Assessing, Organizing, and Indexing Digitized Historical Documents: Constructing a Digital Library and a New Approach to Historical Research

Benjamin Moore

Follow this and additional works at: https://digitalcommons.xula.edu/xulanexus

Recommended Citation
Available at: https://digitalcommons.xula.edu/xulanexus/vol7/iss2/4

This Article is brought to you for free and open access by XULA Digital Commons. It has been accepted for inclusion in XULAnexUS by an authorized editor of XULA Digital Commons. For more information, please contact ksiddell@xula.edu.
Assessing, Organizing, and Indexing Digitized Historical Documents: Constructing a Digital Library and a New Approach to Historical Research

Benjamin Moore, Physics and Spanish
Catherin Morales, Psychology, Pre-Medicine

Faculty Mentor: Dr. Elizabeth Manley, History

Abstract

Over a period of several semesters, Benjamin Moore and Catherin Morales have converted, indexed, compiled, and summarized the beginnings of 10,000 photographs of historical material in order to provide the start of a searchable database of pertinent documents detailing the lives and struggles of women in 20th century politics of the Dominican Republic, as a part of Dr. Elizabeth Manley’s continued research into that time and place. The work consisted of identifying the best way to build the library, building it, and indexing and organizing the materials involved, with an eye toward expanding it in the future. Thus far, Moore worked mainly on these first two tasks, identifying a means to convert the documents to Portable Document Format and a cataloging program that would fit the future library's needs. Morales worked on the interpretation of data, a task Moore may be helping with in the future, considering the size of the library. Both researchers will likely be involved in the organization and integration of documents into the library as it grows. This scholarly note describes each student’s research process, as well as the lessons they have learned from the ongoing project.

Key Terms: Archival Research, Database, Document, Dominican Republic, Digital Library, History

Notes from Benjamin Moore

When I first learned about the project, I was not sure how I could help, and frankly was concerned this would require skills I did not have. After all, I had never done historical research, I had never physically entered a national archive, and I barely had heard of Dominican dictator Rafael Trujillo. To my pleasant surprise, what I was required to do was purely technological. Given roughly 10,000 images in jpeg format, I was to convert these into PDF, assemble them into documents of their appropriate length, and organize them into a searchable database.

Initially, I surveyed all programs that could convert a JPEG image to a PDF file. Because the operating system was Macintosh OSX, the options were limited to programs that would run on that system. Fortunately, OSX natively offers the ability to “print” a document in PDF format. Unfortunately, the file size increases far too much to consider 10,000 such conversions. After an exhaustive search of the options on the internet, I discovered a time-limited freeware conversion program called Dragoman. Within the 15-day trial limit, I easily was able to convert all of the files through a painless 1-click batch process.
To combine the files, I consider using Automator, an OSX native program that allows users to write simple programs by assembling a workflow from a collection of available commands. In the end, though, I ended up using Preview, because Automator, while useful, was not an efficient choice for the amount of files to be combined, and the combining program would have to be rewritten for each run. In Preview, it is possible to drag and drop single-page documents together to create a multi-page one.

After another exhaustive Internet search, FileMaker Pro was chosen as the database software. It is the database software of choice of several institutions, very user-friendly, and highly customizable. I could find no better alternative in the open-source software community. I am now exploring the options with it and find it to be a versatile and powerful, if stubborn program.

When this project was inaugurated, I expected that I would be in a library, physical or digital, running searches or sorting through an ancient card catalog. Instead I tried out programs, compared efficiencies, and used the best tool to solve the problem given. It has been a pleasure to work on the project, and I am pleased to have learned that the skills I acquired in departments far from history, and in my personal life while computer hacking, are useful to research. I will look for and encourage others to look for faculty-student research opportunities. Such exchange offers valuable insight into the application of skills often not valued or recognized by those who possess them.

Historically, I suspect the project will bring to light facets of life that have been hidden in the Dominican Republic’s national archives, and that a statistical correlation of the incoming letters (documents) with their sources and/or their reply rates will reveal previously unrecognized or unconfirmed patterns of governmental attention to or ignorance of the needs of the people.

Notes from Catherin Morales

Being involved with this project has given me first hand experience in the research process. It has been interesting to take part in history, especially researching about the role of women in the 20th century in the Dominican Republic. Participating as an undergraduate helped me better understand life in academia, and to see what type work I might be doing in graduate school or a future career. I also have been able to work closely with faculty members and to develop extraordinary skills of historical research.

I was hired in my freshman year for the Department of History as a student researcher. My contribution to the project is identifying key issues in the historical documents. I read through each one of the 10,000 digital files existent in the catalog. The documents range from letters asking for resources to be sent to various provinces from regional governors, to requests from citizens that certain laws be enforced, usually with regards to the payment of alimony after a divorce. By a careful examination of the Spanish language documents, I scan for key places, important names, and dates. In addition, I write a brief summary of the significant events in the documents. Notably, all documents that reference Trujillo attach to his name the plethora of honorifics typical of a dictator’s correspondence. Many of the documents are letters that public offices and common people sent to state institutions. In those letters, I found many descriptive events concerning women’s contribution to the country. I scan these for names of women who
marked history through their actions in Dominican politics and society. I also read newspaper articles to understand the historical context in which the documents were created and include them in the database. Overall, my part in the project, reading over the documents, is the first step that will lead to building a digital library of historical materials.

I consider the project highly beneficial to my own research training. By reading the materials, I am becoming knowledgeable in Latin American history and more concerned about the importance of women in the 20th century. With this historical project or database, I believe that many students will reap the benefits for future research. The documents I have examined are highly informative and will make students aware of the significance of women’s roles in politics and their contribution to a better society in the Dominican Republic.

The experience has been splendid and educational. Although this was my first time participating in such research, I learned about significant historical issues involving a number of different women. This project clearly will contribute to the generation of knowledge and understanding among many students, as it did for me when I read and assessed the documents.

The project is, of course, ongoing, and so its future will see the completion of the current library, and hopefully its expansion via the addition of more primary source documents. The project may also involve the creation of a web-based interface to open access to other researchers. We look forward to bringing this information from a single archive room that is difficult to access to an easily searchable database available anywhere the Internet can be reached.