Supporting Solidarity:

Appraising and Collecting Online Content Surrounding the

Women's Marches in Maryland

Erin Berry, Caitlin Rizzo, Jen Wachtel

LBSC 785: Documentation, Collection, and Appraisal of Records

May 2, 2017

Dr. Punzalan

University of Maryland

Rationale

Given the task of appraising and collecting online content that documents the National Women's March, our group chose to act locally. We focused our efforts on the attempt to preserve content surrounding Women's Solidarity Marches in Maryland that occurred as a complement to the National March on January 21, 2017. As a group, we felt that while the National Women's March had been well-documented and preserved, work still needed to be done to support the preservation of content created during the smaller marches which sprung up in response to the National Women's March. Of course, collecting materials surrounding all of the solidarity marches would be too broad, so as a group we chose to focus on solidarity marches in the state of Maryland with the ultimate hope that we could accession our finished collection of materials in the repository where we collectively work: University of Maryland Special Collections and University Archives (SCUA). SCUA actively collects material pertaining to the state of Maryland history and culture, historic preservation, and women's history. The women's history materials include a crowd-sourced collection which aims to document the National Women's March, a perfect fit for our finished collection.

Our group decided to contribute web content that fit the scope of the women's history collections at SCUA by appraising a mass of online content, including organizational websites, social media pages, photographs, and videos. Since local groups and individuals organized the Women's Solidarity Marches, our group appraised material that covered local groups and organizations' creation, mission, organizing methods, and final documentation of the March to be of the most value. We focused our appraisal on web content and platforms that captured evidence of the main function of solidarity marches: to bring local women together, to organize large groups of women and supports, and finally to create their own marches. With these parameters in mind, we ultimately decided to crawl Women's March websites, including group organizational pages hosted by Facebook, as well as YouTube and Flickr for images and videos of the actual marches.

Reflection

      The group's completed collection is titled "Women's Solidarity Marches in Maryland" and is available at the following address: https://partner.archive-it.org/663/collections/8832. The final collection documents official solidarity marches hosted in Baltimore, Annapolis, and Ocean City. The collection features 26 seeds, including organizational websites, Facebook pages, YouTube videos, and Flickr images. We captured official organization pages from Action Network, the platform that the Women's March on Washington and official solidarity marches used for sharing logistical information. The solidarity marches in Maryland also created a central page, www.wmw-md.org, which we included in our collection as well.

      Each member of the group ran crawls for organizational websites and hosted media content (i.e. Facebook, Flickr, and YouTube). Each of our seeds required the "Standard" setting on Archive-It, and yet each type of seed required its own specific set of scoping rules and logic and presented unique difficulties. The crawls of organizational websites captured the necessary data including text, photographs, and embedded maps without much trouble at all. While the final capture of social media pages, specifically Facebook pages, produced the desired results of saving both images and continuously scrolling text, there were certainly difficulties to achieve this effect. For example, the Archive-It recommended data limit for Facebook crawls is one gigabyte, so continuous scrolling worked on some seed captures but not others. Facebook group for the Women's March on Annapolis requires individuals to join the closed group in order to view the content. Since Archive-It required that we provide our own personal login credentials to crawl content on password-protected groups, we decided not to publicize that crawl and risk exposing our own personal data.

      We also faced difficulties when crawling YouTube and local newspapers. This content required that we ensure the seed rules for the crawl were set to "ignore robots.txt" in order to bypass files that block web crawlers, per Archive-It guidelines for YouTube crawls. The majority of YouTube videos that were crawled with these parameters were successful, though one video was not able to be captured using these methods. To run the test crawls, each group member created a large group of seeds to be crawled at once. Due to time constraint and the method of crawling groups of seeds, Erin had to save the data of the one video that did not work to be able to capture the other working material. The only content appraised as vital to our collection was the video evidence of the march, the crawls only captured the video data from the

websites and were considered successful. Our group faced similar challenges in our attempt to archive local newspaper pages with unique coverage of spontaneous marches in Garrett County. The newspapers protected their content with "robots.txt" files to prevent crawling. Following the Archive-It guidelines, the group reached out to the webmasters of both local newspapers in order to obtain permission for a crawl. Especially considering our hope that the material would be accessioned into an actual repository, we felt obtaining permission to collect the content was necessary ethically. Neither webmaster responded and ultimately both local newspapers had to be cut from the collection in order to make the deadline.

Our final challenge came from Flickr images. The heavy java encoding on Flickr's website required a special crawl in proxy mode. In order to complete the crawl, it was necessary to use Archive-It's proxy mode to assess the content properly, so the group downloaded an app created by Archive-It to allow for viewing crawls in the proxy mode. This allowed for the successful crawl of vital images of the march which displayed important march ephemera like signs, attire, and group diversity.

Ultimately, we hope that by capturing a variety of sources used by the creators of the movement we successfully created an archive "of the community" rather than simply "about the community," in the words of Shilton and Srinivasan. The group believes that even though we lost content from local newspapers, the collection was ultimately made stronger by including only documents created, organized, and generated by the solidarity marches. We hope the archive itself offers a message of solidarity to the participants in that it will allow their actions to be preserved through their own voice far after the march itself.

References

Shilton, Katie and Ramesh Srinivasan. "Counterpoint: Participatory Appraisal and Arrangement for Multicultural Archival Collections." *Archivaria* 63 (Spring 2007), 87-101.