Archive-It Partner Meeting
Portland, July 25, 2017

Community-based
Collaborative
Collection-development

Cobweb

github.com/CobwebOrg/cobweb

Kathryn Stine
California Digital Library
CDL has been involved in web archiving since 2005

- NDIIPP Web-at-Risk
- Web Archiving Service (WAS)
- End-of-Term Archive
- CA.gov Archive
  [https://archive-it.org/collections/5763](https://archive-it.org/collections/5763)

The 2015 decommissioning of WAS in favor of Archive-It decentralized web archiving activities at the University of California, so we are particularly interested in ways of promoting collaboration

So …
Imagine ...

A fast-moving event unfolding *online* and *on-screen*, as much as on the ground

Harnessing the domain knowledge and technical capabilities of the entire *community*

Enabling *local* collection development decisions based on *global* information

*Complementary* and *collaborative* collecting

Institutional participation at a level commensurate with local *capacities*

Increasing *researcher* and *public* awareness
Centralized *catalog* of collection- and seed-level *metadata*

Establishment of thematic collaborative collecting *projects*

Open *nomination* of topical seeds by interested stakeholders

*Claiming* of seeds by archival institutions intending to harvest

*Holdings* records for seeds actually harvested
Grew out of discussions by an *ad hoc* group of archiving institutions

Collaborative project of *CDL, Harvard, and UCLA*

Funded by *IMLS #LG-70-16-0093-16*

Expanding upon the function of the UNT *nomination tool*

Building on the experience of *IIPC*’s collaborative collecting efforts (e.g., Olympics)
Why Cobweb?

The demands of archiving the web in comprehensive breadth or thematic depth *exceeds* the technical and financial capacity of any one institution.

Curators cannot make *rational* collection development decisions without knowledge of what others have collected or are collecting.

Relevant seeds can be meaningfully contributed by *numerous* stakeholders: curators, archivists, subject area specialists, researchers, event participants, and the public.

Apportioning collection responsibility into granular pieces *encourages* participation by *smaller* institutions and programs.
Where are we?

Staff roles at partner institutions established

- CDL - PI, project management, and outreach management
- UCLA - Technical development
- Harvard - User interface/user experience design

Project kick-off in July 2017!

- Preliminary data modeling
- Establishing and prioritizing user stories
- Confirming functional requirements
Draft Data Model
Where are we going?

Public sharing and presentation of designs and plans for **review** and **comment**

- Outreach at AIT Partner Meeting, AoIR, CNI, IIPC WAC, SAA WAS
- Website and webinars, user testing

Early **prototyping**

Iterative **enhancement**

Rollout at 2018 **IIPC Web Archiving Conference**

Share at 2018 **Archive-It Partner Meeting / SAA**

How can we **sustain** Cobweb as core web archiving infrastructure?
Community-based, collaborative collection development for web archiving

http://github.com/CobwebOrg/cobweb

Questions: kathryn.stine@ucop.edu