

**American West Partner's Meeting**  
**October 24<sup>th</sup>, 2004**  
**Baltimore, Maryland**  
**DRAFT Meeting Minutes**

Attending: Arms (CA), Greenstein (DG), Willet (PW), Urban (RU), Tennant (RT), Chandler (RC), Lee (HL), Landis (BL), Dunn (JD), Johnston (LJ), Davison (SD), Madsen (CM), Brancolini (KB), Llona (EL), Lally (AL).

**Agenda:**

- **Welcome, Introductions, Announcements**
- **Review of Meeting Minutes New Orleans, April 18, 2004**
- **Status Report on American West Project**
- **Aquifer – update**
- **Scope and Content of American West Collection**
- **User / Service Scenarios**
- **Partner Metadata**

**Welcome, Introductions, Announcements**

KB: Announced IMLS grant to start a Masters Degree in digital library specialization at Indiana/UIUC <http://lair.indiana.edu/research/dlib/> (practice-/research based) – focus group later in the week at DLF. Will be surveying those currently working in digital library efforts.

**Review of Meeting Minutes New Orleans, April 18, 2004**

Minutes approved with minor edits (adding Ann Lally to attendee list)

**Status of American West Project (RC / RT)**

- Where we are?  
[https://diva.cdlib.org/projects/american\\_west/presentations/amwest\\_partners\\_rc\\_20041024.ppt](https://diva.cdlib.org/projects/american_west/presentations/amwest_partners_rc_20041024.ppt)
- Assessing value proposition.
- Focus on needs assessment and American West collection planning.
- Year One achievements:
  - Building partnership with collaborators
  - Completion of initial user needs assessment appropriate to collections/tools (18 interview with 45 users in California and Colorado) – ties into work that RT and others at CDL have done on NSDL science portal and a specialized portal for undergraduates. American West project has benefited from assessments across this broader spectrum. Review of principal findings.

- Importance of annotation tools (a very broad concept): narrative essays/lesson plans to contextualize the content (k-12/community college teachers); subsetting/federation tools to integrate with local holdings (librarians); search and retrieval tools/citation extraction (graduate students/university faculty).
- Curation of collections an issue; boundaries and scope/rationale.
- Functions: searching/browsing (topic, chrono, genre, location), selection, annotation
- Collection thus far (32.5 million pages of encoded text and 4 million digital images). Bulk is photographs, focus is specifically U.S. West (rather than broader American West); 4 chrono periods, pre-1800; 1800-1890; 1890-1940 (“post-frontier”-problematic name); 1940-present.
- Gaps: gender (lack of women’s experience in the West); race (supports Anglo-American western experience); missing U.S. Southwest and Hawaii, 20<sup>th</sup> century literature, transnational/borderlands West; music; organized sports; movies/entertainment industry; military-industrial complex; computers/cybertechnology. How to fill gaps? Harvesting 3<sup>rd</sup> party metadata/content; harvesting item-level metadata; harvesting static web pages (to get at more contemporary issues).
- Challenges to subsetting (user-driven construction of browsable views) – challenge that harvested collections have been built for local needs and to localized standards for metadata.
- Functional specifications for tools
- Collection planning/functional requirements.
  - Digital asset management framework (ARK, METS support, XTF (Extensible Text Framework – CDL’s Altavista – support for searching and viewing texts), Digital Preservation Repository for UC campuses soon in testing phase).
  - Harvesting and ingest tools for content and metadata (Voro EAD and Voro Basic); web crawling tools (need to be able to scope and define crawls); curatorial tools for service providers (includes guidelines for data creators and service providers; analysis, normalization, etc.)
  - Access tools (meta search engines – MetaLib: Exlibris, XTF: home-grown capacity)
  - Customization tools (for curatorial efforts and interface customization); branding views and surfacing curated portals for content access)
- Collection building and curatorial tools:
  - [https://diva.cdlib.org/projects/american\\_west/harvest/amwest\\_harvest\\_2004\\_10.ppt](https://diva.cdlib.org/projects/american_west/harvest/amwest_harvest_2004_10.ppt)
  - Metadata harvesting issues:
    - Sets, metadata formats, metadata artifacts, granularity

- Sets: make sense in local contexts but not necessarily appropriate to American West project.
  - Work with partners to create project-appropriate sets
  - CDL create tools for subsetting harvested content to isolate project-appropriate content.
- Formats: simple DC is simply too simple; need partners to surface something more than simple DC
  - LC surfaces MODS as well as DC (richer metadata)
  - CDP surfaces their own richer OAI metadata including refinements and qualifiers. Its called OAI.CDP. CDP can also surface SGML with qualifiers if we are interested
  - Need to reach agreements on more standardized way to surface richer metadata than DC for harvesting. Offering richer metadata is supported by OAI, but the format must have a schema.
  - Develop CDL infrastructure to normalize and enrich harvested metadata.
  - EL: With ContentDM there is a way to get at locally richer metadata content.
    - OSU –Terry Reese – has developed a tool using ContentDM config files to access local tags; UW will use this or provide XML export
  - Contentdm is reportedly willing to enable the application to surface richer metadata for harvesting.
- Metadata Artifacts (unwanted, unintended artifacts for project, even though they make sense in local contexts): idiosyncratic local practices, HTML code, anachronisms.
- Strategies:
  - Work with partners to clean up metadata
  - Create an infrastructure that supports metadata normalization.
  - PW: Kat Hagedorn is doing a lot of data cleanup work for OAIster.
  - RT: Have talked to Kat before, but maybe worth revisiting with her about what kinds of transformations and clean-up she's been doing. NSDL has done some simple cleaning (removing empty and meaningless tags) – have list of transformations that they do, and have worked with guidelines for their contributing partners. CA: Mostly what's out there is simple, and focused on transformations of simple DC.
  - EL: Would you ever keep the higher granularity?
  - RT: First rule is that you never get rid of anything. Currently working at idea/prototyping stage so just working through some of these issues and transforming them into actual infrastructure. If we get good for

enrichment/normalization activities, does CDL want to get into providing enriched/cleaned up metadata back to partners? Do partners want to harvest from CDL their transformed metadata? CDL is always Interested in OpenSource solutions that could then be shared with other DL partners.

- RT: would rather have idiosyncratic local metadata standard than not having rich metadata – as long as it is consistent it can be developed into a transformation profile that can be used anytime metadata is harvested from a certain provider.
- CA: wondering if LC should put up a variety of data providers to expose different depths of richness of extant metadata; LC would want to hear from CDL that we need a second data provider, i.e. a second base url
- DG: We are moving towards profiling from “trusted providers.” We would build a collection profile based on local schemas and we will return to harvest because the content is good and authoritative. The effort to develop transformation profile is worth it if relationship is long term.
- EL: Profiles for local metadata might be useful in the long run in developing metasearching schemes.
- SD: Sheet music consortium has been doing this sort of work. They have developed a profile to harvest the Australian national Libraries sheet music collection based on local schemas. BTW, they can’t harvest CDP because there are no sets; sheet music consortium has just started harvesting new libraries with new metadata standards.
- CA: LC is making sheet music available as its first native-MODS harvestable set.
- Granularity:
  - What is an object? Book vs. individual page?
  - Multiple values packed into a single field/idiosyncratic solutions
  - Determine with partners what is appropriate granularity for project.
  - Develop profile-based process for each provider for combining and breaking apart.
  - Subject terminology differences
  - Date encoding variances.
  - Mapping oddities: format into description; descriptions get stuck in subjects.
- RT suggests that CDL build a more widely available metadata harvesting suite of tools

- Analysis: what do you have? Encoding practices, gap analysis (e.g., missing fields), mistakes (e.g., mapping errors).
  - What can/should be fixed and how (transformations vs. bumping back to provider).
  - Creating procedures for fixing in future, either automated or human-assisted.
  - Some tools out there now: Spotfire Decision site 8.0 – point at chunk of metadata and it'll point out variances; helps you to visualize and pick and variance; for a given field, how many instances of that field exist; analysis of values that fall out of a given range.
  - RT has developed and is testing his own set of tools: Pulling occurrences of a given field, e.g. date field, to scan for variances; counting empty fields for a given field. Ex. of problem: simple DC <date> doesn't allow one to determine which is the original and which is the surrogate. Extracting instances of values of a particular field, looking only at unique values.
  - CA: would be useful for providers to look at their own metadata, e.g., looking at MARC vs. MODS records from LC (one transformation done in automated way).
  - RC/RT: It will be useful for CDL to provide reports for partners on metadata irregularities (based on profiles) as CDL does analysis of harvested metadata.
  - RT: reinforced that CDL always wants the richest format that a partner could expose. Metadata tools will have to deal with a variety of exposed metadata schemes.
  - RU: “trusted” vs. “untrusted”; must be able to do this with both (people as part of project as well as those that aren't at the American West table in order to attempt to fill gaps identified in the collection). DG: efficiencies in working with aggregators or with their sub-aggregators (e.g., CDP as metadata aggregators for a bunch of smaller partners)
  - RC: Do we have a sense of how we define “trusted” vs. “untrusted” providers?
- Subsetting of records: RT developing prototype tool to search raw harvested data and move appropriate hits into American West collection. Need to figure out a way to this kind of thing programmatically at some point – how far can you go without human interaction with subsetting? Process of moving prototype ideas into production will help resolve some of these issues.
    - KB: This will work better with some collections than others (photos with geographic term hooks vs. fiction collections where this metadata isn't predictable. Wright

American Fiction collection does not have subject headings).

- DG: Ways to identify gaps and go to other agencies as a way of taking what's been done and building on it to fill out a previously identified literature canon
- AL: Is this a tool that could be used by end users to subset from the collection?
- RT: possibly; will depend how it works with skinning and slicing interface tools.
- DG: how can we surface these tools so that users can play with them? Views we create won't likely be the views that specific users want, so can our view-creation tools be used in public interface?
- Metadata Enrichment: Adding fields or values, adding qualifications to some fields
  - Normalizing metadata to a standard state. Dates an especially good/problematic example
  - Add institution identifiers.
  - Add geographic coverage – some guesswork involved.
  - Mapping subject terms between thesauri.
  - Linking to authority records from OCLC or other services.
  - Exploding known acronyms that appear in records.
  - Querying gazetteers to enhance geographic metadata.
- Developing OAI best practices for both data harvesters and data providers; Birds of a Feather session at DLF on this topic.

### **Aquifer -- update**

- DG: Contextualization: Where American West overlaps with CDL and Aquifer
  - CDL:
    - What is the library after all the content is digital? Simply a gateway into vendor-supplied current content? User needs analysis shows lots of dissatisfaction with Google-like access to vast quantities of information: still a value for curated collections.
    - Pushed CDL into two areas that weren't originally envisioned ("value propositions"): Suite of tools for libraries to aggregate and build curated collections – a handicraft activity – what tools/infrastructure will enable curation to happen (by libraries, but also by end users for their own purposes – different scope/views of the same activity). Digital preservation: interesting because it is really about collection-building for volatile born-digital content and bring it into library environment and suite of tools that allows process of curation. American West project part of this broader suite of activities that represent a basic, fundamental reshaping of CDL.

- Where do you put the metadata that you harvest? Esp. if you don't want to have to build another database every time you start a new project. Building sustainable infrastructure.
- Aquifer (DLF initiative formerly known as DODL):
  - Took off in August 2004: focused on the future of libraries; themes parallel American West themes; leveraging substantial investments in digital libraries into an “enabling infrastructure that will allow us to build and sustain curated collections”. Enormous similarity.
  - Hopefully will provide deeper pool of development to build tool suite and surface more collections so we get more experience with these issues.
  - May have much more political baggage as it gets off the ground.
  - Gerry Persons' (spelling?) paper on integrating taxonomies into web services model for enriching metadata.
  - CDL sees participation in these initiatives as a way of getting more bangs for their buck, leverage investment in where CDL is going.
  - Hopes to see Aquifer look at areas that will enhance, enrich, compliment the work that is going on in American West.
  - Collection registry issues as different institutions get involved in these collaborations and start contributing collections to various initiatives.
  - DG suggests asking David Seaman to have DLF create some kind of collection profile standard for exposing collections for harvesting. Develop place where institutions where people can place their harvestable collection content where they maintain commitment to all best practices surfaced during AmWest and Aquifer projects that are appropriate to data providers.
  - PW: Michigan (and others) would like to have one known place to post information about harvestable collections, conditions, profiles, etc. and maintain them.
  - RT: OAI should develop more sophisticated (technical vs. descriptive) standard for surfacing collections for harvesting. Ties into DLF best-practice development initiatives.
  - JD/DG: Hope that Aquifer will focus on content files vs. just on metadata, which is the focus of American West project. Chance to help steer Aquifer in complementary directions. DG: Need mental map of the space we're working in and distribute the work between these big projects (so individual institutions can pinpoint where their local initiatives fit on that larger mental map, and also understand how they can leverage their institutional initiatives into larger benefits).
  - Participation in these larger, multi-institutional initiatives as a way of getting more access by end users and a better understanding of how they're using them; also really working

towards building tools that allow end users to find and utilize the collections we build in ways that satisfy their needs. Need more concrete ways to measure impact on end users.

- CA: Another project that can be related to AmWest/Aquifer down the road: national newspaper digitization project will be creating content of potential value to everyone. Use case scenarios for AmWest might be useful in designing these projects so that interoperability is on everyone's radar as content-digitization projects get underway – trying to get them not to focus on creating their content just for use in a project-specific box.
- DG: Need to role out work that AmWest is doing to other institutions. How to get this work onto higher DLF plane so that it is working towards broader protocols and standards – get big library directors to see the bigger standards-development process. Develop collectively a more compelling, bigger picture rationale for this work in terms of creating a broader shared vision for where libraries are going in the future.
- How do we expose what we're learning to faculty, other cultural heritage institutions? Need to find ways to broadcast issues, and standards development.
- PW: Need to involve AmWest collaborators more in ongoing discussion of issues (not just a meeting every 6 months).
- DG: Need to focus on project collaborators. What kinds of building blocks do the 7 of us need to work with our individual communities and institutions? If we can't do that, this hasn't been a collaboration
  - To reach out to faculty (uses)
  - To reach out to contributors of content.
- RU: In approaching the museum and cultural heritage communities, the term "curation" has a different, perhaps richer meaning.
- DG: How do we roll what we learn out into an enterprise- or institution-wide environment in a meaningful way.
- RC/DG: CDL will investigate more digestible ways of putting out information about progress being made and issues being raised by AmWest project.
- LJ: Fedora as example – developing ways of exposing metadata for harvesting, but need to be told HOW to expose, what are best practices being developed?
- Track the discussion points that need to happen between AmWest and Aquifer so that they aren't treading same territory but are working in a complementary way.
- AmWest collaborators need to keep CDL informed about what they need at their own institutions; what discussion points need to be tossed up to the project.

- PW/LJ: need an online forum in which issues can be discussed as they arise;
  - RC: Asked collaborators to provide feedback on what they want.
  - DG: Maybe establish a single communication/discussion venue in combination for both Aquifer and AmWest projects.
  - CA: Need to think in terms of metadata issues, but also user issues that have been explored in AmWest. Probably a different set of people interested in these 2 issues.
  - KB: Much of the user studies stuff done by AmWest is completely generalizable and of interest well beyond this specific project.
  - RC: Maybe press releases with big points of various projects with pointers back to more detailed reports.
  - DG/CA: Need to roll together best practices initiatives together that are underway to promote interoperability. These include:
    - DLF / IMLS OAI best practices
    - NSDL
    - Oaister
    - CDP
    - UVA
    - OAC
    - RLG – Cultural Heritage materials – Guenter Waibel has asked for comment; will be published in December
  - CA/RT: Need to keep track of enrichment that has been done to metadata during project. Store it in some way that identifies it easily.
- **Scope and Content of American West Collection (RC)**
  - CM: Harvard is willing to digitize additional collections for American West Project
  - RU: Other projects affiliated with CDP are creating content
    - Rocky Mountain Project: 2,000 images on New Mexico
    - Western Trails: Z39.5 compatible but Nebraska could be OAI compatible and Kansas using a MARC system could be OAI compatible
  - CA: Bob Allen at Drexel is developing a timeline / gazetteer that would be worthwhile looking at for the project - ("Developing a Query Interface for an Event Gazetteer")
  - CA: University of Chicago collections in LC American Memory are organized by state
  - RC / Group: Reviewed Geneva Gano History Scope and Content and Gap Analysis documents
    - Geography: provide access by state and regions (balanced by chronological period)
    - Gap Analysis Genre – group suggested doing curated views in these areas using Geneva Gano model
      - Music/ Sheet Music:

- UCLA Fronterra Collection – could be made available
- UCLA: Tony Seeger, faculty at the Ethnomusicology Archives could be tapped for sources for folk music that is freely available
- LC: (2) audio collections at LC -- CA will investigate making metadata harvestable
- CDP: Nuropa (sp?) audio collection is a possibility
- Internet Archive
- Ann Lally has knowledge about Public Broadcasting – performances of contemporary music
- Artworks
  - National Museum of American Art
  - Seattle Art Museums – recently received an IMLS grant
- Literature
  - Pursue a planning grant for the canon of American Literature
  - Martha Brogan Study on how to make a collection of literature available
  - Gayle (vendor) is working in this area
- Maps
  - Discuss with David Rumsey
- Natural History
  - UVA has 1000's of plant specimens
  - Missouri Botanical gardens
- Gender
  - CM: Harvard will do some analysis of collections they have relating to women – will provide descriptions to CDL for possible inclusion – have Geneva examine?
  - CA: Mystic Seaport collections concerning women; LC has the marc records and the digital content – have Geneva examine?
- Southwest
  - UVA: faculty member building a collection on Chaco Canyon
  - CDP: new director has contacts with Arizona
  - LC: has a University of Texas American Memory Collection; we could contact Harry Ransom Center (Rich Oram?) at Texas and / or Bob Martin, Director of IMLS
- Topical Views Discussion
  - Term level looks good

- Suggest having list reviewed by other historians and k-12 teachers
  - Examine other examples:
    - Corbus stock photo
    - CA: will send University of Chicago terms
  - LJ: UVA has a taxonomy – she is happy to share with us
- **User scenarios (RC)**
  - Scenarios try to link services required to metadata requirements and hooks needed to make them possible.
  - Curation (aggregation) tools vs. raw search/browse tools vs. pre-packaged portal views.
    - RU: relationships/links between objects in collections and learning resources that use them.
    - RC: role for reuse/access to learning objects/resources via AmWest project.
    - CA: faculty roles vs. reliance on graduate students.
    - SETS: 1) OAI protocol-defined sets of objects made available for harvesting; 2) Preservation of object relationships (how do multiple items come together as an object)
  - Citation structure: Is there a preferred citation structure for the kind of content in AmWest collection. Is it in metadata already or how would it be constructed?
    - DG: Maybe libraries need to own the metasearch layer? More control over what content can be presented to all users (problem of free vs. licensed content).
  - Copyright Information: METS-record approach in OAC content: 1) copyright status indicated at the item level and 2) a link back to repository page for information about permissible uses.
    - CA: At LC easier to say nothing about copyright than to get the lawyers involved. LC could, at minimum, put something in the SET description that could serve as a page to link to that would provide some SET-specific copyright/use information.
    - RU: CDP recommending participants use a URL for copyright information so that it can be changed on the site when necessary without requiring update of individual records.
    - RU: Can AmWest as service provider build some kind of click and dump form that goes off to the right people to provide further information.
    - DG: Approach might not scale well but is worth thinking about.
    - DG: Things to formalize shortly after the meeting.
    - Framework in “Service Scenario” table for a minimum metadata profile for AmWest and maybe other aggregators; table provides rationale in services made possible for minimal set of metadata.
  - RC: Timeframe for AmWest harvesting in production is March 2005.
  - RT: Need good set of samples of richer metadata before then.

- DG: Work backwards from end-of-March harvesting in earnest to create timeline for interim richer metadata and bilateral discussions with collaborators about what is needed to make that happen.
- RT: Support in prototype for one collection and/or one data provider to look at analytical data for just that subset (collection/provider).
- Strategy:
  - Build spec for metadata requirements
  - Have conversation with partners regarding timeline to create richer metadata – and who does what (service provider / data provider)
  - Enrich the metadata
- **Examining Partner Metadata**
  - RC: When would all partner content be harvestable?
  - RT: Want access to richest metadata vs. just simple DC, if others are available.
    - U Va: will know more about schedule next week; but will give us MODS.
    - UW: make available raw XML files for metadata; might be able to do some mapping to MODS prior to the dump.
    - Harvard: Spring 2005 (?), with more detailed schedule for OAI harvestable stuff by December.
    - LC: California As I Saw It will be available next year; (extra information supplied at SET level
    - RT: OAI harvesting of richer metadata is best option (can set up profile-based process), but can also take native metadata files if it means AmWest can get access to richer metadata.
    - IU can have MODS records for Cushman photographs in next several weeks. For Wright should be easy to have MARC XML soon, since descriptive metadata is already in MARC.
    - MARC XML will likely be transformed via MARC-to-MODS.
    - RT: If partner has choice of exposing metadata in local format or MODS, as long as there's no loss of data in that transformation; also would prefer MODS to MARC.
    - DG: Need to include this in an initial set of preferences for AmWest participation.
    - Harvard: Easy to get data in MARC, but might be possible to deliver in MODS.
    - DG: Framework of preferences also gives CDL ability to list suite of tools available for delivering metadata in preferred structure.
    - CDP: Easy to give richer CDP OAI profile metadata, but RU needs to look at the metadata issues.
    - Michigan: PW will check on how easy it would be to subset Michigan's one big set. JPW told RT that John Weise would be good to work with on strategizing ways to subset out the stuff that isn't appropriate for AmWest. MOA stuff has MARC

- records, so should be easy to just deliver in MARC XML. Image metadata will probably come out in DC (might be the best they can get).
- RU: CDP OAI is probably the best metadata we can get; CDP would need to contact the participating institutions directly
  - UWashington can provide an XML export and possibly map to MODS.
- CA: American political prints collection soon online at LC, includes political cartoons that might be useful for AmWest, but would probably require a human culling through it to select.
  - Need to get as much as possible built to as high a specification as possible by end of year 2 so that we can go into year 3 really working on getting people to use this stuff and give us feedback on it.
  - RC: How easy would it be to develop some kind of standard for set collection descriptions.
  - RT: Dovetails into DLF OAI working group.
  - CA: need someone to cobble together schema for OAI set descriptions.
  - LC using Dublin Core schema as framework for OAI set descriptions for AmWest stuff to be harvested;
  - RT likes the richness of LC's set descriptions for harvesting. These are easily browsable on OCLC <errol.oclc.org>.
  - KB: Doing a lot of post-harvest processing makes it more difficult to re-harvest.
  - JD: Major protocol issue is deleted records – potential to harvest thousands of deleted records that you then need to cull through to remove deleted records.
  - RC: Set descriptions, would be good to have robust ones. Could we ask partners to go through and use LC model that CA has talked about as framework for these?
  - Something that should emerge from AmWest: minimum standard for OAI set descriptions.
  - CA: might think about set description differently if it will be used for public consumption as access mechanism vs. just by the harvester for data about the set. Possibly look at model for set descriptions that RLG is using for Cultural Materials set description for harvesting.
  - KB: We'll also need a meaningful way to describe subsets of sets so that you can say what went into subset (parameters for selection) e.g., Western lit subset of Wright collection.
  - DG: Interesting subset definitions only emerge across the AmWest aggregated stuff. Content must have period and geographical coverage and a list of other topics associated.
  - KB: Easy for photos, but maybe not so easy for things like sheet music or fiction that don't necessarily lend themselves well to topical breakdowns. You need to be clear as to what is it that you're interested in by getting at geographical access (e.g., published in West

- vs. about West) – someone needs to come up with decision-making parameters for selection for this kind of material.
- DG: almost suggests that simpler might be better in terms of creating set description.
  - RC: We need to ensure that there is enough metadata to make an object discoverable and useful for further annotation by end users, and that there are links back to repository for more contextual information and for basic information about rights.
  - DG: postpone closure on what topical metadata required for sets for literature, art, and sheet music until after we get a curator to come in and do a Geneva-like treatment of these genres.
  - RU: Are there benchmarks for content files to be delivered to end users? Not really extant outside of individual projects/repositories.
  - RC: We've been wrestling with benchmarks for metadata more than benchmarks for content files.

### **Recap of Decisions / Action Items**

- Draft a collection strategy to fill gaps identified in History curated view
  - Tight Scope
  - Develop a plan to circulate to ask if institutions can participate
  - Requires Contributor's communication plan
- Develop multi-faceted communication strategy/model
  - Partners (systematic)
    - Website
    - Listserv
    - Regularly scheduled conference calls
    - Develop "stories" or press releases about project for partners to leverage in their institutions to further collaboration interests
      - First Example: De-construct the Hewlett 1<sup>st</sup> year report into re-usable chunks for partners in library
    - Post the Collection Development Strategy: History curated view including Scope and Content and Gap Analysis
    - Post collection spreadsheet to serve as collection registry
      - Avoid duplication
      - To facilitate partners adding new content (further digitization)
    - In communications with partners indicated where these activities overlap with Aquifer work
  - Future Contributors of Collections (see Aquifer discussion for more details)
    - Best practices to be observed
    - Strategies for contribution

- Users of Collection (with an eye to future evaluation/testing – see Aquifer discussion for more details)
- Users of Tools (with an eye to future evaluation / testing – see Aquifer discussion for more details)
- Consider five additional curated views of the American West Collection
  - Art
  - Literature
  - Maps
  - Natural History
  - Music / Sheet Music
- American West Project has established a data profile requirement for contributors
  - Expose Native metadata which will include:
    - dates
    - location
    - if supplying MARC, we would prefer MODS
  - Create Sets
    - CDL will work with partner to create project-appropriate SETS
    - CDL will create a harvesting infrastructure to enable subsetting
  - Provide Set Descriptions
  - Copyright information – what is our requirement?
- Align American West Project Goals with DLF Aquifer initiative
- American West Partners expressed interest in using curation tools to integrate selected subsets with “their licensed content”
- Harvard has offered to subsidize creation of additional content by offering their digitization services
  - Work with Christine to identify other Harvard content of interest
  - Consider if there are other UC collections or other partner collections to digitize?