Communities

- Academic Libraries – Law
- Depository Libraries
- Archives
- The Public – Everybody
- Federal Agencies
- Government Transparency
- Teachers
- Cultural Heritages
- State Libraries
- Academic Consortia
- Non-Profit
- Preservation
- Journalists
- FOIA Organizations
- Historical Societies
- Genealogists
- Elected Officials and Staff

Academic and Depository Libraries

- Geo distribution
- Mission alignment
    - Partners with Fed
    - Partners with each other
- Collections – emphasis on publication
- Go-between – agency and users
    - Human
    - Tech
- Dedicated staff
    - Passion
    - FTE assigned
- Focus on accuracy and access
- Academic breadth of collections
- Long Term
- Educational Focus
- Research, service, teaching missions
- Provide context

Federal Agencies (270+)

- Focus on:
    - Permanent record
    - Government publication
- Content creators and collection creators
- Some Federal libraries

- Community and records managers
- Consumers of information
- Deep familiarity
- Built in support system
- Records managers
- All media
- Keep public information public
- Extensive external partners
- <u>Provide Context</u>

We know the pain points for libraries, archives, etc.

- Not for creating agencies

Statutes don't line up with current practice

- o Preservation isn't passive anymore
-

State Government

- Strong working history with public
- Appraisal
- Description
- Organization
- Tight community

Not Present

- Technologists
- Academic Research community
- Info and library vendors

Shared Challenges

Creation impacts ability to preserve

- Create guidelines for creation
  - o Common terms for public info

Outreach

- How to make it usable

FDL's

- Complain about what you're <u>not</u> getting

Thematic Access

- Not provenance

Versioning

- Metadata creation

FDL: Relationship and Role

- Still get print materials
- GPO facilitates training (print)
- Questions – How do we support GPO?
  o Produce guidance documentation
  o Plans for real world
  o Model agreements

How do we get prescriptive when we aren't currently allowed to do so?

Community: Actionable Opportunities

Red = +1   Blue = I would participate   Green = needs $

- Advocacy – raising stakeholder awareness (7,7,2)
- Lead where you can (2, 8, 0)
- Write recommendations document (6, 4, 2)
  o Lifecycle – agency pub guidance
- What is the imperative to document now (not end of term) (4, 2, 0)
  o In a model to last
- Tools (6, 3, 8)
  o Collection development
  o How to collect
  o Description
- Demonstrate use and access to creators (3, 2, 3)
- Environmental scan (8, 6, 6)
  o Who's doing what?
  o Where is it preserved?
  o Where/when is access available?
- Conduct an analysis of collaborative opportunities: Identify (8, 8, 5)
  o How can a group improve things?
- Build off finding from GPO – FRD research (4, 3 , 0)
  o Oct. 2017

1)General Questions

- What are the most difficult/intractable problems for public access to your info?

- Have you been unable to fulfill an information request due to lost bun-digital items?

- What is being done for access?

- How can participation be scoped/ scaled so many can be active, not just folks at organizations or institutions with later digital responsibilities?

- What are the copyright issues, if any, with preservation efforts?

- What role can/should commercial vendors pay to ensure access?

- Is there an existing group that can serve as "home base" for coordination? NDSA?

- Should we worry about duplicating the preservation activities of the Feds? (NaRA, LC, and GPO)

- How does our work relate to and complement the SERI work?

- What risk factors are of most concern under the new "regime" – and how does that alter our work?

- Will Federal Depository libraries be willing to collect digital content?

- Can we improve communications for groups like state archives, state libraries, state historical societies, who share an interest in digital government information preservation?

- How can libraries and archives help toward this HUGE effort in SMALL ways?

- What motivates people to care? Why should they care?

- Since we can't save everything – who is deciding what gets saved?

- How can we be effective in the short term and the long term?

Sometimes librarians scan materials for inter-library loan purposes. Could there be a place where these "quick and dirty" scans could be uploaded for later use and sharing?

2) Who are the content creators? (GPO – FRD 10/17)

- State

- Agencies, commissions, courts, contractors

- Where do content creators look for guidance?

- Federal agency staff must use whatever tools are appropriate for their work.

- Federal, state, local. Contractor? Consultant reports.

- What are their goals when they communicate?

- Contractors?

- State, local, federal, quasi-government agencies.

- What would motivate content providers to cooperate with this effort? (Besides money)

- Private firms hired by government to conduct research on an issue and make a report. (Sometimes copyrighted)

- Commercial vendors. (They are more of a PRODUCT creator using freely available content)

- How are decisions being made by creating agencies?

3) What are the content types?

- Do we care if content types are usable? Eg. GIS, AV, etc.

- All file types used by federal staff to do their jobs.

- How do we advise federal agency staff about sustainable file formats without telling them how to do their jobs?

- How to account for compendia of resources (from interagency task forces, public/ private collaborations)?

- What different modes of capture are necessary for different content types?

- EP6B, DAISY, Mp3, Mp4, PDF, PDF/A2-B, PNG, JP2, TIFF, TXT.

- Description of organizational structure?

- Do you keep multiple formats for the same content?

- Are different agencies producing different formats? And if so, why?

- What formats are being produced?

- What software are content creators using?

- What's protected by copyright or intersects with IP issues? Or PII (Personally Identifiable Information)

- What content types are functionally tied to format?

- How much content relies on proprietary standards?
- Legislative, executive, judicial, all are different, may need different approaches.

- What formats need to be available for immediate use? For transformative use?

- How is it described?

- Copyrighted (state) vs. non-copyrighted (federal).

4) Who is the audience? (For publications and for records)

- User story? User?

- Access demo with wire frame

- Focused, Audience/user, Community

    - Connect to advocacy work

-Preservation user (Future user)

- Needs assessments from users

- Current interests

- What's already been done?

- Researchers: Academic, government, commercial. Future agency personal.

- What common questions or issues arise with public access to government information?

- Who are your most frequent information requesters?

- Do you know about FDLP?

- Audience: Who is it? How do different users group use/access the information? How would this impact preservation expectations?

- How do we account for vulnerable populations, such as those who have experienced hate crimes, are undocumented, who will need immediate and easy access?

- State Government employees- to access their own institution history?

- Researchers engaged in data analysis of words/phrases occurring in BIG DATA collections of texts (i.e. Hathi trust)

- How can we include the research community (historians, political scientist, communications, etc.) in this work?

5) Modes of dissemination/ access?

- Gap analysis?

- Do you have a standard way that the public access this information?

- Discoverability? Do you want open index for local systems? Inclusion vendor systems? Summon, Primo, WMS.

- Versioning? How frequent? What's needed? Authentication?

-What information needs to be searchable/accessible by the public on familiar search engines?

- What formats can be repurposed on transformed for future reuse?

- Are you sharing or making metadata public?

- What systems are content creators using?

- Are publishing systems in-house or contracted?

6) What is the at-risk digital content?

- Non-static evolving data sets.

- Data, publications, and books. Web content, records, social media.

- What does Library of Congress have covered already?

- What does NARA have covered already?

- Undeserved digital collections stored in underfunded digital responsibilities.

- What risks are driven by policy? What risks are driven by technology?

- What did we serve in the "old days" – scope- do we <u>want</u> to save it all or just think that's simpler?

- What falls between publications and records? Policy?

- Preservation v. access. Some creators want preservation but don't want access (or public access). Sensitive classifies.

7) Current Preservation Efforts (Local, State, Regional, National, International)

- NARA preserves permanent federal agency electronic records.

- Role of FIPNET?

- Central registry metadata?

- What are content creators doing vis-à-vis preservation?

- Consortias, one-off efforts, state wide initiative or statutory expectations, any local government efforts.

- Do you have a good sense of your present capabilities to effect long term preservation?

- What gaps exist in current efforts? (Content preserved, agency representation, metadata availability)

- What content isn't being preserved on purpose?

- Identify digital repositories, have they gone through a TDR assessment?

- What are your pain points?

- How are you guaranteed content integrity?

- What roles do existing networks (Stanford, MetaArchive, Chronopolis DPN, etc.) play in this work?

-What are your plans to keep your content accessible by current technology?

- Do you have an auditing procedure?

- What are your obstacles to preservation?

- Can this effort include a concerted push to open up digitized historic state government publications in Hathi Trust? (to get copyright clearance)

- Muckrock.com for FOIA and sunshine law information releases (and denied requests).

- What kinds of description are needed to facilitate access and reuse?

- How are collections development parameters documented?

- How are they funded?

- Capture/collection vs. preservation.

8) Standards

- Better best practices?

- What standards are not being followed by the preserving parties?

- Is content language accessible and accessible to persons with disabilities?

- Standards vs. best practices (Guidance).

- FADGI standards

- Accessibility: S08, ADA, WCAG 2.0, AAA, AA?

- UECMA, other legal stats?

- NARA transfer guidance for file formats and metadata.

- Library of Congress sustainable format guidance.

9) Metrics

- Is it being used? (Stats like counter) What is the volume?

- TDR. ISO 16363:2012

- How do we measure good access?

- Are the intractable information problems lessened or alleviated?

- Does the public know about these efforts?

- Is your content accessible in libraries?

- What amount is created per year (and is it up or down annually)?

- What percent is saved?

- How many copies are saved (and how distributed) in current networks/repositories? (NARA, Stanford-LOCKSS, LC, etc.)

- What agencies produce the most? The least?

- Are the people who do this work receiving adequate support for their efforts?

- Page count in a trustworthy repository.

- How inclusive is this work?

- Content is available to people who need it in the format they want it.

Environmental Scan - Disseminate and encourage Response and Survey

1) BUILD on Internal work
- Here's what I know:
    o Encourage participation
- Define Environmental
- Scan parameters
- Draft proposal points

States/Gap

- Who's responsible for state digital records?
- Relationship with state IT

Preservation	Access
Preserving orgs getting the materials	FDL – getting and providing access
<metadata> - records	
Dynamic public info products	
Who's doing what?	Where do I find it?
Who has capacity, mandate responsibility	When can I use it?