Engaging with Communities and API Driven Accessioning of Digital Folklife

September 2016
Based on the iPres 2016 paper:

From Crawl to Call & Pull: Engaging with Communities and API driven Accessioning of Digital Folklife

David Brunton, LC
Trevor Owens, IMLS
Abigail Potter, LC
Nicole Saylor, LC
Kate Zwaard, LC
American Folklife Center
Preserve & Present Folklife (Public Law 94-201)

“…the traditional expressive culture shared within the various groups: familial, ethnic, occupational, religious, regional; expressive culture includes a wide range of creative and symbolic forms such as custom, belief, technical skill, language, literature, art, architecture, music, play, dance, drama, ritual, pageantry, handicraft; these expressions are mainly learned orally, by imitation, or in performance, and are generally maintained without benefit of formal instruction or institutional direction."
A history of participatory archives
Folk culture on the web
Crawling folk culture on the web

The Library of Congress archive features a collection of web sites, including Serra 45: Presidente do Brasil, Lutheran World Federation, Pew Research Center: Numbers, Facts, and Trends Shaping Your..., and Slaves of the Immaculate Heart of Mary. The collections are composed of sites selected by subject specialists to represent web-based information on designated topics. It is part of a continuing effort by the Library to evaluate, select, collect, catalog, provide access to, and preserve digital materials for researchers today and in the future.
Approaches to acquiring web data

CRAWL (web archiving)

- Harvested via Heritrix, saved in WARC
- Crawls a list of seeds URLs (the preservation objects), created by a curator
- Acquisition and processing is often invisible
- Follows internal links to a specific depth
- No verification against original source is possible
- Manual quality control

CALL & PULL (born-digital datasets via API)

- Develop individual scripts to pull data from web APIs
- List of files (the preservation objects) pulled from the API
- Negotiation and action required of third parties or public
- No internal links in data
- Data was created or shared in an app
- Website presentation is not the object of preservation
Engaging with Communities and API driven Accessioning of Digital Folklore

#FolklifeHalloween2014

Schiller Park, Columbus, Ohio, October 30, 2014. Photo by Katherine Borland
Public is asked to post photos on Flickr with a specific #hashtag and a creative commons license the Flickr API and ingested into the LC system

Carla Tomaszewski's hands. Photo by Edwin Remsberg

#mytradition

Megan dances the Sailor's Hornpipe at the Scottish Cultural Center in Scotia, NY. Photo by Kathy Park.
Oral narratives go global

Record and share the stories around you.

Choose someone to interview. Pick great questions. Find a quiet place to record. Listen closely.

When you're finished, share your interview with the world.

Help create an archive of the wisdom of humanity.
Oral narratives go global

- Record and share the stories around you.
- Choose someone to interview. Pick great questions. Find a quiet place to record. Listen closely.
- When you're finished, share your interview with the world.
- Help create an archive of the wisdom of humanity.
Engaging with Communities and API driven Accessioning of Digital Folklore

Call & pull

Boys' tug of war, Fourth of July celebration. Vale, Oregon [http://hdl.loc.gov/loc.pnp/fsa.8a30297](http://hdl.loc.gov/loc.pnp/fsa.8a30297)
Common path for new development

1) Manual download
2) Manually kick off script
3) Automate kicking-off script
4) Build script into application
5) Continually improve script
Opportunities

- Web archiving is still the dominant method for capturing born-digital content
- Developing scripts to pull from APIs and negotiating with third parties to provide metadata complements the web archives and broadens the content available to researchers
- Other advantages to this approach:
  - Makes acquisition and preservation more active and transparent
    - Embedding preservation practice into third-party providers’ workflow
    - Engaging the public and/or specific communities in selecting born-digital content for preservation
  - Fixity generally reflects completeness or quality of an object. Ingest and QA can then be automated
Excerpt: “Will you do me the greatest honor, and will you marry me?”

In an excerpt from a longer interview, listen to the moment Rory Miller asks his girlfriend Asya Adcock to marry him.

Rory and Asya, both 25, sat together on July 8 in their Tennessee home and talked about their relationship. They shared memories of their life together and their first kiss in high school that she described as “pleasant” while he was inspired to write a song about it. But that was it for the two of them for the next seven years until they reconnected at Bonnaroo in 2011. Rory had no

https://storycorps.me/interviews/excerpt-will-you-do-me-the-greatest-honor-and-will-you-marry-me/