Description and Rationale

Title

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

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 1B to digitize the remaining material due for destruction in Summer 2022 and thus allow for a more complete archive to be processed during the remaining phases of the project.

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.

The 2020 proposal stated that the funding was to be used to clean up, organize, prepare, and scan only the materials scheduled for destruction, which as of October 2021 no longer exist. Halfway through scanning everything that met the criteria, one of my assistants and I made two discoveries. First, we discovered many poorly labelled boxes; and second, during the time the offices were closed to the public, the lawyers at the different municipalities, having heard that the digitization project was forthcoming, did a thorough clean up and sent the last of the cases and paperwork from the 1990s through 2007 from their respective units. These two events were definitely unexpected. We believed we were scanning a little over 70% of the material available based on our estimates after the tour and going through some boxes, but it was actually closer to 55%. Near the end of what is now called Phase 1A, we inspected, reorganized, and relabeled every box by year and municipality. This was done in the hopes of scanning everything of value before moving on to Phase 2 and before this final batch of cases is removed from the archives and destroyed in Summer 2022.

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)
- 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

As of now it is still planned that I will be working with an assistant from the Archives and Special Collections Digital Initiatives Department and graduate student/s 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 complaint. 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 1A: Digitization and OCR [est. time frame]: 4 months with two digital lab technicians working 30 hours/week

Phase 1B: Digitization and OCR [est. time frame]**: 3 months with two digital lab technicians working 25 to 30 hours/week and the Commissions Archivist working up to 10 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 1A** of the project, cases will be organized by case size and length to prepare for digitization. The assistants 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 processed and prepared by either removing staples, paper clips, notes, etc., and 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, so the library staff can redact sensitive information during the anonymization phase, and for researchers can do full text searching when the files are available to the public. 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. The remaining cases in **Phase 1B** have already been organized by year and municipality and are ready to be digitized. After this is done, we can move to Phase 2 since we would have digitized all pertinent cases from the 1990s and early 2000s.

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.

In **Phase 3,** 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 1B** salaries to cover two Mexican nationals who will be digitizing and organizing files based on the naming file conventions established during the 2017 pilot and expanded in Phase 1A.

Budget:

	Time Period	Salary/hour	Total
Two Digital Lab Technicians	3 months—30 hrs. per week	US$15.00 per hour	US$9,000.00

The $9,000 requested is based on the Commission's offer to help with some of the labor this time around in order to finish digitizing the remaining 45% of materials. **USD is also allowing me to spend more time onsite** given that COVID-19 is becoming less of threat. If all goes well, we are positive we can commence on Phase 2.

I continue to have access to a pool from which to interview people 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 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 at least one full day per week of the project's duration
- 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 8 to 10 hours/week up from a few hours/month to assist with preparing cases and discarding these once they have been digitized.

Timeline:
> Phase 1B Timeline: January 2022 - March 2022
> - Two (2) assistants working 25 - 30 hours/week working with two scanners with feeders and one flat scanner

- Goal is to digitize the remaining cases, around 125 boxes (of different sizes) or approximately 60 linear feet in order to move to Phase 2 with all of the 1990s and early 2000s cases having been digitized

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 the project's Phases 2 and 3.

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, 125 boxes.	Condition Excellent to Fair

Source and Holdings Notes

Not applicable.

Physical Details

Paper archives of human rights violation cases. 125 boxes = approximately 60 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 cases by anonymizing the records before they are ingested into DigitalUSD. (See their attached letter of support for this project)

Processing Requirements and Costs

Note: All processing costs values will only be visible to CRL members if this material is posted on any CRL websites.
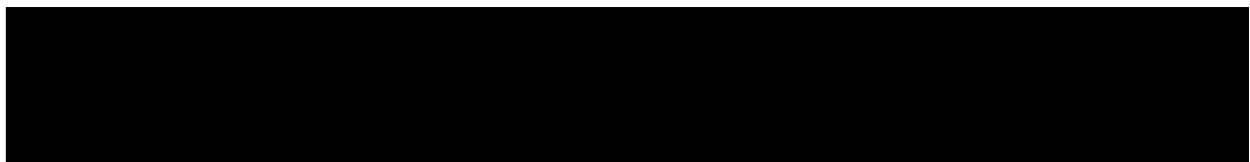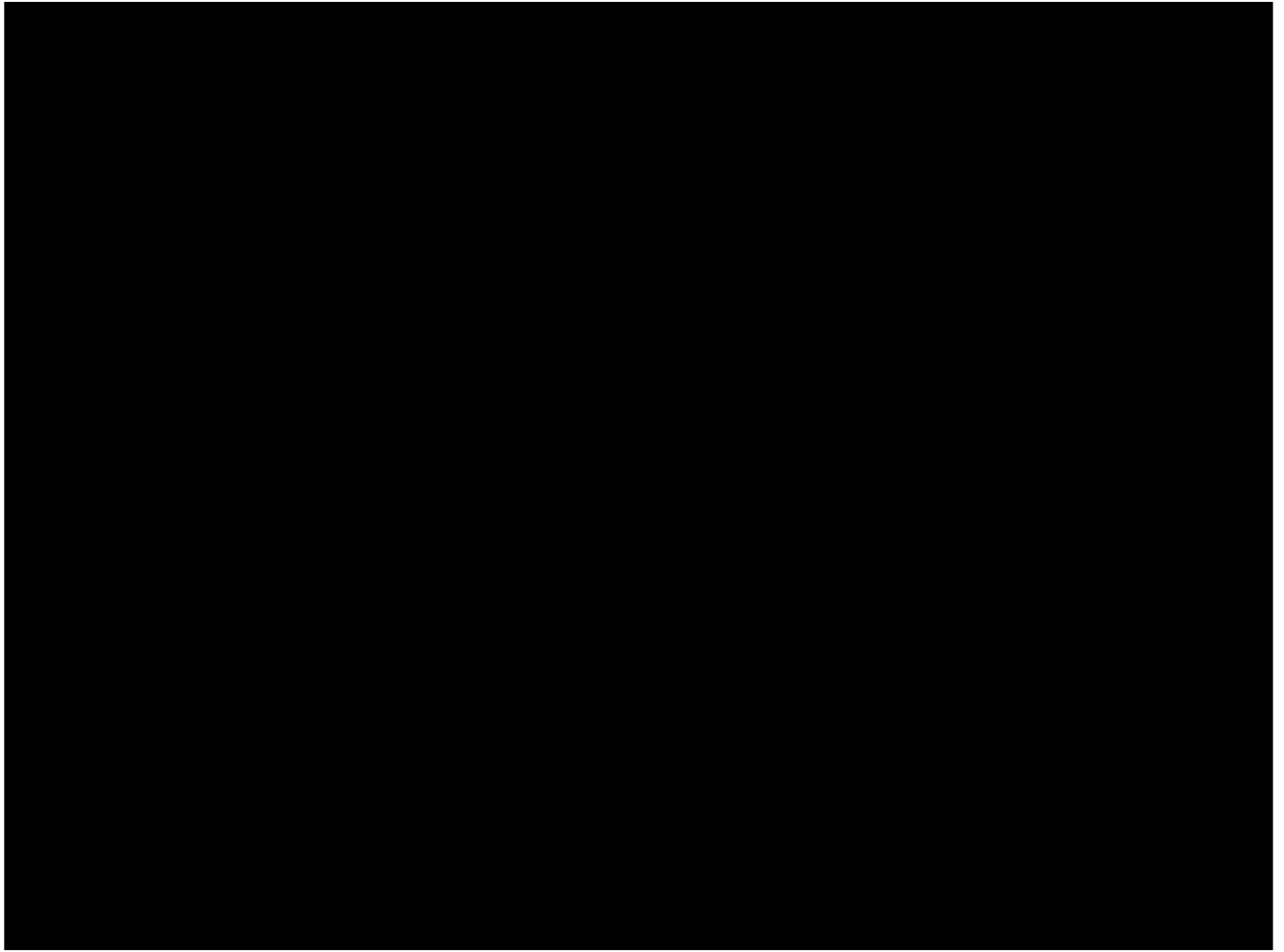
COMISIÓN ESTATAL DE LOS DERECHOS HUMANOS DE BAJA CALIFORNIA

.

PRESIDENCIA
Asunto: El que se indica
Número de oficio: CEDHBC/TIJ/OT/121/2021

Tijuana, Baja California, a 04 de noviembre de 2021

"30 años escribiendo la historia de la defensa de los Derechos Humanos en Baja California"

LARRP ADVISORY COMMITEE
P R E S E N T E

Por medio de la presente, aprovecho la ocasión para enviarle un cordial saludo, asimismo manifestarle la intención institucional de continuar apoyando el proyecto de digitalización de la documentación con que cuenta esta Comisión, dicho proyecto lo apoyamos ya anteriormente en el año del 2017 y este año 2021, con el cual quedamos muy satisfechos; por lo que seguimos apoyando para que se vuelva a aplicar a finales del año en curso, o en 2022.

No omito manifestar que, al momento de que desarrollen las actividades del proyecto en esta institución a mi cargo, contarán con un espacio físico adecuado a las necesidades para que puedan desarrollar sus labores y puedan utilizar el equipo con el que cuentan para realizar el trabajo correspondiente, asimismo contarán con el apoyo de una persona de esta institución, por un periodo de tiempo de 8 a 10 horas por semana. Como ya es de su conocimiento podrán utilizar los servicios de energía eléctrica y demás servicios necesarios que proporcionaremos.

Es de suma importancia recalcar que los expedientes que han sido integrados en este organismo público autónomo son documentos que cuentan con datos personales y pertenecen a esta dependencia, otorgando la autorización a la Universidad de San Diego de colgar los documentos digitalizados **(después de ser anonimizados)** por lo cual se solicita que sean omitidos los datos personales tanto de las personas civiles como de las autoridades y no sean expuestos al público en general sin excepción alguna.

Sin otro particular, me despido reiterándome a sus apreciables órdenes.

ATENTAMENTE

MIGUEL ÁNGEL MORA MARRUFO
PRESIDENTE

C.c.p- Minuta
MSR

November 5, 2021

Center for Research Libraries
LARRP Executive Committee
6050 S. Kenwood Avenue
Chicago IL 60637-2804
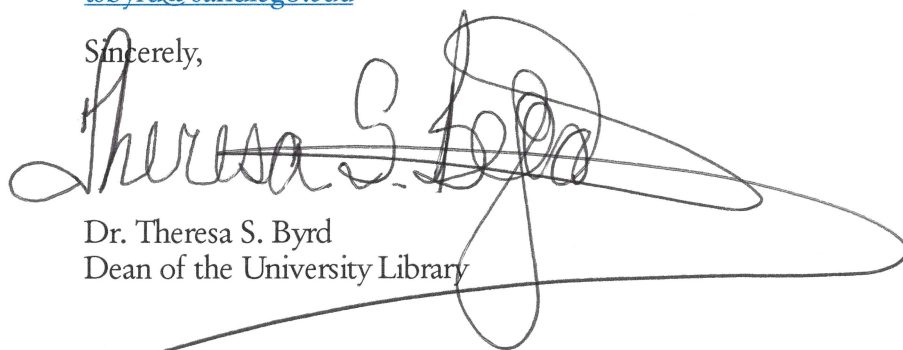
Dear Members of the LARRP Executive Committee,

This is in support of Dr. Alma Ortega for the Comision Estatal de los Derechos Humanos de Baja California (CEDH-BC) archives digitization project and its mission to inform the local community and world scholars on issues of human rights on the California / Baja California border.

The project entails digitizing Baja California human rights cases in the Commission's Archives located in Tecate, BC. Once digitizing is completed, these cases will then be OCRed. After this step, a copy of each case will be made with identifying information omitted. The redacted cases will then have metadata applied and only then can the cases finally be ingested into Digital USD, Copley Library's institutional repository, for researchers to access online.

Dr. Ortega is seeking funding for two assistants, Mexican nationals, in order to finish Phase 1: digitizing the remainder cases and OCRing. Time permitting, they can also assist with jumpstarting Phase 2: redacting and applying metadata.

As the dean, I am donating Dr. Ortega's time and travel budget for this project. Dr. Ortega will be onsite one day per week for the duration of the estimated 10-week project. Please let me know if you have additional questions of me regarding Dr. Ortega's grant proposal. I can be reached by email at tsbyrd@sandiego.edu

Sincerely,

Dr. Theresa S. Byrd
Dean of the University Library