanophones; half were international students or permanent residents.

<sup>7</sup> SharePoint is a web application framework and platform developed by Microsoft which facilitates collaborative projects. <https://products.office.com/en-us/SharePoint/collaboration>

- *Use of an Authority File to find Controlled terms:* The students were instructed in the use of the *Virtual International Authority File (VIAF)* to test whether they could effectively select the appropriate name headings for the item in hand.

### III. Helping Students to Populate the Record Elements

So far, we have outlined how the materials in the collection were described. To make the items more discoverable we would like to additionally provide access points for persons and corporate bodies, and some type of subject access. Can students with only rudimentary training provide such access? We assumed that our students would be able to furnish uncontrolled access, meaning names and terminology derived from natural language usages, and not from any pre-determined list. Keywords, “tags”, free-form notes (including summaries), corporate names in natural language order, and personal names in a simple Surname(s), Forename(s) format are all examples of uncontrolled access. We also inferred that the students’ knowledge of the Spanish language and of naming practices in the Hispanophone world would help them formulate names correctly. Then, another question arises: If the students are successful in assigning uncontrolled terms, can they be trained to utilize controlled vocabularies such as the RBMS genre terms, Library of Congress Subject Headings, devise headings, or copy heading from an authority file?

Since destined for the bibliographic catalog, we wished the records to conform as much as possible to established bibliographic norms. At the same time, traditional cataloging requires mastering numerous complex standards such as rules for description, headings, and subject strings. This paper focuses on developing approaches which enable minimally trained students to populate bibliographic fields. We enjoyed the challenge of determining just how close we could come in enabling students to produce records which would approximate those produced by trained catalogers. We experimented with several strategies.

#### 1. *Uncontrolled Values*

Some elements were satisfied with uncontrolled values; that is, the students input content depending on their judgment. Such values cannot later be indexed, and are only useful for keyword searching of the records, providing imprecise, inconsistent retrieval, faceting, and collation.

- *Subject keywords:* Students were instructed to consider how they would “tag” the item, and input into an open ended *Keyword* field those terms *they* would use if searching for the item in hand. Students reported being comfortable with this task, and that the information on the item preliminaries often adequately indicated the item’s nature and contents. In other cases, they demonstrated resourceful ingenuity in searching the Internet for information about the item itself or the entities named on the item, to appreciate the historical context enough to determine appropriate keywords.
- *Summaries:* For the proclamations and broadsides, students input summaries describing the nature of the item. At times, this summary could be taken verbatim from a vendor supplied description.
- *Uncontrolled Titles:* Students were asked to input the titles as found on the item. They were not required to identify the original title if the item was a translation and the records will be

lacking uniform titles (if required), unless subsequently supplied by a cataloger. We considered these elements too difficult to require from students.

- *Uncontrolled Author Names*: In the first stage of constructing records, students were asked to input the author name in the *Author* field in the form [*Surname(s), Forename(s)*], but otherwise as found on the item. We thought this would be an easy task -- but it was found to be the most difficult, as discussed below.

## 2. *Controlled Values from Drop-Down Lists*

One of the reasons SharePoint was selected as a tool for hosting our record database is that it permits embedding in the fields drop-down lists of possible values. We used terminology from standard controlled vocabulary lists; or, if a name, we used the authorized form found in the authority file. Database administrators may also implement rules controlling whether a selection is required, or if multiple values are permitted.

- *Static lists*: Embedded lists of controlled values were embedded in these element fields: *Country, Language, Illustrations, Ornaments, Font Color, Binding Types, Author role, and Title Source*. Students were given instructions and definitions to enable them to select appropriate values.
- *Static lists. Genre*: A list of collection-appropriate genre terms selected from the *Rare Books and Manuscripts (RBMS) Controlled Vocabularies* was embedded in the *Genre* element field.
- *Static lists. LCSH Subjects*: Broad subject headings selected from the *Library of Congress Subject Headings* were used in the *Subject* element field.

|       |   |
|-------|---|
| Genre | <input type="checkbox"/> Account books.<br><input type="checkbox"/> Administrative regulations.<br><input type="checkbox"/> Almanacs.<br><input type="checkbox"/> Autobiographies.<br><input type="checkbox"/> Baptismal records.<br><input type="checkbox"/> Biographies.<br><input type="checkbox"/> Breviaries.<br><input type="checkbox"/> Broad­sides.<br><input type="checkbox"/> Calendars.<br><input type="checkbox"/> Catechisms.<br><input type="checkbox"/> Catechisms.<br><input type="checkbox"/> Confessions.<br><input type="checkbox"/> Contracts.<br><input type="checkbox"/> Devotional calendars.<br><input type="checkbox"/> Devotional literature.<br><input type="checkbox"/> Eulogies. |
|-------|---|

Figure2: Genre Element with Embedded List

|              |  |
|--------------|--|
| Subjects/Geo | <input type="checkbox"/> Mexico ðx Antiquities.<br><input type="checkbox"/> Mexico ðv Biography ðy 16th century.<br><input type="checkbox"/> Mexico ðv Biography ðy 17th century.<br><input type="checkbox"/> Mexico ðv Biography ðy 18th century.<br><input type="checkbox"/> Mexico ðx Commerce ðy 19th century.<br><input type="checkbox"/> Mexico ðx Commerce ðy 16th century.<br><input type="checkbox"/> Mexico ðx Commerce ðy 17th century.<br><input type="checkbox"/> Mexico ðx Commerce ðy 18th century.<br><input type="checkbox"/> Mexico ðx Commerce ðy 19th century.<br><input type="checkbox"/> Mexico ðx Economic conditions ðy 1540-1810. |
|--------------|--|

Figure3: Subject Element with Embedded List

The greatest advantage is that such lists permit us to control the values the students may input, ensuring compatibility with values used in existing cataloger-produced records. This promotes indexing and faceting when the new records are amalgamated into the catalog. As a minor matter, the use of controlled vocabulary lists also reduces typographical errors on the part of the students.

Such lists are most useful when working with smaller or homogeneous collections where the number of applicable terms or headings is limited. Conversely, a broad collection like the Mexican Colonial requires more terms, yet this creates a problem: a larger list of specific terms that becomes unwieldy when embedded in a template field, and which requires more student training for selection of appropriate term. The only way to revert the list to manageable size is to use broader terms. Furthermore, if the cataloger must examine a large number of items in a heterogeneous collection to develop a large list of precise terms, this is so labor intensive that he/she might as well just assign the subject headings in the traditional manner.

To overcome these problems, we attempted to modify this strategy and experimented with making lists dynamic, expanding them as needed via feedback provided by the students.

- *Dynamic lists. Example – City of publication:* For the *City of Publication* element, a list of city names in the form authorized by the *Name Authority Cooperative Program (NACO)*<sup>8</sup> authority file was developed, comprising the cities likely to appear on the items in the collection. But if the students encountered a city of publication that was not on the embedded list, they could select the “*Specify your own value*” option and input the new city name as found on the piece. A cataloger would periodically review the database to find these new city names, look up the authorized form of the city, add the authorized form of the new city to the list as a value likely to occur on items in this collection, correct that record by removing the ‘*Specify your own value*’ choice, replacing that choice with the authorized form of the city, now on the list.

- *Dynamic lists. Example – Authorized name of printers:* Similar to the controlled city list, we developed a list of authorized names of printers likely to be associated with this collection.

Using the same process as described for the list of cities, new printer names could be added to the dynamic list as needed.



Specify your own value:

<sup>8</sup> <http://www.loc.gov/aba/pcc/naco/>

Using feedback generated by the student's examination of the items permitted us to provide larger numbers of controlled terms, but still this strategy is more useful when working with a more homogeneous collection, i.e. the *Primeros Libros*, which is limited to 16<sup>th</sup> century imprints and thus is associated with a limited number of printers. For the larger *Colonial* collection, as the number of printers grew throughout the 18<sup>th</sup> and 19<sup>th</sup> centuries, it became increasingly difficult to modify the list in a timely manner.

### **3. Student Use of Online Authority File**

To avoid the impracticality and labor of embedding an extensive list in fields we attempted to test whether students could learn to effectively search name authority files and discover the authorized form of a name for themselves.

For a subset of records which did not already have name headings provided by one of the previous strategies, the students searched the *Virtual International Authority File* (VIAF). They copied the 1XX field containing the authorized form of the name into the SharePoint template fields. This strategy required an investment in training the students to navigate the authority file and choose the correct record corresponding to the needed name. There are benefits of training the students to directly use an authority file discussed in the conclusion of the paper. However, occasionally the students had difficulties identifying when they had the correct authority record for the name found on the item.

## **IV. Discussion**

### **1. What was easy?**

- Title element: because students just copied what they saw.
- Elements for which we were able to provide drop down lists.
- Even the more dynamic list was easy to manage until the amount of feedback was overwhelming.

### **2. What was difficult?**

- *Occasionally the student had difficulty identifying the author(s):* In the case where several names were mentioned on the item, students were sometimes misled by the wording on the page into selecting the wrong individual as the author.

*Example:* in Spanish the word *por* can mean either 'for' or 'by', depending on the context. Usually the name of an author is introduced by the word *por*, as in:

Don Quijote de la Mancha *por* [by] Miguel de Cervantes

However, word *por* can also mean 'for'. In funeral sermons where the title statement appears in the form 'Oraciones funerales *por* [name of the defunct]' students automatically assumed the individual named after the word *por* was the author. The word *del* can also be misunderstood in the same way, as in the example '*Solemnes exequias del illmo. señor dr. d. Jose Gregorio Alonso de Hortigosa, obispo que fue de la ciudad de Antequera ...*' for which Hortigosa was entered in the record as the author, when he is really the defunct. This mistake happened frequently enough to provoke the humor of our paper's title

## WHAT CATALOGERS KNOW ABOUT SPANISH NAME HEADINGS

Making access points for personal names can be daunting, even to people familiar with the language of the items being cataloged. Catalogers have the *National Authority File* and the rules set forth in *Resource Description and Access* (RDA) and the *Anglo-American Cataloging Rules* (AACR2) to guide them, but non-catalogers would be unaware of these resources, and would require complicated training to use them. This project happened to involve Spanish language materials and Spanish speaking students were chosen with the hope that their knowledge of Spanish would permit them (with only basic instructions) to formulate names correctly. That did not always happen.

Some Spanish names consist solely of Forename(s) Surname. As is the case with names in English, a basic Surname, Forename(s) heading would suffice – as long as the design of a project does not require that each name is unique.

But most Spanish names are not so simple. Most people in the Hispanophone world use the surnames of both parents. The father's surname comes first and, thus, serves as the beginning of the access point (for example, Juan Antonio Pérez Vásquez, whose father's family was Pérez and whose mother's was Vásquez; his heading would be Pérez Vásquez, Juan Antonio). Prepositions such as “de”, articles (el, la, los, las), or the two in combination (del, de la, de las, de los), or conjunctives such as “y” may occur as part of a name<sup>1</sup>. Often, when a woman marries, she appends her husband's first surname to her family's surname(s) with “de”. Thus, if Isabel María Ortiz y Pino marries Juan Antonio Pérez Vásquez, her full name would be Isabel María Ortiz y Pino de Pérez; her surname begins with Ortiz.

According the rules for establishing Spanish names in both AACR2 and RDA, the heading should begin with the first element of the surname, excluding prepositions like “de”.

Before the 17<sup>th</sup> century, some people simply appended the name of a place or an attribute to their forenames. An example of this is Pedro de Alcántara, a 16<sup>th</sup>-century Spanish saint.

It is customary among Catholics to name their children after saints; so on occasion people are named after saints with these phrase-like names, e.g. *Pedro de Alcántara Martínez*. It is easy to think that “de Alcántara” is part of this man's surname, but in this case, it is part of his forename; he's named after the 16<sup>th</sup>-century saint. The name for him is *Martínez, Pedro de Alcántara*. It requires some expertise to input correctly names such as this.

Variations in Spanish form or spelling frequently occur during the time period covered in this collection, i.e. “Yturriaga” vs. “Iturriaga”: “Leyva” vs. “Leyba”, “Florez” vs. “Flores”; “Gerónimo”, “Gerónimo”, and “Jerónimo”: “Xavier” vs. “Javier” etc. The popular forename José was usually rendered “Joseph” or “Josep” as late as 1800. Another potential difficulty is the versions of a name appearing in other languages. Pedro Martínez wrote in Latin, using the form “Petrus” instead of “Pedro” on the items. “San Juan Nepomuceno” is the Hispanicized form for a saint Bohemian-born saint, John of Nepomuk.

Titles of nobility add yet another layer of complexity, particularly if a person tended to use his or her title on documents rather than their name. The Colonial Mexican collection contains a document by General Francisco Xavier Venegas addressed to “Excmo. Sr. Duque del Infantado”; the Duke's name does not appear in or on the piece, and the student entering the metadata only gave Venegas' name. It required research to reveal that the Duke's full name was Pedro de Alcántara Álvarez de Toledo y Salm Salm.

Other categories of names present problems in any language: those represented by initials only, and pseudonyms. For example, one would need to look in the authority file to know that “Ignacio Tomay” (also spelled Tomai or Thomay) was the pseudonym of José María Genovese, a Sicilian-born Jesuit living in New Spain

- Formulating name headings, without help from the VIAF file. The following types of mistakes were frequent.

| Type of Error  | Entered  | Correct   |
|--|--|---|
| <i>Incorrectly selecting an element as the first element of the surname.</i>   | Antonio de Oviedo, Juan Jose<br>De Urdanivia, Cárlos Joseph Diez   | Oviedo, Juan José Antonio de<br>Díaz de Urdanivia, Carlos Jose  |
| <i>Incorrectly including abbreviations of titles as part of the name</i>   | De Castro, D. Juan Francisco<br>Br. D. Juan Gabriel de Contréras   | Castro, Juan Francisco de<br>Contréras, Juan Gabriel de   |
| <i>Beginning a name heading with "de"</i>  | de Contreras, Juan Martin<br>de Jesus Soria, Rafael  | Contreras, Juan Martín de<br>Soria, Rafael de Jesús   |
| <i>Incorrect use, or lack of, diacritics</i>   | Perez Quixano, Agustin   | Pérez Quixano, Agustín  |
| <i>Lack of recognition when a person's forenames include the name of a saint</i>   | Nepomuceno Rivas, Juan   | Rivas, Juan Nepomuceno<br>( <i>Nepomuceno</i> is really part of the man's forename as he was named after St. John of Nepomuk) |
| <i>Lack of recognition of non-Spanish forms of names; i.e., names in Latin, Italian, or French, or even archaic spellings of Spanish</i> | José Gregorio Alonso de Hortigosa<br>Joseph Gregorii Alfonsi de Ortigosa<br>(Latinized form of the name) | Alonso de Ortigosa, José Gregorio<br>(heading from Biblioteca Nacional de España)   |
| <i>Non recognition when an individual had only a forename (despite training)</i>   | Transfiguracion, Francisco de la<br>Sales, Francisco de  | Francisco de la Transfiguración<br>Francis de Sales   |

### 3. *What was impossible?*

- Works by corporate officials are entered under their official title.* This is something we would not expect a non-cataloger to know. For example, a document issued under the authority of the viceroy Martín de Mayorga should have as the author the form *New Spain. Viceroy (1779-1783 : Mayorga)* and not *Mayorga, Martín de*. We accepted that records for official documents would incorrectly personal name forms in the element fields.
- Translations.* Without substantial training, and extensive searching in OCLC, it would be difficult for students to determine when a publication was a translation, or what the original title would be. So we accepted that these records would not have translation uniform titles and would not collate with the original editions.
- Pseudonyms.* Although sometimes -- with the help of the VIAF files, or Internet searching for information-- the students were able to find the real name of an author appearing on the item as a pseudonym or initialism, this was a difficult task we didn't require.

## V. Conclusion.



The purpose of this paper is not to argue for the strict use of controlled vocabularies and headings in bibliographic descriptions, either in traditional catalogs or other databases using non-MARC based records. We understand that it is acceptable for record standards to vary depending on the type of access needed for a particular collection or project. And it is clear that there is a lack of trained personnel to create records at the highest level for backlog of items which need to be processed. It was not our intention to demonstrate the value of the work of LC, OCLC (via VIAF), Dublin Core, ALA, etc. of promoting the sharing of authority records and of controlled vocabularies— yet this project has reiterated to us the importance of their efforts.

In doing so, these associations enable project managers to focus effort from the difficult, expensive, time consuming investment of training non-catalogers (usually temporary personnel) to construct headings or use vocabularies according to standards, to the lesser task of training staff to navigate the authority record databases and select the correct record showing the authorized form of the topical or name heading needed. Users of the authority file records can see the correct authorized form of a heading, including diacritics, order of the elements of a name, correct spelling, etc. And thus they do not need to know how to construct a heading. Professional catalogers — who are trained in the rules and standards for constructing headings and vocabularies -- can dedicate themselves to building the authority records.

Use of the online authority and vocabulary resources frees us from training non-catalogers how to construct a heading, but the focus shifts to training them to identify the appropriate record. Experienced catalogers can identify non-obvious bits of information in AACR2 authority records (including in the citations) which help them determine that a record corresponds to the individual on the item in hand. Non-catalogers, who lack the training and experience, cannot see these clues. Our students selected incorrect records due to their inability to read these clues. The new RDA attributes<sup>9</sup> permit a non-cataloger to better identify which record is the correct one, thus permitting the staff to focus on inputting the proper information rather than debate among several possible records. It was not our original intention to demonstrate the value of RDA authority record attributes, but after observing the students, the value of the attributes to non-catalogers was demonstrable.

#### ADDITIONAL READING

Almuzara, Leticia Barrionuevo, M. Luisa Alvite Díez, and Blanca Rodríguez Bravo. "A Study of Authority Control in Spanish University Repositories." *Knowledge Organization* 39, no. 2 (2012): 95-103.

Caimei Lu, , Jung-ran Park, and Xiaohua Hu. "User Tags Versus Expert-Assigned Subject Terms: A Comparison of LibraryThing Tags and Library of Congress Subject Headings." *Journal of Information Science* 36, no. 6 (December 01, 2010): 763-779.

---

<sup>9</sup>Associated dates, Associated Place, Address, Field of Activity, Associated Group, Occupation, Gender, Family Information, Associated Language, Fuller Form of Personal Name.

DeZelar-Tiedman, Christine. "Exploring User-Contributed Metadata's Potential to Enhance Access to Literary Works: Social Tagging in Academic Library Catalogs." *Library Resources & Technical Services* 55, no. 4 (2011): 221-233.

Maron, Deborah, et al. "Lo-Fi to Hi-Fi: A New Way of Conceptualizing Metadata in Underserved Areas with eGranary Digital Library". *Proceedings of International Conference on Dublin Core and Metadata Applications*. Austin, TX (2014) 37-42.  
[dcevents.dublincore.org/IntConf/dc-2014/paper/download/259/228](http://dcevents.dublincore.org/IntConf/dc-2014/paper/download/259/228)

Yusuf, Felicia and Christopher Nkiko. "Involvement of Non-Professionals in Cataloguing: A Survey of Practices in Three Academic Libraries in Nigeria." *Malaysian Journal of Library & Information Science* 15, no. 1 (2010): 83-90.