

## **Daily Breeze Digitization Project**

### **Executive Summary**

The Los Angeles County Public Library is pursuing the digitization of a portion of the historic *Daily Breeze* newspaper collection for the preservation of the collection and to increase public access.

Dana Vinke, Redondo Beach Public Library Director and Timothy Parker, Technical Services Librarian met with Brian Geiger director of UC Riverside's Center for Bibliographical Studies and Research and Wendy Crutcher, Library Administrator, Technical Services, of the Los Angeles County Public Library (LACPL) and Amy Boteilho, Assistant Library Administrator in Technical Services for LACPL to discuss the possible digitization of the *Daily Breeze* newspaper.

Based on the results of that discussion, LACPL has prioritized the project to digitize their historic *Daily Breeze* newspaper microfilm collection spanning from 1904 to 1977. This project will preserve the newspaper for decades and increase awareness and access to important local cultural and historic events.

### **Background**

The *Daily Breeze* newspaper as we have come to know it was founded as a weekly paper titled The Redondo Breeze in 1894 by Redondo Beach local S. D. Barkley. The newspaper's primary service area was the local community or Redondo Beach. As the population of the South Bay grew, so did the newspapers coverage area. The newspaper became a daily in 1922. The paper went through a variety of name changes including the *Redondo Daily Breeze*, *South Bay Daily Breeze*, and simply the *Daily Breeze*. The paper also had several owners including the Copley Press, the Hearst Corporation and the Los Angeles Newspaper Group (LANG).

Previously, Susan Anderson, former Redondo Beach Public Library Director, had worked with Kelly Riddle, Digital Projects Coordinator for LACPL on a potential *Daily Breeze* newspaper digitization project. At that time, the strategy was for LACPL to ship all or parts of the collection to Newspapers.com. Newspapers.com would digitize the newspaper and provide access to the Library.

An affiliate of Ancestry.com, Newspapers.com states it is "the largest online newspaper archive consisting of 912 million+ pages of historical newspapers from

24,400+ newspapers from around the United States and beyond". According to Wikipedia, Newspapers.com increase their databases by working with libraries, publishers, and historical organizations to scan publications for free to include in their database. Many organizations view this process of free scanning as an easier, cheaper and quicker way to get their publications online than other alternatives.

Ms. Anderson retired in 2022 and Ms. Riddle is no longer with LACPL. Dana Vinke, Redondo Beach Public Library Director, subsequently began to revisit the digitization of the historic *Daily Breeze* newspaper microfilm collection.

### **Assessment**

The procedure of transforming traditional print newspapers or microfilm from their analog format into digital imagery is known as newspaper digitization. The majority of historic newspapers are available in physical paper or microfilm formats. In order to make these collections available online, print or film of the newspaper are scanned and the resulting digitized pages are analyzed (though not always) with optical character recognition (OCR) software to generate text files containing the newspaper's content and to enable the newspaper to be searchable. OCR is an essential component of newspaper digitization, as without it, the newspaper is only accessible by date. An analogy would be searching Google by the date something was posted instead of using words and phrases.

There are several major challenges posed by the digitization of the *Daily Breeze* newspaper. These include: copyright and permissions; physical possession of the microfilm; scanning and image capture or digitization; OCR; digital asset management and storage.

The first challenge in a digitization project is to determine where the physical archives reside. The only physical record of the early years of the Breeze newspaper are the microfilm holdings at LACPL at the Hawthorne Library. There is no known archive of the physical papers. The Redondo Beach Public Library does not own any *Daily Breeze* microfilm. LACPL maintains the largest, albeit incomplete, collection of the *Daily Breeze* newspaper on microfilm. In many cases, the Library maintains the only remaining copies of the newspaper. LACPL recently performed an inventory of their *Daily Breeze* newspaper microfilm collection and found that from the paper's inception, 1894 to 1977, LACPL houses 539 rolls of

microfilm spanning from 1904 to 1977. Rolls vary in size making the ability to determine how many images may result from digitizing such a collection difficult to estimate.

The second challenge in a digitization project is to determine the rights to digitize the newspaper microfilm. The term “public domain” includes those materials which are not protected by intellectual property laws such as copyright. No individual owns these works; instead, they are owned by the public. Newspapers published before 1926 are in the public domain. Depending on whether a publisher registered with the U.S. Copyright Office, newspapers published before 1978 may or may not be in the public domain. If a newspaper is not in the public domain, permission of the copyright holder is required to digitize and bring such a paper online. The Redondo Beach Public Library does not currently have permission to digitize and make available the *Daily Breeze* newspaper online.

Once the challenge of copyright has been addressed, one of the next challenges is determining how the microfilm will be scanned and in what format. It is important to note the distinction between digitization and scanning. Although both convert a hard copy into a digital file format, only digitizing allows for a document to be OCR'd and provides additional meta-data points. This difference is why most preservation projects digitize materials instead of only scanning them.

It is generally recommended that images be digitized to TIFF file format and made available in PDF format. TIFFs are the highest quality image deliverable. They use a lossless compression algorithm to preserve image quality. They are a standard in the field of preservation. TIFFs require more storage space than PDFs and are not ideal for web use due to their large file size. PDFs are ideal for access because they are not platform-dependent, ubiquitous, OCR compatible, and require less storage space.

It is recommended that digitization be performed by professionals given the rarity and condition of the physical newspaper microfilm. While volunteers or staff and locally sourced equipment could be used, any number of issues could arise that would necessitate the use of specialized equipment and professionally trained staff. Currently, the Redondo Beach Public Library does not have the staffing, time, or funding to accommodate either an in-house staffed or a volunteer digitization project.

Another challenge is selecting a digital preservation platform. Digital preservation platforms combine policies, strategies, and actions to ensure access to digital content. They involve planning, resource allocation, and the application of preservation methods and technologies. There are several platforms available and they include both open and closed source software and there are costs associated with both types of digital preservation platforms. For example, the Redondo Beach Public Library currently utilizes Laserfiche to host its collection of Redondo Compass (1892-1893) and the Redondo Reflex (1906-1970). The Library also utilizes Montage, an Auto-Graphics, Inc. software to host a collection of local High School Yearbooks. Maintaining multiple platforms is not ideal as it can be costly, inefficient, and not user friendly.

### **Recommendation**

It is our recommendation that Redondo Beach Public Library support LACPL in their endeavor to digitize the historic *Daily Breeze* newspaper. LACPL has been digitizing unique items and collections since 2012. Most of the Library's digital collections come from 3 sources: the Library's own collections, archival collections at other County departments, and material contributed by communities within Los Angeles County. LACPL takes steps to ensure that it has appropriate rights to digitize and provide access to content. Where possible, LA County Library secures rights to use works that are in copyright. LACPL will work with their County Counsel to ensure the collection is properly insured and that any access agreement is fair to the Library and their patrons.