

# Engaging with Communities and API Driven Accessioning of Digital Folklife

September 2016



LIBRARY OF CONGRESS

Based on the iPres 2016 paper:

## **From Crawl to Call & Pull: Engaging with Communities**

David Brantley, LC  
**and API driven Accessioning of Digital Folklife**

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

STANDARD FORM No. 14  
APPROVED BY THE PRESIDENT  
MARCH 10, 1926

## TELEGRAM

OFFICIAL BUSINESS—GOVERNMENT RATES

ALAN LOMAX \* RADIO RESEARCH PROJECT

FROM \_\_\_\_\_  
LIBRARY OF CONGRESS

BUREAU \_\_\_\_\_  
03x8031 - 29.10

CHG. APPROPRIATION \_\_\_\_\_

U. S. GOVERNMENT PRINTING OFFICE 10-1723

MR. FLETCHER COLLINS,  
ELON COLLEGE,  
NORTH CAROLINA. *by airmail express*

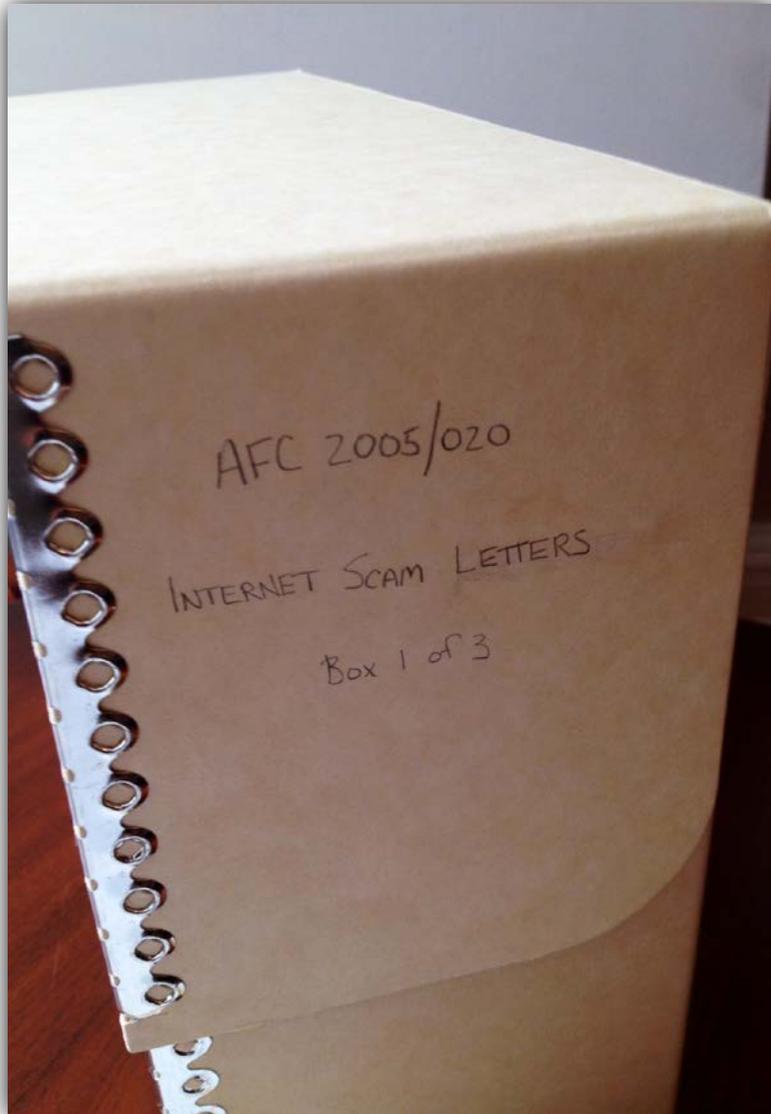
CAN YOU SEND US WITHIN 48 HOURS RECORDINGS OF REACTIONS OF FOUR OR FIVE AVERAGE MEN AND WOMEN TO JAPANESE AGGRESSION. GET COLORFUL PEOPLE WHO SPEAK WELL. LIBRARY WILL PAY COSTS OF DISCS WHICH WILL BE USED FOR HISTORICAL RECORD AND POSSIBLY RADIO BROADCAST.

ALAN LOMAX

*A.L. \$1.60*

*North. N.C.*

# Folk culture on the web



# Crawling folk culture on the web

Library of Congress > Archived Web Sites

Format  
**Archived Web Sites**

[Search Archived Web Sites](#) [Collections with Archived Web Sites](#)

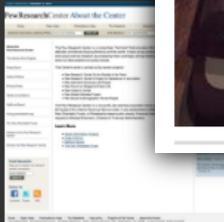
**Featured Content**



Serra 45 : Presidente do Brasil



Lutheran World Federation



Pew Research Center :  
Numbers, Facts, and Trends  
Shaping Your ...



Slaves of the Immaculate Heart  
of Mary

## Angry Splash Mountain Lady - Zoom In

0 0

[Share this Image](#)
[Submit](#)
Like us on Facebook! Like 674K

**PROTIP:** Press the ← and → keys to navigate the gallery, 'g' to view the gallery, or 'r' to view a random image.




[Zoom In](#)

## Collections with Archived Web Sites 1 - 22 of 22

The Library of Congress Web site archives are composed of sites selected by subject specialists to represent web-based information on a designated topic. 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

#FolklifeHalloween2014

10.30.2014

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



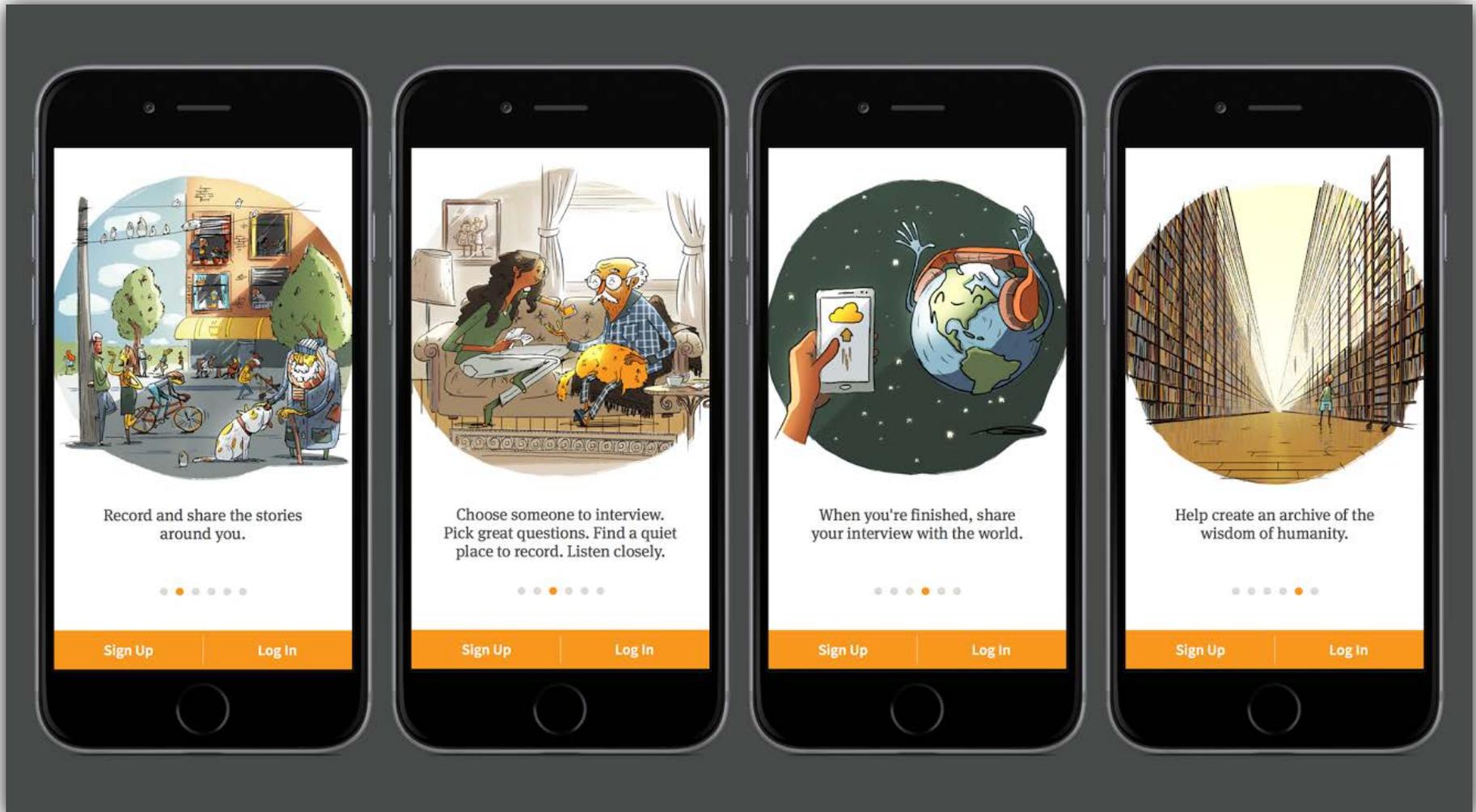
Carla Tomaszewski's hands. Photo by Edwin Remsberg

**#mytradition**

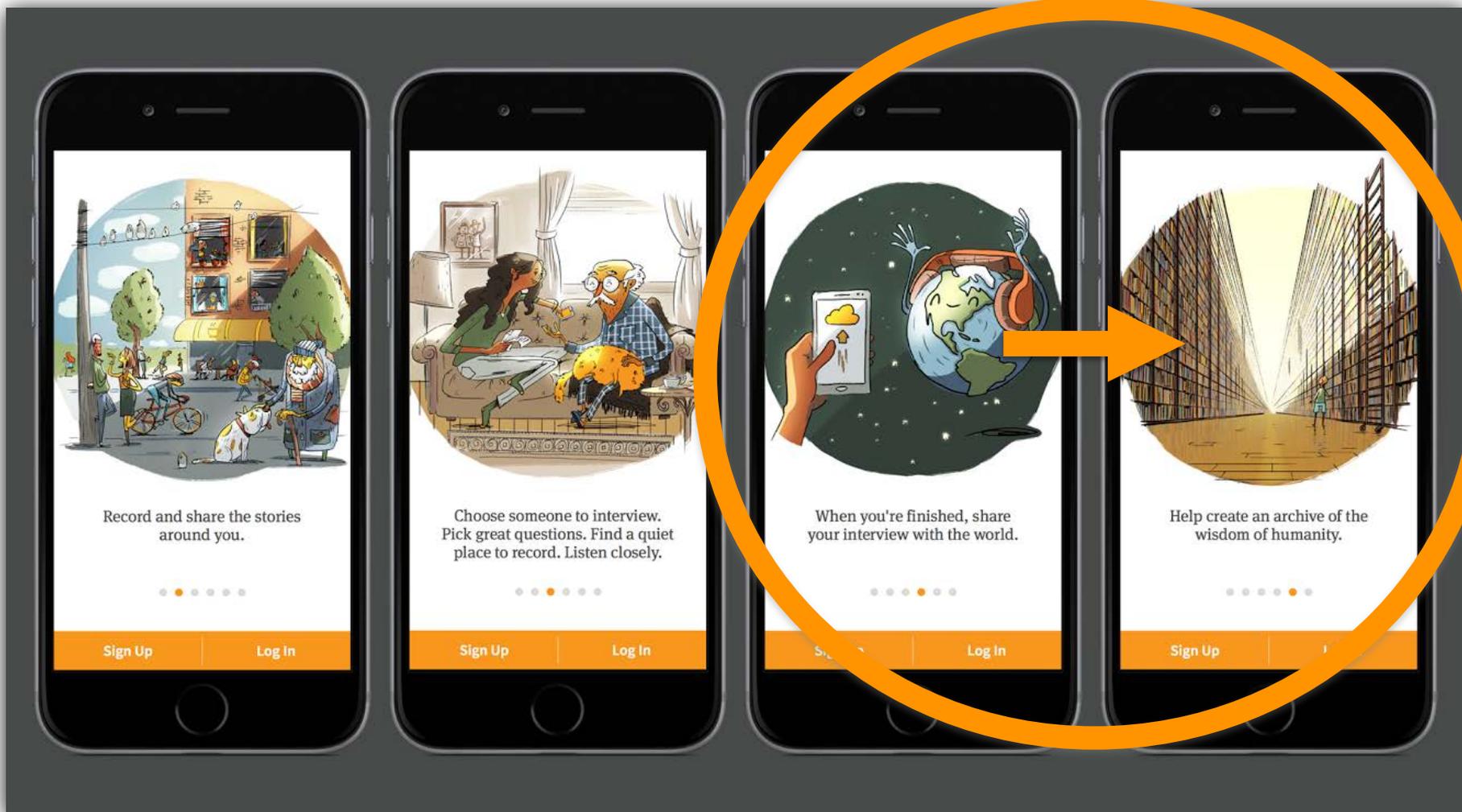


Megan dances the Sailor's Hornpipe at the Scottish Cultural Center in Scotia, NY. Photo by Kathy Park.

# Oral narratives go global



# Oral narratives go global



# Call & pull



Intervie  
w  
packets

Boys' tug of war, Fourth of July celebration. Vale, Oregon <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

StoryCorps .me

Home

Browse

More

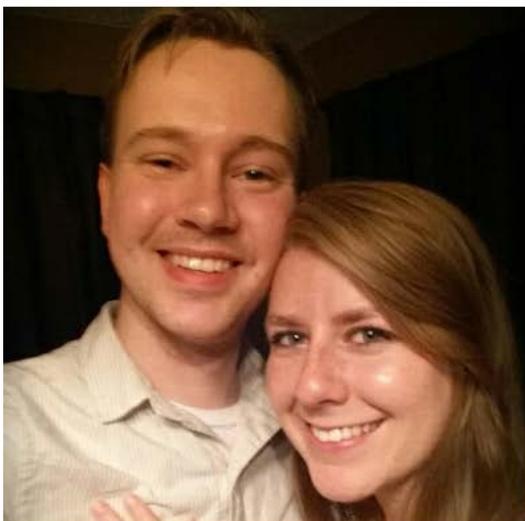
Log In

Sign Up

Donate

StoryCorps  
July 10, 2015

Not Following



## Excerpt: “Will you do me the greatest honor, and will you marry me?”

1:58 167 5749



### Keywords

### Summary

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

### Share



### Participants

StoryCorps

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