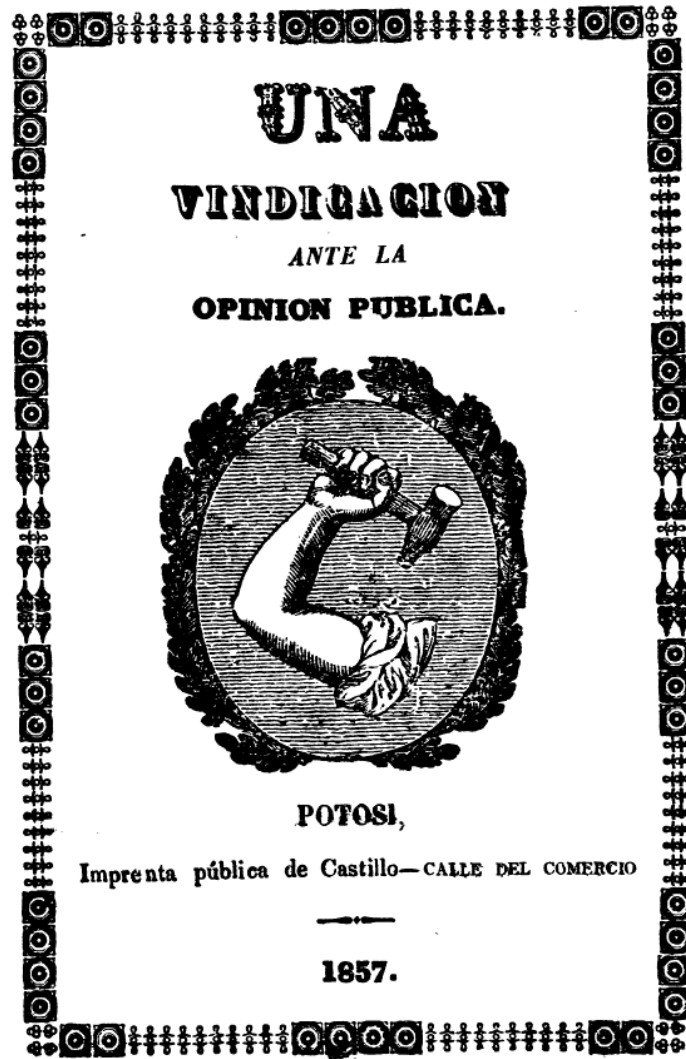


LDI Round 4  
Widener Library Project to Digitize, Process, and Save Widener's  
Latin American Pamphlets

Final Report



Collection Development Department  
Widener Library, Harvard College Library

Submitted by Project Manager Dan Hazen,  
Maggie Hale and Richard Lesage  
November 29, 2006

## TABLE OF CONTENTS

Project Staff .....	3
Widener's Latin American Pamphlets .....	4
Project Benefits .....	4
Selection and Cataloging .....	5
Reformatting .....	12
Storage .....	16
Staffing .....	16
Project Schedule .....	16
Budget .....	17
Access and use .....	17
Appendix 1: Cataloging Statistics .....	20
Appendix 2: Project Workflow .....	21



**POTOSÍ.**

Illustration from *28 de diciembre de 1869 en Potosí*. [Potosí, Bolivia ]: Tip. del Progreso, 1870.

## **PROJECT STAFF**

### **Project Manager:**

Dan Hazen, Associate Librarian for Collection Development, Harvard College Library

### **Project Associates w/Name, Title and Contact Information:**

Lynn Shirey, Librarian for Latin America, Spain, and Portugal, Collection Development Department, Widener Library, Harvard College Library

### **Project Team:**

Preservation and Imaging Services, Harvard College Library

Richard Lesage, Project Cataloger, Bibliographic Services

Katherine Sweeney, Preservation Cataloger, Bibliographic Services

Ann Kern, Senior Cataloger, Bibliographic Services

Amy Jackson, Preservation Assistant, Bibliographic Services

Todd Bachmann, Associate Head, Imaging Services

Maggie Hale, Project Coordinator, Librarian for Collections Digitization

Ethel Hellman, Collections Conservator for Widener Library

Mingtao Zhao, IT Specialist

Scanning and Microfilming Services

Digital Imaging and Photography Services

### **Project Advisors**

Robin Wendler, Metadata Analyst, Office for Information Systems, Harvard University Library

Stephen Chapman, Preservation Librarian for Digital Initiative, Weissman Preservation Center, Harvard University Library

Wendy Gogel, Digital Projects Program Librarian, Office for Information Systems, Harvard University Library



**Latin American Pamphlet  
Digital Collection**

## WIDENER'S LATIN AMERICAN PAMPHLETS

Harvard is one of few institutions to have consistently collected pamphlets as part of its research collections. Our extensive holdings from Latin America, many of them scarce or unique, particularly detail the region's historic moments and mentalities, rhetoric and realities. Faculty, researchers and graduate students, as well as undergraduates preparing term papers and honors theses, figure among the principal users of these primary resources. The David Rockefeller Center for Latin American Studies, through its programs for visiting scholars, research grants for Harvard faculty, and summer library grants for academics from other institutions, directly and indirectly sponsors many of these researchers. The Andrew W. Mellon Foundation's program to support Harvard's graduate students in Latin American History has likewise intensified local interest and demand. While Widener's pamphlet holdings are extraordinary, however, their use has until recently been artificially depressed.

When the Latin American Pamphlet Project took shape, Widener's Latin American pamphlet holdings were estimated at 336 pamphlet volumes and 414 pamphlet boxes. Each of the pamphlet volumes held as many as thirty or more pieces; most of the boxes—a deceptively elegant name for what were (and are) often shabby and acidic manila envelopes—contained somewhat fewer. Latin American pamphlet holdings from the volumes alone totaled about 4,400 items, representing nearly 170,000 pages. Cataloging for the materials was extremely limited, since neither the volumes nor the boxes were represented in HOLLIS and most of the separate pieces likewise lacked bibliographic records. The only record for many volumes or boxes was a very cursory listing [e.g. “SA 3447.1 -- 6P -- Pamphlet Vol. -- Mexico. History 1860-66.”] in the old printed Widener shelflist. All of these pamphlets, sitting on open shelves without bibliographic control, were at once inaccessible to most users and invitations for theft.

### PROJECT BENEFIT

As a result of the LDI funded *Project to Digitize, Process, and Save Widener's Latin American Pamphlets*, and the continuation of the work with Harvard College Library funding, users now have greatly improved access to over 4,500 of Widener's Latin American pamphlets. The LDI funding allowed for the cataloging, preservation, digitization, and safeguarding of the first 3,000 of these Latin American pamphlets. Full online bibliographic records in HOLLIS have been created for each pamphlet, with a link from each record to immediately accessible digital images of the entire piece, delivered through Harvard's Page Delivery Service (PDS).

In addition to accessing the pamphlets through the HOLLIS records, it is now possible to access the Latin American Pamphlets through the *Latin American Pamphlet Digital Collection Virtual Collection*. This web-based catalog provides a means of searching only the records of digitized Latin American Pamphlets from Widener Library's holdings.

Browsing in the stacks traditionally allowed users to gain an overview of complete collections. They could also carry out cursory inspections of previously unknown items and then decide what might deserve a closer look on the basis of such cues as typeface and page layout, the level of language and tone of exposition, and the presence of footnotes and illustrations. But today, as

collections are more widely dispersed and as more and more materials are housed offsite, browsing is ever less adequate in fulfilling either of these functions. Users must instead rely on catalog records.

MARC records, even when enriched with extensive subject indexing and very full descriptive notes, rarely convey the significance of a publication as vividly as even the most abbreviated contact with the piece itself. Nonetheless, these records are central to user success and satisfaction. Librarians therefore need to make them as evocative as possible. Using digital technology to link full, traditional bibliographic records to page images of pamphlets has accomplished much the same end as browsing in the library's stacks, with the vast additional advantage that users from all over the world can view the materials.

The enhanced access to the pamphlets that has resulted from this project has also benefited collection development. With the pamphlets in Widener's holdings cataloged and digitized, it is now possible to avoid duplication of effort and expense. Lynn Shirey, Librarian for Latin America, Spain, and Portugal, reports locating titles of pamphlets being offered to Widener and being able to avoid acquiring duplicates.

In addition to greatly enhancing access to the Latin American pamphlets, this project contributed to the development of new infrastructure (Virtual Collections); the incorporation of national best practices to facilitate broad discovery and use (through the use of the Digital Library Federation Registry); and the refinement of production workflows, including cataloging, outsourced film scanning, and structural metadata production. The refinement of workflows has allowed the work on the Latin American pamphlets to move from Project to Program. With additional funding from the Harvard College Library, the cataloging and digitization of pamphlets will continue.

## SELECTION AND CATALOGING

### - Notes by the Project Cataloger, Richard Lesage

*"For many days have been cataloguing pamphlets and books received at the College Library. The title of each pamphlet is entered as minutely as that of the most valuable book. Pamphlets are the most valuable part of a Library, which has reference to posterity."* John Langdon Sibley, April 1, 1846

This quote from the private journal of John L. Sibley, then Assistant Librarian of the Harvard College Library, summarizes my endeavor as cataloger for the Latin American Pamphlet Project.

Harvard has indeed consistently collected pamphlets for posterity, but many of these pamphlets have either not received the attention of a John Sibley, or made it to the card catalog. The bulk of the pamphlets selected for the project were acquired by the Harvard College Library under the leadership of William C. Lane, Librarian from 1897 to 1928. The pamphlets were part of historical libraries of famous Latin American bibliographers and book collectors such as Luis Montt<sup>1</sup>, Nicolás Acosta, José Augusto Escoto, and others. The accession of these huge

---

<sup>1</sup> Some Chilean bound pamphlet volumes, were interleaved with ms. notes by Montt, possibly intended for his unfinished *Bibliografía chilena*.

collections within a relatively short period might explain why most of the pamphlets were never processed adequately, and ended up being the Cinderella of the Widener stacks awaiting Digital Prince Charming!

In the following notes to address some of the issues and challenges faced in the cataloging phase of the Latin American Pamphlet Project. I would also like to add some observations which could be useful from a curatorial standpoint.

### **Descriptive cataloging**

The first challenge was inherent to the pamphlet “format” in itself. Pamphlets are generally works where authorship is not straightforward. Very often, they have no title page, or a title page that does not follow the conventions of modern-day publishing. This meant a couple of things for cataloging:

- more time has to be spent extracting data from the whole publication, to fulfill the requirements of a bibliographic record;
- the unconventionality of the title page layouts often made it somewhat of a challenge (hence more time consuming) to follow AACR2 rules. It seemed easier to use DCRB (Descriptive Cataloging of Rare Books) core standards and this was done at the beginning of the project but abandoned.<sup>2</sup> The peer editing practice within the Reformatting Section was producing hybrid records: other catalogers of the team were either unfamiliar with DCRB, or uncomfortable editing records using this standard.

Another challenge was related to the condition of the pamphlets. Many items had lost their title page or cover, the preliminary leaves being the most vulnerable to wear and tear. Cataloging these items required more elaborate searches of bibliographic utilities, comparison of copies, and consultation of annotated bibliographies of Latin American countries (René-Moreno’s *Biblioteca boliviana*, and *Biblioteca peruana*; Valenzuela’s *Biblioteca guatemalteca*; Medina’s works on the printing presses of various Latin American cities, etc.). In many instances perfect copies were obtained for filming (either Widener duplicates or copies from other Harvard libraries) and holdings linked to proper bibliographic records, reflecting the work as such. In some cases, photocopies of title pages and/or missing pages, were obtained through Inter-Library Loan (ILL), and filmed together with Widener’s imperfect copy.

Another difficulty resulted from the prior partial processing of the pamphlets. One third of the items had no bibliographic records whatsoever. Another third had stub records, with virtually no access points. Several items had been reclassified many times over, without proper annotation in HOLLIS. In this respect the project was really the clearing up of the backlog built up during the first quarter of the 20<sup>th</sup> century, and inadequately addressed since. Yet another difficulty resulted from the reconversion of the card catalog to HOLLIS, with incorrect bibliographic records exported from OCLC, and not reflecting the copy in hand. The reconversion had also generated duplicate records. Thus, there was also an important “recon clean-up” dimension to the project, as for any project dealing with HOLLIS.

---

<sup>2</sup> Another reason for using DCRB was the possibility of a more detailed pagination, useful for navigation in the Page Delivery Service

In general each item cataloged brought up a series of problems (determination of issues, editions, fixing of headings, etc.) requiring the examination and cataloging of other related pamphlets. In addition to complex cataloging and recon clean-up, “cluster cataloging” would be an appropriate term to describe the cataloging of the Latin American pamphlets.

### Subject analysis

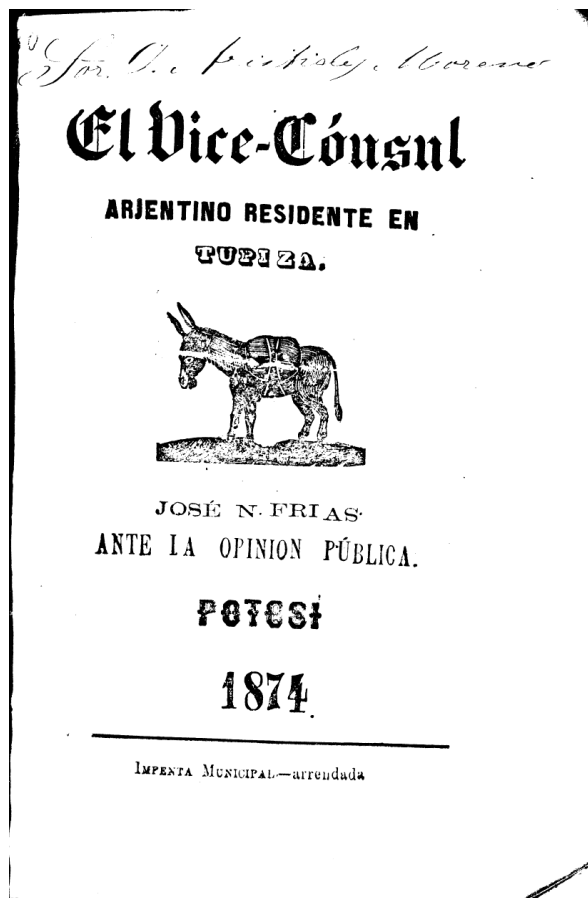
The pamphlets document the emergence of the Latin American colonies as independent states, and all aspects of the social and cultural life of the people of Latin America. Many pamphlets are devoted to the conflicts between Latin American countries all through the 19<sup>th</sup> and beginning of the 20<sup>th</sup> century, boundary disputes and territorial expansion, the description of previously unexplored territories, the relation between Church and state in the nascent states’ quest for identity (the question of allegiance to the Crown; conflict of laws relating to marriage, funerals, burials, freedom of religion, appointment of bishops, etc.). The pamphlets are also a witness to the work of printers fighting for the freedom of the press, like Cajetano R. Daza, the printer of Santa Cruz (Bolivia).

The HOLLIS catalog being the main tool of discovery of Widener holdings, pamphlets in the project were cataloged to full level. Subject analysis was thus thorough, and as precise as possible using LCSH. Moreover in the case of pamphlets with a generic, non-descriptive title, the subject matter was summarized in a note to improve chances of discovery by keyword search in the OPAC.

### Illustrations

The pamphlets have generally few illustrations, except for guides, and other advertising brochures, from the end of the 19<sup>th</sup> century forward, at a time when photomechanical processes facilitated illustrations. Many earlier pamphlets have decorative title vignettes, reflecting the contents of the text, or the geographical location of the printing press (judge’s hammer for legal pamphlets, the Bible and dove symbolizing the Holy Spirit for religious pamphlets, a donkey for satirical pamphlets, a volcano for imprints from Potosí, etc.). Some of the vignettes have, curiously, captions in English, which may point to contacts between Latin America and the United States (possibly missionary presses) in the development of printing.

**Decorative title vignette. Arraya, Francisco.**  
*El vice-cónsul argentino residente en Tupiza.*  
Potosí: Impr. Municipal, arrendada, 1874.



Some of the pamphlets contained photographs. Ten albumen prints have been identified. The bibliographic records for these pamphlets have been given an appropriate form genre heading to denote the presence of albumen prints. This will enable curators to locate photographs, assess their condition, and determine appropriate conservation treatment if need be. More valuable South American photographs, housed in albums or loose envelopes, were found during my visits to the stacks. These, although not selected for the project, were removed from the stacks for security reasons, and set aside for proper cataloging and preservation.



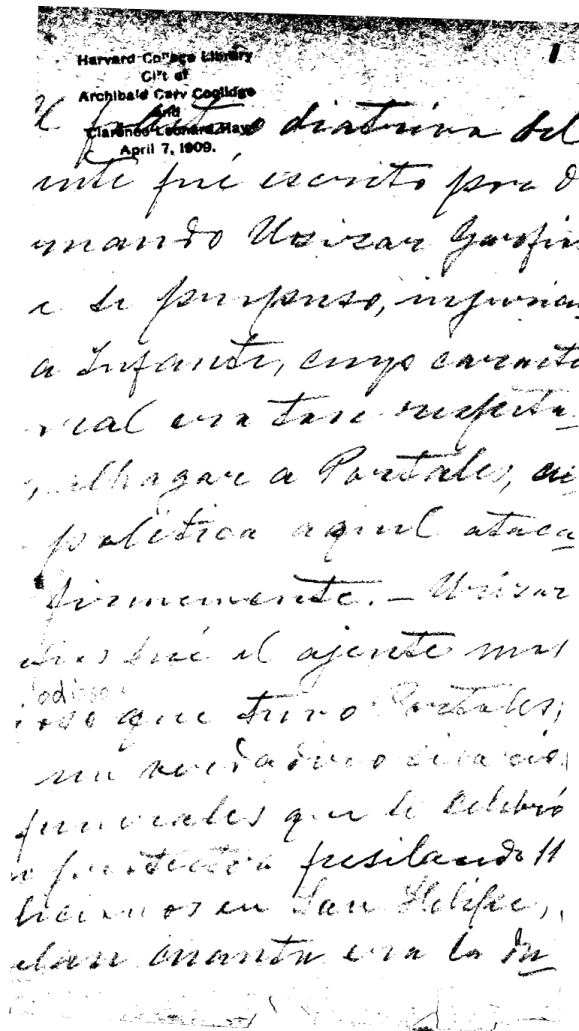
**Albumen print from *Rasgos biográficos del Excmo. Sr. general D. Eusebio Puello y Castro y exposición que meses antes de morir dirigió á S.M. el Rey. Habana: Impr. militar de la viuda de Soler y Cía., 1872.***

### **Provenance**

Holdings for the pamphlets were annotated with copy specific information. Manuscript inscriptions pertaining to ownership (as well as annotations on authorship when found in anonymous pamphlets) were transcribed in extenso, especially those belonging to famous Latin American book collectors. Although a departure from Widener cataloging practice, this was deemed appropriate, considering the historical value of the pamphlets. Provenance information allows the virtual reconstruction of these personal libraries. It can also help elucidate the connections between authors and owners, and interpret the collections as a whole (each item in relation to each other). For practical reasons, the former owners were not always given added entries in the bibliographic records, but owner's names in provenance notes are keyword searchable in HOLLIS, through public or staff mode.

The majority of the pamphlets with provenance information (482) are from the library of Chilean bibliographer Luis Montt. Next in order of importance are the collections of José Augusto Escoto, of Matanzas, Cuba (386) and Bolivian bibliographer Nicolás Acosta (200). A significant number of pamphlets come from the collections of Blas Garay/Edwin V. Morgan, Charles Sumner, and John B. Stetson. The cataloging process also uncovered interesting association copies: to name but a few, a copy belonging to Daniel Florencio O’Leary, Bolívar’s aide-de-camp, and two copies belonging to the library of Nicaraguan poet Rubén Darío, on the theme of Central American federalism. These pamphlets were brought to the attention of Rare Book Cataloger David Whitesell, a Darío specialist. These pamphlets will be transferred to the Houghton Library, and rejoined to Darío’s collection.

**Inscription by Chilean bibliographer Luis Montt found in Urizar Garfias, Fernando. *Contestacion al num. 114 del Valdiviano federal.* Santiago [Chile]: Impr. Araucana, 1836.**



### Scope of project

LDI project funding covered the cataloging and digitization of 3,000 pamphlets. Additional HCL funding covered continued cataloging and digitization for the Latin American Pamphlet Project. By September 2005, 4,250 pamphlets had been cataloged, with slightly over 3,000 digitized. These notes and cataloging statistics reflect the work on these 4,250 pamphlets. At that point the LAP Project was concluded, and cataloging and digitization of Latin American pamphlets continued as a Harvard College Library program for pamphlet digitization. The series title for bibliographic records of digitized pamphlets was thus changed from “Latin American pamphlet digital project at Harvard University” to “Latin American pamphlet digital program at Harvard College Library.” As of October 2006, 4,500 pamphlets have been digitized and are publicly available.

Pamphlets for the Latin American Pamphlet Project were originally selected by the bibliographer for Latin America, from the SA class (Latin America) in Widener Library Shelflist (no. 5, v. 1) published in 1966. The shelflist listed collections of uncataloged pamphlets. Argentina, Bolivia, Chile, Cuba, Mexico and Peru were identified as priority countries in the selection plan. As materials were being cataloged, it was found that Argentina was under

represented. Additional Argentine pamphlets were thus located and added to provide a balance. Additional HCL funding allowed the inclusion of pamphlets from Uruguay, Ecuador, Paraguay, and Guatemala. Thus, with regard to the SA class, the project has covered most of the Latin American countries, with the exception of Brazil<sup>3</sup>, Colombia, Venezuela<sup>4</sup>, Panama<sup>5</sup>, the Dominican Republic, Puerto Rico<sup>6</sup>, and pamphlets relating to Central and South America in general. Some 600 pamphlets relating to countries not covered have been set aside for processing, and will be part of the ongoing reformatting and preservation program for Latin American pamphlets.



**Istituto coloniale italiano. *A ricordo della commemorazione centenaria della indipendenza della Repubblica Argentina.* [Roma]: Istituto Coloniale Italiano, [1910].**

It should be noted that many pamphlets, not part of the original selection from the shelflist were included as a more thorough inventory by country was being conducted. These added items came from some pamphlet volumes discovered in the Widener stacks which were neither listed in the Widener shelflist or the HOLLIS catalog. Also added were pamphlets not marked as “pamphlets”<sup>7</sup> in the shelflist, but truly pamphlets in need of adequate cataloging and

<sup>3</sup> Not selected for the project

<sup>4</sup> Most of the older Colombian and Venezuelan pamphlets have been previously transferred to the Houghton Library

<sup>5</sup> Not selected for project. Pamphlets dealing with the Panama Canal could also form a thematic project of its own

<sup>6</sup> Pamphlets relating to the Dominican Republic and Puerto Rico (with the exception of some pamphlets in the Cuba class range) were not selected for the project

<sup>7</sup> Only pamphlets with no card catalogue are annotated as “pamphlets” in the Widener Shelflist

preservation. Related pamphlets were also added to the project, as in the case of a polemic where pamphlets respond to each other. If a pamphlet was about a specific work, that work too was added if the page count did not exceed 150 pages. The inclusion of related pamphlets was necessary both for the interpretation and intelligence of the texts, and also as a matter of fairness to both sides in a debate.

It should also be noted that some 400 Latin American pamphlets, microfilmed in the 1970's and until recently stored at NEDL, are practically inaccessible. Collective records (with no access to title/author) exist for the microfilms only, and these, unfortunately, have not been properly targeted, rendering their use very difficult.

Finally, a word about the SAL (Latin American Literature) class. One hundred rare pamphlets with SAL call numbers, not part of the original selection for the project, were included in the project. Several hundred more of these brittle pamphlets are awaiting proper cataloging and reformatting. They are the literary counterpart to the historical and political events which make up the subject matter of the SA class. The SAL class pamphlets, as well as the pamphlets from NEDL could also form part of the new ongoing reformatting and preservation program for Latin American pamphlets.

### **Copyright**

With regard to copyright, the cut off date for our selection was 1922. Without going too much into the specifics of copyright law, 1922 clearly demarcates works in the public domain from works under copyright, in the United States. For works published outside of the United States, the clear line of demarcation is 1909, but most works published until 1922, without subsequent republication with a copyright notice, are also in the public domain. The project includes 4,074 items published before 1923. After careful consideration, and a realistic risk assessment by the Latin American Bibliographer, it was decided to include some pamphlets even beyond the cut off date of 1922 in the project. Thus, 176 items have publication dates ranging from 1923 to 1937. In the unlikely event that copyright will be claimed for any of these items, digital access to the related pamphlet will be removed.

### **Some technical issues**

#### ALEPH

It was definitely a challenge to start the project at a moment when the HOLLIS database was being migrated to ALEPH. The cataloging was steadfast, in spite of the flaws, inadequacies, and uncertainties of the new system. For almost two years, authority work was impossible, or not recommended, because ALEPH could not protect local fields in HOLLIS. This created an authority work backlog which will need to be addressed. In spite of this drawback, some 250 authority records were created in order to break conflicts for author headings.

On the positive side, the great advantage of ALEPH was the expanded capabilities of searching. This proved to be helpful in identifying works, when the item in hand was imperfect. ALEPH was also very useful in gathering bibliographical data for the project (in addition to the spreadsheets maintained for cataloging, and other workflow statistics).

#### OCLC participation and DLF Registry

Holdings for digital masters in HOLLIS were created for all titles in the project. Currently, these digitized versions are only available to the public via bibliographic records in HOLLIS. We expect that at some point in 2007 all records of the completed titles will be exported to OCLC where a broader scholarly community can discover and link out to the digital objects. In addition to discovery, it should make it easier for other institutions to avoid duplication of digitizing effort and/or incorporate records and links directly into their OPACs. Due to early technical constraints and the need to refine and provide additional bibliographic data elements, these records will need to be treated as a separate batch load to OCLC. The additional data will be added to the HOLLIS records at the creation of the batch load to ensure that all records are compliant with the Registry of Digital Masters.

### **Documentation**

Spreadsheets were created for the project, documenting various aspects of the reformatting and cataloging process, namely: disbinding, HD requests, number of items present in volume, items missing, related items added to the project, original records created, stub records brought up to full level, records exported from OCLC, added Widener holdings for existing bibliographic records, duplicates within the project, fragmentary copies, broadsides, post-1930 copies falling under copyright, deletion of duplicate records, authority records in need of creation, conservation treatment, and ILL borrowing. The spreadsheets are organized by countries and call no. Another spreadsheet was created for items in need of direct digital capture.

## **REFORMATTING**

Cost-effective production digitization of bound printed pages was a challenge at the time the Latin American pamphlet project was being planned. Following exploration of reformatting options for this project, it was determined that using a hybrid approach – microfilming materials first and then scanning the microfilm – would best meet the project's functional requirements for page images with legible text at the lowest possible costs. Direct digital capture was used for pages with meaningful color. 3,000 pamphlets were digitized in the LDI funded project, containing 114,270 pages. 1,311 images were selected for direct capture.

### **Preparation**

Preparation and cataloging of the pamphlets was done by the Reformatting Preparation Unit of the Preservation and Imaging Services Department of the Harvard College Library, closely following established workflows for microfilming preparation. Select pages, such as pages with color images, were identified during preparation for direct digital capture following filming. Volumes in need of treatment in the conservation lab prior to filming were also identified at this time.

### **Conservation treatment of pamphlets**

The Collections Conservation Unit of the Preservation and Imaging Services Department of the Harvard College Library performed any conservation work necessary to prepare the pamphlets for filming. The goal of any conservation work performed on the bound volumes of pamphlets or on individual pamphlets was to get the materials to the point of 100% openability (opening to 180 degrees) for microfilming.

Although many of these pamphlets have artifactual value, many had already been damaged in the binding process. Conservators used their judgment regarding treatment. If, due to prior intervention, condition of material was not pristine, some trimming and destruction of spine folds and cover margins was done. Individual pamphlets were not disbound unless their openability was limited. 52 volumes, with an average of 24 pamphlets per volume were prepared for reformatting, including disbinding and stabilization, by Collections Conservation.

After reformatting, materials were returned to conservation for final rehousing. Resewing or repair of bindings was minimal. The majority of materials remained in slings (paper folders) and in envelopes with no repairs. Volumes that were not disbound were placed in boxes as necessary. All materials were then sent to the Harvard Depository, made non-circulating to be served through the Phillips Reading Room. As page images will be available for all pamphlets, future use of the original pamphlets is expected to be extremely limited.

### **Microfilming of pamphlets**

The Microfilming Unit in the Preservation and Imaging Services Department of the Harvard College Library created preservation microfilm of all pamphlets. All film was inspected for physical quality, legibility and bibliographic accuracy. An extra duplicate negative for digitization was initially made of reels sent to the vendor. Halfway through the project, it was recognized that the negative reels of film were coming back from the vendor without a significant amount of damage. The decision was made at that time not to make duplicate negatives, but to send out the original negative film to the vendor for scanning. Negative films were transferred to the Harvard Depository after scanning.

### **Digitization**

The scanning of the microfilm was originally outsourced to Preservation Resources in Bethlehem, Pennsylvania, with Preservation Resources working as a sub-contractor to the Digital Imaging Group of the Department of Preservation and Imaging Services. By the end of the first year it was determined that a change of vendors was necessary. Although the quality of the digital images produced was consistent, there were delays in production, inconsistencies in file naming and missing files. The vendor was extremely slow in responding to problems, and despite assurances that work would improve, there was no change. The contract was terminated and a new agreement established with Hudson Micrographics.

600 dpi, bi-tonal TIFF images of the pamphlet pages were produced by the vendor. Page images of pamphlets were organized in directories named with the pamphlet's HOLLIS number. Page numbers for each title and the corresponding image file names were recorded in a tab delimited text file. Implied page numbers were entered in brackets. This information was used by the Digital Imaging Group (DIG) in creating the structural metadata needed to deliver the object through the Page Delivery Service. All two-up page images were split at the gutter into individual page-images.

Meaningful color and grayscale images were captured at the studio camera. If an image could not be interpreted in black and white it was considered for direct digital capture. Pamphlets with pages marked for direct digital capture (color pages, etc.) were pulled after microfilming and sent

to the Digital Imaging Lab for scanning. The color digital image was added to the digital object, replacing the bi-tonal image captured from film.

Keyword searchability of the text on the pages was not a requirement of the LDI project, so OCR of the digitized text was not done in the initial project. At the point the work shifted from project to program in 2005, OCR, having become part of Imaging Services' established workflow for published materials, was incorporated into the work.



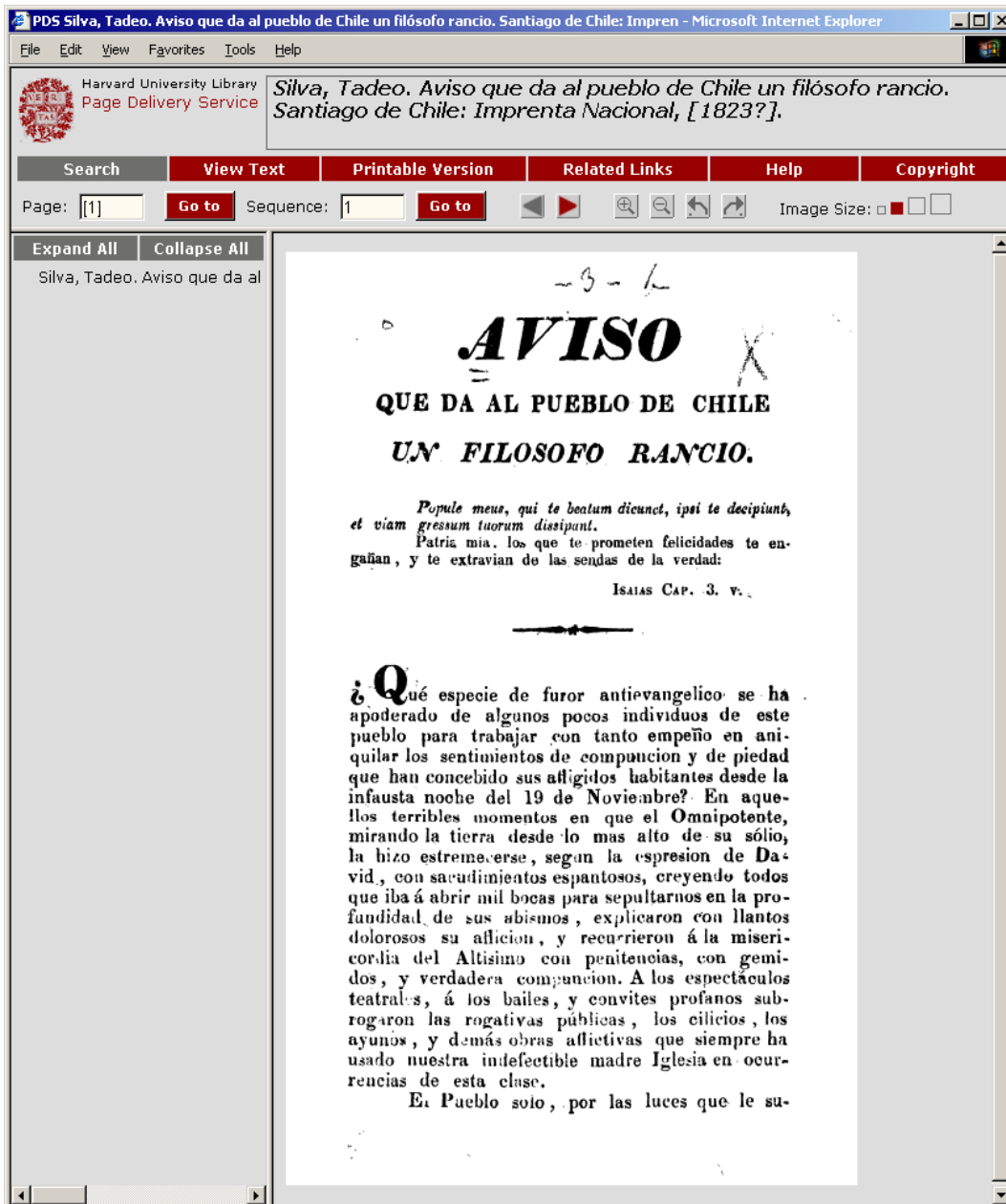
Pages with meaningful color, such as this pamphlet cover, were digitized through direct digital capture at a studio camera. Anderson, T. J. *Forty hours on southern seas to the beautiful land of winter rest.* [Houston?]: Southern Pacific Steamship Lines, [1905].

#### **Processing of digital images created from microfilm**

Digital files on CDs were delivered to the DIG for final processing. Page images only were created. No machine readable and searchable “full text” files were created. Page image files created by DIG through direct scanning (color pages, etc.) were placed in the appropriate directories. Administrative metadata was created based on DIG specifications. XML files containing curatorial information, label information, and structural metadata for delivery of the page images through PDS were created. For quality control purposes, page counts recorded by Reformatting Preparation staff were compared to the number of digital images delivered by the vendor. Additional quality control measures included selective rendering and the use of a validation program. Digital images and accompanying metadata files (MOA2/METS files) were deposited by the Digital Imaging Group into the Digital Repository Service (DRS) for storage.

## Metadata for page turned objects

Metadata needed for label of the page turned object was collected through the use of a program developed by the DIG. The program pulled the label data directly from select fields in the catalog record. This descriptive information, along with the page number information provided by the film scanning vendor was used to create the MOA2/METS structural metadata file needed for the delivery of the pamphlets through the Page Delivery Service (PDS). The majority of the pamphlets are simple objects, with no indication of structure. A small number of pamphlets had a second level of structure defined for better access through the PDS. The program developed by DIG also populated a workflow database used to track the pamphlets through the reformatting process.



View of pamphlet in the Page Delivery Service.

### **Linking digital objects to catalog records**

Once deposit of digital images and accompanying metadata into the DRS was confirmed, catalog records were updated to include a link to the networked resource – the page images in PDS.

### **Quality Control**

Extensive review of the first batches of digital images, including comparison with the original pamphlets, was undertaken at the start of the project to confirm that the product produced from the film scanning was satisfactory and met the desired goals of the project.

Once in production, quality control was performed at the time the link was added to the HOLLIS record. The match between the digital object and the HOLLIS record was confirmed, the total number of pages checked against the record, and select page images reviewed.

### **STORAGE**

Digital images and accompanying metadata files were deposited by the Digital Imaging Group into the Digital Repository Service (DRS) for storage. The Harvard College Library is covering the costs of DRS storage. Widener Library is retaining the original pamphlets at the Harvard Depository.

### **STAFFING**

Richard Lesage was hired as Project Cataloger in 2002 and worked full time on the project for the two and a half years funded by LDI. He has continued to catalog Widener's Latin American Pamphlets in Imaging Services half time on HCL funding. Richard was assisted in the cataloging of the pamphlets by Preservation Cataloger Katherine Sweeney, who reviewed all records for the Latin American Pamphlets. Amy Jackson performed the Quality Control review of the pamphlets in the Page Delivery Service and added the links to the HOLLIS records. Todd Bachmann managed the production workflow. Maggie Hale provided assistance in project planning, management and reporting throughout the project. Dan Hazen was the Project Manager. As Dan's position changed from Librarian for Latin America, Spain, and Portugal to Associate Librarian for Collection Development in the Harvard College Library, Lynn Shirey's involvement in the project increased.

### **PROJECT SCHEDULE**

The project was initially planned as a two and one half year project, starting in July of 2002. Steady progress was made in cataloging and the cataloging remained on schedule throughout the project, with over 3,000 pamphlets cataloged by December 2004.

There were delays in reformatting at several points in the project. There were delays in microfilm digitization early in the project due to problems with the original vendor; this was finally resolved by selecting a new vendor. Delays in digitization in 2004 resulted from slow production

rates in DIG for items requiring direct digital capture. New equipment was added in August 2004 and rates improved. Additional delays in reformatting resulted from flooding in the microfilm labs in August 2004. This stopped production for three to four weeks. Digitization and linking of the initial 3,000 pamphlets in the project was not completed until March 2005.

Harvard College Library funding has allowed for the continued digitization of Latin American Pamphlets. By October 2006, an additional 1,500 pamphlets had been cataloged and digitized.

The Latin American Pamphlet Virtual Collection was completed and made public with approximately 3500 records in April of 2006. Additional records were added in November of 2006, bringing the total in the Virtual Collection to 4,512.

## **BUDGET**

Project Cataloger (2 ½ years FTE): \$XXXX salary and fringe.

Digital Projects Librarian, HCL project development and coordination: \$41,773.

HCL Conservation, materials preparation services: \$4,200.

HCL-DIG Imaging Services including microfilming, scanning creation of structural metadata and deposit to DRS: \$97,800.16

**Total expenditures: \$274,173.13**

(Project cap: \$286,000)

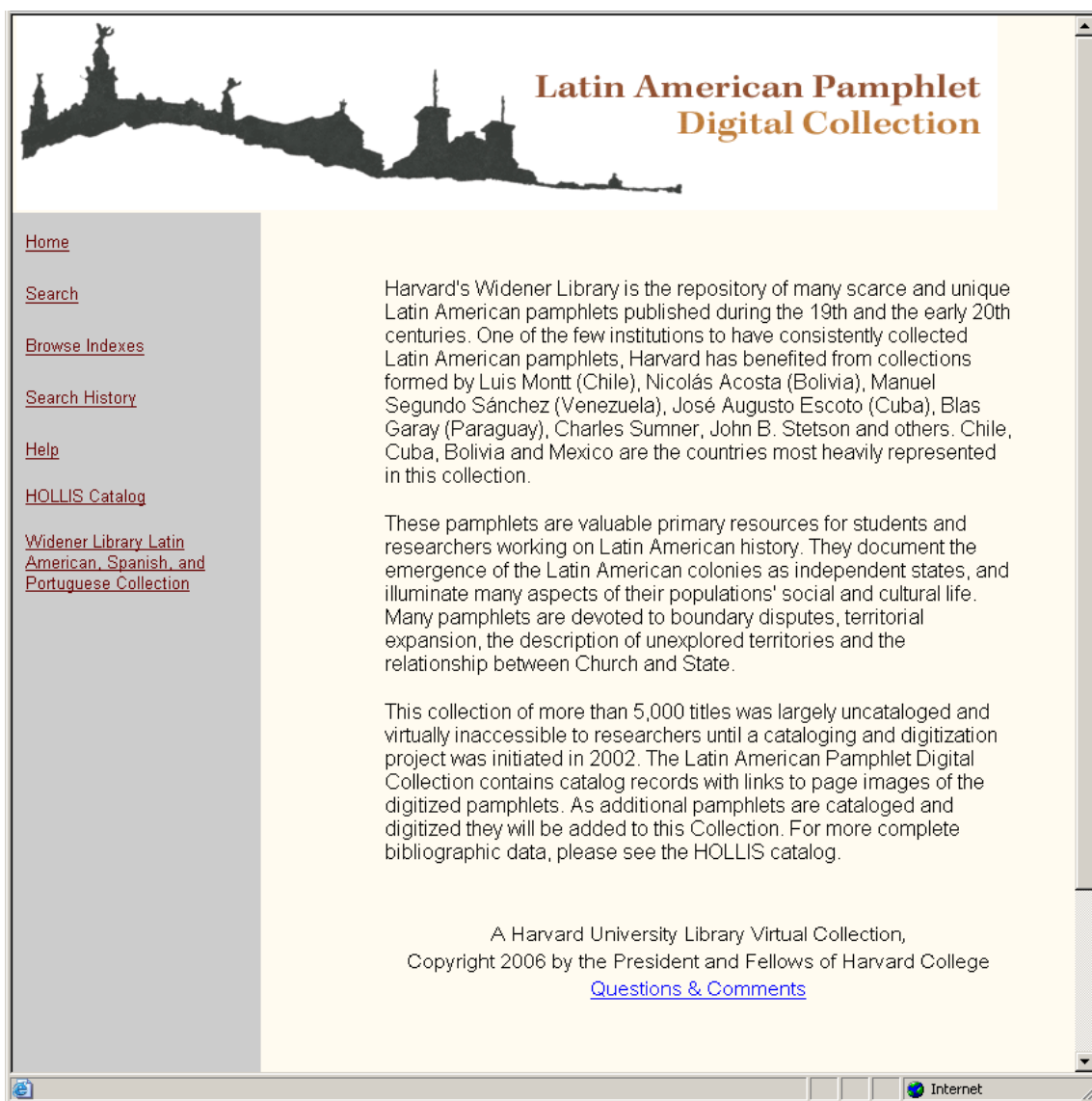
Review of average number of pages per pamphlet was reviewed in November 2003 when figures were available for approximately 1200 of the pamphlets in the project. The page average per pamphlet appeared higher than the page count used in project planning and budgeting (forty pages per pamphlet). Figures indicated an average of fifty-two pages per pamphlet. The number of direct digital captures of illustrations and/or color images was also higher than anticipated. At the time HCL committed to providing the additional funding that may be required to complete the project. Additional funding from HCL was not needed to complete the original 3,000, as average page counts dropped, however committed funding was used to continue the work beyond the original 3,000 pamphlets.

## **ACCESS AND USE**

As the project began, there was concern that HOLLIS, with some ten million records, would inevitably defeat unwary users poking around after our 3,000 pamphlets. There was a need for a mechanism that would permit users to identify easily the specific materials subsumed within the project. Various options were discussed during the course of the project, including the development of a web page that would provide a general description of the project, a description of the type of materials included (potentially organized by country), and how to locate the pamphlets in HOLLIS.

As the Latin American Pamphlet project was reaching its end, HUL OIS development of the Virtual Collections service was underway. Virtual Collections is a service that can harvest descriptions and links from HUL union catalogs and provide a web-based search and display interface of these materials for the user. Recognizing that this service met the initial need for easy access to the Latin American pamphlets expressed by Dan Hazen when this project was first proposed, OIS used the Latin American Pamphlets as one of two pilot projects for the Virtual Collections service.

In addition to accessing the pamphlets through HOLLIS catalog, users can now search only the records for digitized Latin American Pamphlets through the “Latin American Pamphlet Digital Collection” Virtual Collection at: <http://nrs.harvard.edu/urn-3:hul.eresource:latampdc> . This resource provides catalog information harvested from the HOLLIS records, including links to page images of the digitized pamphlets delivery through the Page Delivery Service.



Home page for the “Latin American Pamphlet Digital Collection” Virtual Collection.

Dan Hazen and Lynn Shirey have promoted the project on a number of listserves and spoken about it at meetings and conferences. Lynn has advertised it via the SALALM listserv (Seminar on the Acquisition of Latin American Library Materials); the LANE listserv (Latin America Northeast library consortium); and has spoken about the collection at meetings of each of those groups. She has demonstrated the Latin American Pamphlet Digital Collection Virtual Collection at orientation sessions for visiting scholars at the Rockefeller Center for Latin American Studies. Lynn will also be contacting the Biblioteca Nacional de Chile to inform them of the Latin American Pamphlet Digitization Project/Program, and the online availability of pamphlets containing manuscript annotations by Luis Montt, bibliographer and director of Chile's national library between 1886 and 1909.

Recent examples of use of the collection include a visit to Widener Library by Prof. Josep M. Barnadas, a scholar and bibliographer from Bolivia, who requested to consult pamphlets belonging to Nicolás Acosta. Prof. Barnadas was enthusiastic about the cataloging of the Bolivian pamphlets, and the ability to view the page images through the web. There has also been interest from an American researcher in the pamphlets formerly owned by Luis Montt.

The Latin American Pamphlet Digital Collection is currently accessible through the Harvard College Web Page, as an e-resource, and from the Digital Collections Highlights section of the Harvard Libraries Site (see: <http://digitalcollections.harvard.edu/>).



## APPENDIX I

### LAP 4250 statistics (grand total).

The total number of items processed = 4958  
broken down as follows 3806 (number originally selected for the project) + 1152 (number added in course of project)

The number of missing items = 103 This does not necessarily mean that all of these have been lost; some items might have been reclassified without proper annotation in HOLLIS; some items might be in the group of pamphlets recently retrieved from NEDL)

The bib records created in HOLLIS equal 2070 (1505 new original records + 565 records exported from OCLC)

The original/extensive cataloging amounts to 2944 records ( 1505 new original + 1439 extensively upgraded stub records) That amounts to 69.3% of the pamphlets.

The cataloging statistics do not include routine clean up work done on records for works not included in the LAP Project

213 holdings records were added to existing bibliographic records

73 duplicate records were deleted

Some 275 authority records were created (all authority records created are kept in a folder). The total number of records counted in column T in the various spread sheets indicate the number of authority records to be created. If we subtract the 275 already created we will have an idea of the authority record backlog.

651 items were not filmed because they were duplicates within the project, out of scope, fragmentary, under copyright, or had more than 100 pages.

There were 195 duplicate pamphlets (i.e. duplicates within the project)

202 copies were not digitized because of copyright. Some of them were filmed for preservation.

Items filmed separately for preservation only (i.e. filmed but not digitized) equal 85

## Appendix II

### Latin American Pamphlet Project Workflow

#### Inventory of pamphlets (Reformatting Preparation Unit)

- Pamphlets will be processed by country, in the following order: Argentina, Bolivia, Chile, Cuba, Mexico, Peru, Uruguay, Venezuela, others.
- Going in call number order, pull 100 volumes at a time to review.
  - Confirm that volumes from the preliminary inventory are on the shelves
  - Verify call numbers on piece/s
  - Verify numbers of pieces within pamphlet volumes, boxes, and envelopes
  - Determine whether it is necessary to disbind volume for microfilming; flag volumes to be disbound
  - Identify and flag materials not in the public domain
  - Identify and flag materials with more than 150 pages
  - Collate: examine page by page for completeness, and to identify features requiring additional attention in microfilming and/or scanning (e.g., oversized or color illustrations).
  - Address any immediate physical concerns
  - Identify any ephemeral materials; flag for review by bibliographer
  - Re-house all loose pamphlets in acid-free folders
  - Place all materials on shelf for cataloging

#### Tracking of pamphlets (Reformatting Preparation Unit)

- Set up spreadsheet for tracking pamphlets
- Take pamphlet group to be cataloged (i.e. one call number for pamphlets bound in a volume, or loose in a box, or envelope/s), and search in HOLLIS bibliographic records for particular call#
- For volumes to be disbound, and loose group of pamphlets proceed as follows:
  - If collective bib record exists, charge volume to Reformatting Prep., by creating item record, and assigning dummy “barcode” if need be
  - If collective bib record does not exist, create record, and charge materials as above
  - For groups of less than three loose pamphlets do not create group record, but charge each item individually upon cataloging
- For volumes not to be disbound proceed as follows:
  - If collective bib record exists, charge volume to Reformatting Prep. as above (All items in volume will be linked to collective record, by means of H77 in their holdings records)
  - If collective bib record does not exist, do not create one, but charge first pamphlet in volume to Reformatting Prep. by creating item record, and affixing real barcode to volume. (All remaining items in volume, will be linked to first item in volume by means of H77 field in their holdings record)

### **Cataloging of pamphlets (Reformatting Preparation Unit)**

- Determine whether each piece qualifies for project (considering content)
- Search titles for each pamphlet, before cataloging (according to searching procedures of Reform. Prep., not detailed here)
- Catalog pamphlet (according to cataloging practice of Reform. Prep., not detailed here)
- Perform all necessary authority work pertaining to bibliographic records
- Populate spreadsheet with cataloging data
- Place materials on shelf for editing
- Edit bibliographic records
- Place materials on shelf for reelization
- Reel programming (determine pamphlets to appear on reels). Assign series and master negative numbers.
- Flags with sequence numbers placed in pamphlets.
- Pamphlets in need of disbinding delivered to Collections Conservation, approximately 40 pamphlets (1 reel) to 120 pamphlets (3 reels) at a time. (This number will be determined as project gets under way.)

### **Pre-filming conservation treatment of pamphlets (Collections Conservation)**

- Pamphlets received in lab and treated. Treatment of materials sent to lab:  
Conservators will disbind as necessary.  
Disbound pamphlets will be placed into individual slings.
- Pamphlets sent back to reformatting Preparation Unit.

### **Final preparation of pamphlets for microfilming (Reformatting Preparation Unit)**

- Receive pamphlets back from Collections Conservation.
- Review sequence; verify materials are in correct order.
- Create and print out targets for each pamphlet and place with pamphlet.
- Save a file containing records for all titles on the reel. File will be saved from the OPAC in MARC Tags format. File will be renamed with the master film #. File will be saved in unicode font as a text file in H: \P\_and\_I\latin\_pamphlet\ (The average file will contain 30 to 50 pamphlets.) The Digital Imaging Group (DIG) will run weekly scripts on the H: \P\_and\_I\latin\_pamphlet\ and generates node

files and directory structure necessary for the XML structural metadata file used by the Page Delivery System (PDS). See **Appendix: “Capturing information for MOAII file and workflow database”**

- Send materials to Microfilming Group. Materials will be accompanied by flags indicating sequence number on reel, flags marking features requiring additional attention in filming and/or scanning, and targets. Three reels worth of materials will be sent at a time (batch size may be adjusted during project).

#### **Microfilming of pamphlets (Microfilming Unit)**

- Microfilm pamphlets.
- Microfilm inspected.
- Extra duplicate negative made for scanning.
- Microfilm packaged and sent to vendor. (Microfilming Unit will package film, Digital Projects Librarian will send out by Fed Ex.)
- Pamphlets returned to Reformatting Preparation Unit.

#### **Post-filming processing of pamphlets (Reformatting Preparation Unit)**

- Pamphlets stamped to indicate they have been filmed.
- Pamphlets with pages flagged for direct scanning (color, etc.) pulled and sent to Digital Imaging Lab. Each pamphlet should be accompanied by a printout of the item’s MARC record with the HOLLIS ID highlighted and the item’s Film Master number noted. Flags should mark the pages identified for direct digital capture. Each batch of pamphlets should be also be accompanied by two copies of an inventory of the physical items delivered to DIG in the batch. (Both copies will be signed upon receipt of materials. DIG will keep a copy, and one copy will be kept by Reformatting Prep. When materials are picked up, DIG’s copy must be signed by the recipient of the materials and the date noted.)
- Pamphlets are sent to Conservation for post-filming treatment.

#### **Post-filming conservation treatment of pamphlets (Collections Conservation)**

- Pamphlets repaired as necessary and placed in final housing:
  - Repairs will be minimal.
  - Final housing will be individual envelopes.
  - Volumes that were not disbound will be placed in boxes as necessary.
- Item records for pamphlets will be created and pamphlets will be barcoded.
- Pamphlets sent to HD Transfer.

### **Digitizing of microfilm (Preservation Resources sub-contracting with the Digital Imaging Group)**

- Microfilm received by vendor.
- Microfilm scanned.
- Two-up page images split at gutter. Page images cropped.
- Page images placed into directories named with HOLLIS ID number.
- Tab-delimited text file created for each title, with page numbers and corresponding image file names.
- Files written to CD.
- Return CDs and microfilm to DIG.

### **Processing of digital images created from microfilm (Digital Imaging Group)**

- Microfilm and digital images received from Preservation Resources.
- Digital images processed.
- XML Structural metadata file (MOA2/METS) created using information gathered through DIG's Data Collection Tool, pagination information supplied by Preservation Resources, and data directly entered by DIG.
- Image files w/administrative metadata and XML structural metadata file (MOA2/METS) deposited to DRS.
- Confirmation reports received, Reformatting Preparation Unit notified of deposit and confirmation.

### **Quality Control and linking digital objects to catalog records (Reformatting Preparation Unit)**

- DRS Deposit report received in Reformatting Prep Unit.
- From DRS Deposit Report, spreadsheet created with URN and Owner Supplied names (HOLLIS System #) for all pamphlets in batch.
- Add link to HOLLIS record and perform quality control step as follows, one pamphlet at a time:  
Add link to HOLLIS Record:
  1. Using System # copied from spreadsheet, pull up record in ALEPH and enter cataloging mode.
  2. Add NET holding. Open template created for the NET holdings in the project. Copy the URN from the spreadsheet and paste into the template.

3. Save changes. Close record.

Perform Quality Control steps:

1. Copy the HOLLIS System # from the spreadsheet, and use it to search the OPAC and pull up the HOLLIS record for the pamphlet. Use link to bring up the digital object.
2. Check for match between the title and author in the HOLLIS record and the title and author in the label text for the pamphlet in PDS.
3. Check for match between label text and the title page of the pamphlet.
4. Check the first few pages of the pamphlet, until standard page numbering begins.
5. Check every 10 pages of the pamphlet, confirming that page numbers are correct (PDS page number matches number on page image). Also check for legibility, skew, and other problems with pages.
6. Check last few pages of digital objects. Confirm that the digital object appears to end with a "last" page.
7. If problems are noticed in the above review, continue to review more pages of pamphlet.
8. Note QC issues in project database.

### **Additional tasks (Bibliographer and Digital Projects Librarian)**

- Collect complete cost data for the full range of project activities (ongoing).
- Invoicing (ongoing).
- Project reporting (ongoing).
- Project publicity.
- Analysis and evaluation of project.

### **Workflow Appendix:**

#### **Capturing information for MOAII file and for workflow database**

- Following the assignment of a master file # to a group of pamphlets, the Reformatting Preparation Unit will save a file containing records for all titles on the reel. File will be saved from the OPAC in MARC Tags format. File will be renamed with the master film #. File will be saved in unicode font as a text file in H: \P\_and\_I\latin\_pamphlet\ (The average file will contain 30 to 50 pamphlets.)

- DIG runs weekly scripts on the H: \P\_and\_Mlatin\_pamphlet\ and generates node files and directory structure.
- Information captured for the MOAII file as follows. If any of the fields included in the file is unpopulated in the MARC record, leave blank.

Pamphlets will be simple objects.

**Item Control Number** will not be used.

**Object Type:** monograph

**Object Display Label** will appear as in the following order, with punctuation as indicated:  
 Creator, title. Place of publication : publisher, date.

Example: **Viale, César, Congreso pro-creación Consejo Nacional de Colonización. Buenos Aires: Impr. N. Spinelli, 1934.**

This is a concatenation of the following MARC fields (see highlighted fields in the sample record below):

- Personal Creator: 100, subfield a; **or** Corporate Creator: 110, subfield a; **or** Conference/Meeting Creator: 111, subfield a (*if neither the 100, the 110, or the 111 field is populated, the title will be the first element in the label*)
- Title: 245, subfield a (*add a period at the end of the title, unless there is no subfield b or c, which case the period will be already there*)
- Place of Publication: 260, subfield a
- Publisher Name: 260, subfield b (*there should be one space between the colon following place of publication and the publisher name*)
- Date of Publication: 260, subfield c (*followed by a period*)

[Note: the page number, which will be following the date of publication in the label, should not be preceded by any punctuation.]

**Object Title:** 245, subfield a and subfield b

**Personal Creator:** 100, subfield a

**Corporate Creator:** 110, subfield a

**Date:** 260, subfield c

**ALEPH ID #:** SYS

**Place of Publication:** 260, subfield a

**Publisher Name:** 260, subfield b

Sample record: Fields for use in MOII file are highlighted.

```
FMT          BK
LDR          00552nam 2200169Ka 4500
```

001	007102513-8
005	20020906120016.0
008	960829s1934 ag 000 0 spa d
040	a HLS  c HLS  d HMM
1001	a <b>Viale, César,</b>  d 1881-
24510	a <b>Congreso pro-creación Consejo Nacional de Colonización</b> :  b discurso de clausura de las sesiones deliberativas /  c por el Sr. Vice-Presidente del comite ejecutivo César Viale, Buenos Aires, 7 de julio de 1934.
260	a <b>Buenos Aires :</b>  b <b>Impr. N. Spinelli,</b>  c <b>1934.</b>
300	a 12 p. ;  c 18 cm.
61010	a Consejo Nacional de Colonización (Argentina)
71010	a Consejo Nacional de Colonización (Argentina)
830 0	5 MMF  a Latin American pamphlet digital project at Harvard University ;  v 0030.
CAT	a CONV  b 00  c 20020608  l HVD01  h 0755
CAT	c 20020722  l HVD01  h 0229
CAT	a WIDLESAGE  b 50  c 20020801  l HVD01  h 1501
CAT	a WIDSWEENE  b 50  c 20020806  l HVD01  h 1509
CAT	a WIDSWEENE  b 50  c 20020806  l HVD01  h 1522
CAT	a WIDKERN  b 50  c 20020814  l HVD01  h 1626
CAT	a WIDJACKSO  b 50  c 20020906  l HVD01  h 1200
H48	a ORK1020ASWID2Box03512  b 19960829  c 19960830
H016	a BIF6808
H03	a MHBIF68086HU
PST7	0 HOL  1 HVD60-009014192  n 7  2 ZHCL  b WID  c GEN  h SA 5003.10  4 Widener
PST8	0 HOL  1 HVD60-011033379  n 8  b MMF  c GEN  h Film Mas 32653  4 Master Microforms
PST8	0 HOL  1 HVD60-011033383  n 8  b MIC  c GEN  h Film W 17667  z In process  4 Microforms (Lamont)
<b>SYS</b>	<b>007102513</b>

- A workflow database will be maintained for the project. This will be used for tracking following the creation of the microfilm. DIG will run a script to populate the workflow database from this file, using the System ID # (SYS); the Film Master number (PST 8, subfield b MMF and subfield h); and the Series number (830, subfield v).

Fields used are underlined in the sample record below:

FMT	BK
LDR	00552nam 2200169Ka 4500
001	007102513-8
005	20020906120016.0
008	960829s1934 ag 000 0 spa d
040	a HLS  c HLS  d HMM
1001	a <u>Viale, César,</u>  d 1881-
24510	a <u>Congreso pro-creación Consejo Nacional de</u> <u>Colonización :</u>  b discurso de clausura de las sesiones deliberativas /  c por el Sr. Vice-Presidente del comite ejecutivo César Viale, Buenos Aires, 7 de julio de 1934.
260	a <u>Buenos Aires :</u>  b <u>Impr. N. Spinelli,</u>  c <u>1934.</u>
300	a 12 p. ;  c 18 cm.
61010	a Consejo Nacional de Colonización (Argentina)
71010	a Consejo Nacional de Colonización (Argentina)

830 0 |5 MMF |a Latin American pamphlet digital project at  
 Harvard University ; |v 0030.  
 CAT |a CONV |b 00 |c 20020608 |l HVD01 |h 0755  
 CAT |c 20020722 |l HVD01 |h 0229  
 CAT |a WIDLISAGE |b 50 |c 20020801 |l HVD01 |h 1501  
 CAT |a WIDSWEENE |b 50 |c 20020806 |l HVD01 |h 1509  
 CAT |a WIDSWEENE |b 50 |c 20020806 |l HVD01 |h 1522  
 CAT |a WIDKERN |b 50 |c 20020814 |l HVD01 |h 1626  
 CAT |a WIDJACKSO |b 50 |c 20020906 |l HVD01 |h 1200  
 H48 |a ORK1020ASWID2Box03512 |b 19960829 |c 19960830  
 H016 |a BIF6808  
 H03 |a MHBIF68086HU  
 PST7 |0 HOL |1 HVD60-009014192 |n 7 |2 ZHCL |b WID |c GEN |h  
 SA 5003.10 |4 Widener  
PST8 |0 HOL |1 HVD60-011033379 |n 8 |b MMF |c GEN |h Film Mas  
32653 |4 Master Microforms  
 PST8 |0 HOL |1 HVD60-011033383 |n 8 |b MIC |c GEN |h Film W  
 17667 |z In process |4 Microforms (Lamont)  
SYS 007102513