

Description and Rationale

Title

State of Baja California Human Rights Commission Archives Case Digitization Project Phase 1/
Proyecto de Digitalización de Casos del Archivo de la Comisión Estatal de los Derechos
Humanos de Baja California Fase 1

Project Description*

After a successful pilot during the summer of 2017, the University of San Diego (USD) - Copley Library will digitize the case backlog on the Fall 2020/Spring 2021 destruction schedule. Cases go as far back as the 1990s before there was a Comisión Estatal de los Derechos Humanos de Baja California (CEDH). These cases hold information on the types of abuses that were filed during that time along the Baja California/California border. The data in these cases, many of which were terminated, closed or dismissed before full investigations were completed, will provide a snapshot of the region for border scholars and historians alike. The goal of this project is to eventually make all of these older cases available for research and data mining online via DigitalUSD, USD's Institutional Repository. This proposal is asking for funding for Phase 1 to digitize all of the material due for destruction within the next fiscal year 2020-2021.

Reasons for Consideration*

Border scholars on the border region of Mexico and the United States are interested in data on human rights violations. Border scholars in the San Diego region have expressed interest in the possibility of data mining Human Rights cases—except these haven't been digitized. Only the recommendations (cases that make it to the National Human Rights Commission) are digitized, and only the most recent recommendations have been anonymized. The CEDH human rights lawyers have also expressed interest to see past cases digitized. Currently, they only have access to the details of these cases in a minimal format: case number and case resolution. Digitization and anonymization of the data would allow researchers and citizens alike to learn the details of these cases.

Description of the project/Background of the project:

During the spring of 2017, the director of the USD TransBorder Institute (TBI), Dr. Everard Meade expressed interest in the cases kept in boxes in the archives of the CEDH. He had already approached the then CEDH President, Licda. Melba Adriana Olvera Rodríguez, and explained the value these cases had for not only their own Human Rights lawyers but to border scholars and Human Rights researchers. Dr. Meade knew he had to involve the library because he lacked any experience with large digitization projects. We explained to him that this type of project was more than simply scanning and uploading to a server. The first step was arriving to an agreement with the CEDH and gaining access to the archives to evaluate how much material there was and in what condition. While the CEDH's headquarters are located in Tijuana, the archives are located within the Human Rights Branch Office in Tecate. During the summer of 2017, the University of San Diego (USD) - Copley Library ran a pilot to digitize the case backlog scheduled for removal for the smallest state municipality, Playas de Rosarito. Although the pilot was successful and a copy of all of the digitized materials was given in February of 2018 to the lead complaints lawyer, the project stagnated due to sabbaticals and personnel changes at both institutions during 2018 and the beginning of 2019.

By summer of 2019, the CEDH had hired its first official Archives Manager, Lic. Miguel Ángel Lerma under the new CEDH president, Lic. Miguel Ángel Mora Marrufo. Before Mr. Lerma's arrival during the summer of 2019, there was only a part-time employee who stacked cases and filed them, as best she could, while also receiving new cases to file twice a week. Mrs. Lerma continued her work and improved upon the reorganizing that I employed with my team during the pilot in summer of 2017. The cases continue to be organized by Municipality, then by Year, and no longer simply by Year. Mr. Lerma in his time at the CEDH's Archives has successfully separated the CEDH's financial records and old training manuals from the cases. The cases are no longer interspersed all over the limited space available in the one room archives.

Throughout 2018, I kept in touch with the Tecate Lead Human Rights Abgda. Laura Gutierrez. She connected me to the new Director of Complaints and Orientation, Licda. Margarita Suárez Rodríguez, who had been using the digitized and OCR'ed cases and wanted to inform me that they were still interested in setting up a collaboration between USD to continue with the digitization project. The new incoming CEDH President is open to having all of the old cases digitized for everyone to learn about the type of work they do at Baja California's CEDH.

Scope and content:

Scope of the project: The goal of this project is to make available Human Rights violation cases filed from early 1990s to the early 2000s with the CEDH Complaints units all over the state of Baja California for research and data mining online via the University's Institutional Repository, DigitalUSD. To accomplish this goal, this project will digitize, create metadata descriptions for the digitized files, anonymize personal data in these files, process and upload the content to the Institutional Repository to make it accessible as an Open Access collection. This proposal focuses on Phase 1 of this project: Assistance with digitizing these cases in collaboration with the CEDH of Baja California.

Description of content: These human rights cases go as far back as the 1990s before there was a Human Rights State Commission for Baja California. Cases then were created by the Procuraduría Estatal de Baja California. The information held in these cases ranging from: depriving a person of their freedom to discrimination against women, children, the elderly, the disabled, and indigenous peoples, to violence against members of LGBTTTIQ community and Human Rights activists to human trafficking and the disappeared paint a picture of the situation regarding human rights violations that were filed during the period of the 1990s through 2010 along the Baja California/California border. The data in these cases, many of which were terminated before full investigations were completed, will provide a snapshot of the region for border scholars and historians alike besides assisting the interested CEDHs own lawyers gain a better understanding of how case reporting has evolved.

Access

DigitalUSD the university's institutional repository is hosted on bepress. All the files uploaded for this project will be accessible as Open Access.

The CEDH is the institution responsible for the custody of these cases and all of their pertaining information. They are the ones with the authority to grant permission to digitize and make the cases available to the public, as long as USD removes all personal information on cases by anonymizing the records before they are ingested into DigitalUSD. (See their attached letter of support for this project)

Metadata as applicable

- The metadata scheme to be used to describe digitized files will be Dublin Core:

- Creator: Comisión Estatal de los Derechos Humanos Baja California
- Language: Spanish
- Subject: Human Rights Violations including Visitaduría Classification 1-5 (types of Human Rights abuse/s) (See Appendix A for classifications)
- Coverage: Ensenada, Mexicali, Tecate, Tijuana, Playas de Rosarito, Baja California
- Type: Text
- Format: Text
- Date: YYYY-MM-DD
- Creator: Comisión Estatal de los Derechos Humanos Baja California
- Contributor: LARRP
- Publisher: University of San Diego
- Rights: Comisión Estatal de los Derechos Humanos Baja California

I will be working with an assistant from the Archives and Special Collections Digital Initiatives and graduate student to apply the metadata in the pre- and post-production of this part of the project, Phase 3.

Archiving/Plans for long-term sustainability and stewardship:

The cases are all paper based, approximately 75% of these are on 8-1/2 x 11 and the 25% is on legal-size paper. Over ninety-five percent of the material is in Spanish. The remaining material is in indigenous languages or English. All cases are placed on a retention schedule and will be destroyed by CEDH as scheduled. The cases that we at USD are seeking funding to digitize are already on a destruction schedule due to space issues. CEDH has indicated that they are open to holding the cases longer in order to have them digitized now that they have seen the value of using internally the digitized cases for Playas de Rosarito. They do not have a budget for digitization, they work with a minimal budget and a very dedicated team of lawyers and staff. The CEDH is not permitted by law to donate the physical materials to any foreign entity.

Amanda Makula, the Digital Initiatives librarian at USD, stated that bepress (bp) follows a comprehensive preservation approach, with persistent URLs, and a robust infrastructure with dedicated staff to manage it 24/7/365. The Digital Commons maintains multiple backups and bp storage is also LOCKSS compliant. Additionally, USD university's archives also has its own dark archive through campus Information Technology Services where multiple copies of this collection will also be kept.

Work Plan, including (where relevant)

Principal Investigator(s) and institutional information:

Alma Ortega, PhD from the University of San Diego's Copley Library. USD is a private Catholic university located in the city of San Diego in California, 21 miles from the Mexican border.

Project Phases:

Phase 1: Digitization and OCR [est. time frame] 4 months with 2 two digital lab technicians working 30 hours/week

Phase 2: Anonymization of data [est. time frame] six to eight months

Phase 3: Metadata description [est. time frame] six months

Phase 4: Processing and upload content to IR [est. time frame] 3 months

Phase 5: Public Launching of project [est. time frame] 2 months

In **Phase 1** of the project, cases will be organized by case size and length to prepare for digitization. Each assistant and I will be processing and digitizing one municipality at a time, except for the Municipality of Tijuana. This municipality is the largest and by far has the most

cases thus a larger team is needed to digitize its cases. Each case will be separated either by removing staples, paper clips, notes, etc., then the case will be digitized. Notes will also be digitized and placed back in the order they were found within the case file. Cases will be placed in new boxes to keep track and free up space until these are to be destroyed. Once cases have been digitized these will be OCRed for ease of use to library staff redacting sensitive information during the anonymization phase and for researchers when topic searching in the final product. With a team of three people, two of which will be working 30 hours a week, it will be feasible to scan the vast majority of the cases, if not all, in 16 weeks.

For **Phase 2** a copy of each PDF file will be made, this copy is the one that will be redacted using Adobe Acrobat Pro DC. Free options were not explored because these tend to change or lose support once a new free software is released. For consistency Adobe Pro will be used because it can remove data, not just mask it. These cases need to have the sensitive data removed permanently before uploading to the institutional repository, otherwise victims, lawyers and judges involved in the cases may be endangered. The CEDH and USD are committed to protecting the identity of all the people involved in these cases. The workload will be carried out by a librarian, a full-time assistant, and a graduate student. Due to the nature of this sensitive work the timeline will be longer in order to allow for multiple double-checking and ensure victims' names and addresses are not missed. For this phase six to eight months will be needed. **Phase 3** will be when the cases will have metadata description applied to each case following the Dublin Core metadata standard. This work will again be carried out by a librarian, a full-time assistant, and a graduate student. Approximately six months will be needed for the cases to have metadata descriptors added. During **Phase 4** the data will be double-checked and organized for ingestion into DigitalUSD. It is estimated that around three months will be needed for this phase. In **Phase 5** the cases will finally be launched publicly for all types of users worldwide. We estimate that we will need about 2 months to finish this work. This includes testing the launch as well as marketing it to the scholarly community at large.

Detailed budget: USD is seeking funding for Phase 1 salaries to cover two Mexican nationals who will be digitizing and organizing files based on the naming file conventions established during the 2017 pilot.

Budget:

	Time Period	Salary/hour	Total
Two Digital Lab Technicians	4 months—30 hrs. per week	US\$15.00 per hour	US\$15,000.00

I have a pool of qualified people to interview for these positions. Baja California is a more expensive state than central and southern ones in Mexico. The cost of living is higher in the north and to hire qualified and productive employees a minimum of \$15/hour needs to be paid.

Funding from USD includes:

- Librarian’s salary and travel to Tecate, BC to digitize cases along with hired staff
- USD-Copley Library will cover the librarian’s time spent on this project
- Purchase anonymizing software [Adobe Acrobat Pro DC license]

Once the project is finalized, USD will be donating to the CEDH the equipment and purchase anonymizing software in order for them to continue digitizing newer cases.

The CEDH will be supporting this project with:

- Working space
- Electricity
- Internet access
- Salary of CEDH Archivist for a few hours a month to assist with discarding the cases once they have been digitized and quality of scan has been reviewed.

Timeline:

Phase 1 Timeline:

Original: July – October or August – November 2020

Due to COVID-19 now December 2020 – March 2021 or later depending on pandemic's development

- Two (2) assistants working 30 hours/week working with two scanners with a feeder and one flat scanner
- Goal is to digitize 276 boxes or approximately 115 linear feet.
- Cases are first scanned as TIFF files which are then converted to PDF files
- During the 2017 pilot which lasted nine weeks approximately 60,000 pages were digitized or 33.33 linear ft

Post-project Assessment Plan: Describe what methods are going to be used to assess the outcome of the project (measure(s) of success):

- Amount of linear ft. of material digitized, anonymized, described, and uploaded to DigitalUSD
- Statistical Data collected from DigitalUSD on visits, downloads, etc. and monthly reports will help assess users' interactions with the digitized material

Additional sources of funding: Complete list of funding already secured and/or pending: USD's Transborder Institute (now department within the Institute for Peace and Justice) and the Dean of the School of Peace, Patricia Marquez, have stated that they will provide a bilingual Graduate Student in Peace Studies to assist with Phases 2 and 3 of this project.

Authors

Not applicable.

Publishers

Not applicable.

Proposal Contributors

Not applicable.

Details and Sources

Resource Types*

- Archival materials

Regions*

- Latin America and Caribbean
- North America

Countries of Origin*

Mexico

Major Languages

Spanish

Source Format*

- Paper
- Other Paper

Target Format

Digital

Identifiers

Not applicable.

Links

Not applicable.

Source and Holdings

Location	Format	Holdings/Range	Notes
Comisión Estatal de los Derechos Humanos Baja California – Archivos en Tecate, BC	Paper legal case files	Archives of human rights cases going back to the early 1990s through the early 2000s, 276 boxes.	Condition Excellent to Fair

Source and Holdings Notes

Not applicable.

Physical Details

Paper archives of human rights violation cases. 276 boxes = approximately 115 linear feet.

Existing Reformatted Materials

Not applicable.

Intellectual Property Considerations

The CEDH is the institution responsible for the custody of these cases and all of other pertaining information. They are the ones with the authority to grant permission to digitize and make the cases available to the public, as long as USD removes all personal information on