

# FROM REDACTION TO ACCESS

## *Navigating Challenges to Unlock Houston's LGBTQ Media History*

**Emily Vinson**

*University of Houston Libraries  
United States  
evinson@uh.edu  
0000-0002-8902-2454*

**Bethany Scott**

*University of Houston Libraries  
United States  
bscott3@uh.edu  
0000-0002-8242-6384*

**Abstract – *The Gulf Coast LGBT Radio and Television Digitization Project* launched in 2020 to digitize, preserve, describe, and make accessible thousands of hours of Houston's LGBTQ broadcast history. The authors explain the significance of the programs selected for inclusion in the project and describe the steps taken to balance the goal of equitable access to unique materials created by and for a marginalized community, while maintaining adherence to copyright restrictions.**

**Keywords – LGBTQ Community, Broadcast archives, Audiovisual archives, Digital preservation, Accessibility**

**Conference Topics – Digital Accessibility, Inclusion, and Diversity**

### 1. INTRODUCTION

The University of Houston Libraries (UHL) received the multi-year grant for *The Gulf Coast LGBT Radio and Television Digitization Project* in 2020. This grant, funded by the National Endowment for the Humanities, Division of Preservation and Access, Humanities Collections and Reference Resources Program, supports the digitization, preservation, and online publication of Houston's LGBTQ broadcast history. The selected programs cover a wide range of topics, including politics, activism, health, and LGBTQ identities. This poster provides an overview of the project, its historical context, and the steps taken to ensure access, with a particular focus on our approach to redacting copyrighted content in access copies.

### 2. PROJECT OVERVIEW

Now in its final year, *The Gulf Coast LGBT Radio and Television Digitization and Access Project*, through post-custodial and traditional archives relationships, has reformatted and is in the process of transcribing and describing over 3,500 hours of locally produced radio and television created by and for Houston's LGBTQ community.

The project includes various radio programs produced at Houston's listener-sponsored Pacifica radio station, KPFT, such as "Wilde 'n' Stein," "Lesbian & Gay Voices," and "After Hours." Each program contributed to the representation and empowerment of the LGBTQ community in Houston. The earliest program included in the project, "Wilde 'n' Stein," focused on community outreach, activism, and education about safer sex during the AIDS crisis. "After Hours," a late-night program that aired for thirty years, offered a blend of music, news, activism, entertainment, and an intersectional perspective on the LGBTQ experience.

One television program is included in the project. Broadcast between 1998 and 1999, "TV Montrose" was designed to appeal to all audiences, with a particular focus on LGBTQ life and culture.

#### A. *Project Partners*

By combining existing collections with post-custodial collaborations, our goal was to assemble a comprehensive collection of Houston's LGBTQ media history. Donors and project partners had not only collected and housed the source materials, but also

contributed to the creation of these unique recordings.

### B. *Project Workflow*

As our project launched in May 2020, the challenges caused by the COVID19 pandemic necessitated significant changes to our work plan; however, the broad strokes of our plan and the goals of the project remained the same: to make Houston's LGBTQ media history accessible.

*Inventory:* Due to the University's hiring freeze, the project PIs took on the task of shifting and inventorying assets from our partner, GCAM. Following safety protocols, we moved analog recordings from GCAM's storage to UHL custody. To expedite the project, we brought tapes home for inventory, transcribing show notes from cassette labels. For digitization, we outsourced to George Blood Audio Video Film, receiving files in various formats with technical metadata and checksums. Samples were assessed for quality assurance.

*Redaction:* In our planning phase, we obtained permissions from KPFT and the Pacifica Network Archives, but including third-party content in the show episodes posed a challenge for online publication. Licensing agreements held by KPFT did not extend to UHL's digital collections repository, especially concerning commercial music and rebroadcasted news briefs. However, considering the historical significance of the shows, we decided to redact the third-party content to make the recordings available online worldwide. We conducted a fair use analysis and believe that the educational and informational nature of the recordings justifies their online display. The transcripts, descriptive metadata, and transformation of the original works enhance their educational value and promote scholarship.

Though we considered using an AI-based approach to identify music for redaction, we found that most shows included hosts and guests talking over music and original mixes. In order to not remove original show content, we developed a workflow with three student staff members. Over 3,600 files were evaluated for content to be redacted using visual inspection of the waveform and spot-checking the sound. Student employees determined which content required redaction, such as commercial music or news briefs, and which did not, such as the show's theme music. For redacted music,

we included a few seconds of the intro and outro with a few seconds of silence in between. These redactions are also noted in the transcripts.

Redacted versions are edited access copies available for online public access. Preservation master files remain unedited, and unedited access copies are accessible in the UHL Special Collections Reading Room or upon request online with restricted access.

*Accessibility:* Accessibility was a priority, and we aimed to include transcriptions and/or captions for all materials. Machine-generated transcripts alone proved inadequate, leading us to allocate a portion of the budget to professional transcription services. 3PlayMedia was selected to create captions for the television program and transcripts for a portion of the radio programs. For the remaining radio programs, we used an in-house workflow where redacted audio files were submitted to the Otter.ai platform and corrected by student employees. To enhance accuracy, we created vocabulary term lists based on tape label transcripts and shared spreadsheets to track names, places, and acronyms. These terms also inform the descriptive metadata, ensuring consistency throughout the project.

*Preservation:* We are committed to preserving the media and transcription files included in the project. We have incorporated the project files into the UHL digital preservation program, which follows the OAIS reference model. For each item in the project, we generate a submission information package (SIP) that includes the preservation master file(s), the caption or transcript files, a metadata CSV, and a persistent identifier/permalink. These SIPs are then exported to Archivematica for preservation ingest and storage.

*Descriptive Metadata:* In collaboration with the UHL metadata unit, we've strived to balance the demands of a large project with the need for descriptive metadata. We include label transcriptions in descriptions when available, and for those without, keywords and host names from transcripts serve as the basis for our metadata records. To ensure inclusive language, we consult subject matter authorities such as Homosaurus and project partners. Metadata records, along with a streaming link to the UHL repository, will be shared with the American Archives of Public Broadcasting for enhanced discoverability.

*Publication:* UHL's media collections are available online through our instance of Avalon Media System. This system allows us to upload captions for videos and transcripts for audio recordings in bulk. To support new research methods, such as "collections as data," we will make all transcripts available as plain text files in the UHL Dataverse Repository. This will allow researchers to explore the collection using computational analysis.

In addition to making the recordings accessible to the community that created them, we also hope to increase their reach by connecting with faculty and researchers who may use them for teaching and scholarship. It is only through preservation, long-term access, and reuse of these important collections that we help to diversify the historical record for generations to come.