CCBER Workshop

The idea is to get the highest amount of exposure as possible for things we have stuck, all kinds of institutional docs and field notebooks.

Thank UCSB Library, Mirea too.

Peter Alagona
There and Back Again
UC NRS History and Archive Project
Project 5-6 years, meta archive for UC NRS, custodian of the largest system in the world, 5 x size of SF, all over state, earliest in 1930's, wide variety of materials, in all conditions.

Peter gave presentation back in fall for 50th anniversary of NRS, father is Joseph Grinnell, first NRS reserve Hastings, wouldn't be realized until a century later, which is today.

Building archives, collecting specimen collections, accrue over time by collecting.

Fast forward to today, kinds of reserve records, many, as a whole contains an abundance of wealth for the environmental history of CA

Land use records prove essential to unraveling environmental relationship

Gold Standard, Harvard Forest Data Archive, David Foster director doing it for over 100 years, excellent job.

Stanford has done an excellent job with Jasper Ridge Reserve, site worth that much more for researchers today.

NRS, results to date are mixed, are we living up to Grinnell's vision, beyond Hastings?

For some reserves, someone gets interested and took it upon themselves to work

Off site collections very important to some reserves

Historical Preservation Efforts, around the state there have been efforts for environmental history and environmental science, working to establish digitized databases of specimens

Wieslander Vegetation Type Mapping Project
http://vtm.berkeley.edu/#/home

California Digital Library, uncovering California's Environmental History Collections
http://www.cdlib.org/cdlinfo/2009/12/07/uncovering-california%E2%80%99s-environmental-collections/

Grinnell Resurvey Projects can go back to same sites because of preservation of meta data
http://mvz.berkeley.edu/Grinnell/

UC NRS Data Infrastructure Projects

Boyd Deep Canyon Reserve, got idea in 2005
What is role of university run research station in shaping environmental policy and history?

Tried to develop meta archive, Laurie and Tim and Andrew involved, conducted basic archival surveys, at 39 reserves, UCOP and campus archives related to UC systems. The idea was to document and catalog NRS materials
To Make them more accessible
To Conserve for future use
To Promote the collection of additional materials

They created collection records for around 80 groups of NRS historical materials not previously identified or described

Connected NRS staff with university and museum archives to transfer sensitive materials for permanent preservation

Facilitated the creation of new collection records for these materials on the on line archive of CA

Developed a web portal that will serve as a clearing house for historical information related to the NRS

They will be augmenting this site with off site collection information

Producing a set of guidelines for reserve management to develop their collections and manage future materials

OAC
Guide to the UC Riverside NRS records, connected to UC archivist and produced full collection records

Some materials at Department of Special Collections before the addition from the reserves

Robert Norris was a Geography faculty, Robert M. Norris Papers MS-13, retrieved and worked with collection records many of which pertain to UCSB reserves, history

Brand new web site, reserves, basic description and links to resources, direct links to on line collections and some off site collections as well, abstract and tags, location, where to find it, creators and contributors, types, arrangement and physical descriptions

Collections, can go into collections and see what is held in the collection, let people know that things exist, if can't find than might as well throw in dumpster

Becomes part of infrastructure and history of the reserve

Challenges in developing an NRS system wide archive

1) convincing reserve directors that collecting and curration records is an essential part of their job
2) Compelling researchers to submit their data and publications
3) Reconstructing histories after the fact in cases where materials were not gathered in real time
4) Figuring out how best to manage materials on site at reserves
5) Deciding which materials to transfer to campus archives
6) Determining which materials to digitize and how to make that happen
7) Developing better records management strategy for capturing current and future data
8) Instituting a common culture and set of best practices across a diverse and widely dispersed system

Holding paper on site is not a good solution, we have to think about digitization. Labor, space, access, permanence. Hopefully digital materials have a sense of durability.

Resources, we need a community of volunteers, to develop a sense of importance that this is a knowledge base that develops over time.

Question: Took meta data and put into 80 different collections. Meta data collection effort as first step. Will encourage digitization.

Can you sign up as OAC, any institution can contribute, yes, set up relationship with them, can contribute records to that database, all kinds of special collections are described.

Management document, draft isn't really ready, going up to SNARL and talk to new reserve director and get on agenda for manager's meeting, UC records management policy, reserves have and need, feedback from managers. Document very useful.

Active vs Inactive materials, not being used for regular use, active consult regularly. Pass legal threshold, document active, inactive, beyond concern of law, end up in dumpster or part of collection.

Field Notebooks, copyright and permissions, intellectual property issues. Donor correspondence, real estate transactions, who owns land and what way, rare and endangered collecting black out. Land ownership gets you to land use, can be important.

Shallow dive to deep dive, basic entry level way, first step in what would be longer process.

What are the next steps for archive project database?

Trying to make sure things don't disappear, curated and preserved and try to get them on OAC. Look for off site collections. Discover and linking them into the web, one stop shop for what's out there.

Michelle Koo
Christine Fidler UC Berkeley

Processing archives from research field stations:
Cataloging material and digitizing data for integration and visualization with other materials

Collection policies, unique record with intrinsic value that are not organisims, interesting oddities, wide range of materials

Workflow, similar to what collections managers do, appraisal, accession, processing, critical to document provenance, there is a formality, then there is processing, survey, arrange intellectually, not physically, preserve original order, describe and catalog, preserve materials get rehoused and paper clips removed, make them available through discovery like OAC, looking at other methods of delivery and try to understand what researchers are looking for.

Movement to item level records, find a balance, taking archive in new direction.

MVB
Field Notes, bound 700 volumes
Unbound 200 volumes
Personal papers

Correspondence
Annotated maps
Archived images
Reprints

Can develop policies together and workflows together with archivist and researchers.

Joseph Grinnell's Vision, assuming that our materials are safely preserved, knowledge and resources to preserve from archivist.

Intern processed collection at Hastings, exposed to elements and pest damage. Work with manager at Hastings and removed and brought to museum. We know that materials they create now will eventually go to the museums. The materials are very vulnerable, to establish a repository is crucial. Paramount to establish partnership, materials come to museums and are curated and put into finding aid, intellectual guide to collection OAC. Best practices and description for archives, encoded XML file, archivist should be able to do for you, some training. How to relate our collections, how to connect our data, describe our relationships, expose relationships, DHL and field book registry larger preservation challenges, what is the lifespan of digital media, analog will last longer worth taking steps to property house and preserve.

Preservations Resources, Preserving archives and manuscripts, archival care and management, society of American archivist,
Developing and maintaining practical archives, Gregory Hunter, Western Archives Institute, two week training program, June, sign up in March.

CLIR grants to MVZ Archives
2012-2015 Hidden Collections
2016-2018 Digitizing Archives

Cal-Photos, A Century of Natural History visualized, OAC

1st grant in 2003, NSF to start scanning archives, specifically our field notes. Efforts to digitize and integrate were quite early, in charge of manage archives, trained archivist like Christina, 2012 jump start cataloging at Hastings, Hidden Collections, get that to OAC has added to value and perceived value to administration

Digitizing archive grant this year, partner with like-minded archives, digitizing field notes that didn't get done over a decade ago, haven't embarked quite yet. Also have an IMLS grant to refurbish room in MVZ used as storage facility to really create a true archival storage space.

Part of Hidden collections, create a variety of database integration, Artos, web portal and read field notes in their digitized organized fashion. Search functions allow you to look at original materials without having to come to museum.
Eco reader.berkeley.edu

Standardize meta data files ingested into eco reader and help organize scans and digital service.

HOLOS, rescuing and Integrating Biological and Environmental Data in the Face of Global Change

100+ years of data collection
Field notes
Species Lists
Photos
Climate Sensor Arrays
Experimental Data
Berkeley Eco-informatics Engine

Explore Tool, expose all of diverse datasets, points over raised, combination of specimens and photographs, references by collections, can allow user to come in and use.

Can quickly jump to NRS sites

Informatics Manager, Becca Fenwick, working with for NRS web portals for individual field stations

Show Field Book registry, IDigBio, our own field notebooks, there is the Smithsonian Field Book Registry, growing trend to search field notes, 9 partners to grant all ultimately live in DHL portal, own distinct portal. All on Github right now.

Meta data, archival description of field notebooks, digital object record, DAC to create it.

Question: What went into making the digitization project happen? What advice to proceed?

Couple different aspects, digitization archival material, have to be catalog first, however also in the midst of working on eco engine, research questions helped prioritized, species to collection event, species or place. Brand new work, same sort of prioritization join up the proper-curated collections with the right material. Focused on 4 NRS reserves, censored data another separate system. Working with ANR, couple of grants through them, sending students out to do tree specimens.

Few of reserves want to keep their own material, creating database portal for them, ARCTOS, works out for NRS and curation help from students.

Hastings 30 banker's boxes took 9 months, just cataloging, none of those are digitized. Point count data in written format, transcription project that helped digitize some of that transcription into digital format. FROM THEPAGE, active in archive. Can talk to him about projects, active participant in IDigBio, lot of potential here to extract data from actual archive.

Question; Reprints, digitize and get ride of originals.
Most of the monographs not digitized and have a lot of grey literature.

Jon, Map librarian at UCSB Natural Reserve System Society of American Archivists, good resource

Manage really large archive of Ariel photographs, part of effort to find off site collections

Individually one by one NRS has already paid for these, 90 percent of users are off campus

Oldest air photo 1858, 4 years after invention of photography
1906 image after earthquake, restored ands canned at high resolution
Fairchild Aerial Surveys oldest surveys

2.4 million physical images
1927-2012 physical
born digital since the 90's
scanned which physical don't exist

scan at 600 DPI, some 1200 DPI
Grayscale target

Digital collection created upon demand
UC affiliate, can give you image
XL 1000, Epson Scanner, film is light source adaptor to shine light through it
Move to camera capture,

Physical Storage-Film Storage
Nitrate film, very flammable
Acetate film
Acetate polyester color

Expired Nitrate film, must be treated as hazardous, if you film is pre-1950 and it does not say safety film it is likely nitrate

Mapping projects, not sure what to do with them

Vinegar syndrome
Affects post 1950 and pre1970 acetate film stock
2012 Vinegar assessment, litment paper stick test strip can evaluate how long film will live, has to be less than 60 degrees

Haynes Grant
Collection sat largely untouched since 1980's, archivally re-house and create finding aids for 40,000 pictorial images, can assess as budget allows
Start to transfer non-aerials

Framefinder, how do we spacilize this, centerpoints for images with links to their scans
A new offline geodatbase and a legacy MySQL driven website
Our main access mechanism
Pre 1940's only more indexes than photos, some photos gone or destroyed

Center-point for all and get list from point, set level and image level records combination to find in collection

More visualizations coming, airscalifornia.com, can visualize decade by decade for whole state, charts, graphs, maps

Interdisciplinary Research Collaboratory
Drop in computing and data access
Social Science and Spatial data services
Workshops and Community
Research Data Curation Service
Consultation on best practices and disciplinary experts in spatial and humanities

No map services at library.

Born digital backups-big issues.

Finding aids, image by image record of what those records are, five lines describing what photos are, log book able to scan and OCR.

Photograph map, controlled room, use Photoshop, take fewer images and stitch together. Some equipment becoming more affordable, state library talking about investing and driving around and set up to digitize collections. Could be an IMLS grant. Drum Scanners, for maps, hesitant to do with paper or anything old.

---

Patrick Randall, Biodiversity Heritage Library

Expanding Access to Biodiversity Literature
BHL
Patrick Randall
Increase access to biodiversity material on line by making BHL an on ramp for content providers who want to contribute to DPLA

Digital books, adapted to include book like items
50 million pages of material

DPLA is US material BHL is world wide

Hasn't been mechanism to get into BHL, no funds or resources
Grant DPLA grant to worthy contributors, offer assistance to digitization

BHL an on ramp for DPLA, everything in BHL will be harvested in DPLA and when finding, may already be material in BHL showing up in DPLA

Specific goals, content provider for researchers all over the world, greater diversity of institutions, serve as aggregator, preserve and provide access to small collections and improve and provide meta data

100 first time contributors, legacy literature, everything in public domain pre 1923, starting to get post 1923, new with this grant, taking lead on licensing

Get in touch, figure out what want to contribute, provide funding, initial amount of $500, internet archive to do work, institution can send 4000 pages, scanned and digitized by internet archive they would pay and put in their collections.

How to prepare to get your data into BHL, scanned images and metadata together.
How it might apply to archival material. Patrickrandall@fas.harvard.edu
Mariah Lewis, mleweis@nybg.org

New York Botanical Garden

Metadata, content
BHL software can send directly to BHL
MARC Record
MARCXML
To convert files, alternative ways to upload CSV files

Title level metadata plus item level metadata, plus page level metadata equals one digital book or volume

Macaw, BHL internet archive for harvest into BHL, facilitate files to export, cloud version, BHL members and affiliates

Create new, manually enter item,
Upload pages and enter page level metadata

Scanning deliverables

Curation, accurate metadata is the key to discovery, post scanning is important part of process

Edit through administrative dashboard, curation editing, good curation

BHL selection criteria, scope of BHL, defined very broadly, biodiversity related, valid item types, don't' take photographs, artwork, maps, specimens, beginning stages to accommodate artwork and photographs, can't be a duplicate, have to have a MARC XML record, Leslie Peroia at Smithsonian, metadata from scratch for field books, have to be able to create one, image quality has to be high, open access in public domain or written permission from owner.

Get text files, PDF and images, creative commons license, no qualification on how stuff can be used.

Explain relationship with Internet archive, any drawbacks to outsourcing, work closely, all images hosted on IA, BHL has own collection on IA except all metadata isn't there, IA does good job scanning, always potential for abuse, outage at IA.

Provide training, csv template, or Mccaw , dropbox image files

Requirements, fee 10K, own collection on BHL, have BHL homepage for institution, discount on scanning on IA, can serve on BHL committees,
Link BHL membership page, BHL affiliates, some of same privileges, great way to include everyone, discussing how to include images, through other special project, paying you to digitize so we can put on BHL. Trying to make it easy for non BHL members to get stuff into BHL. If you have monographs or journals can harvest title record from catalog and goes to BHL and then do page level data themselves, traditional material very little have to do as long as catalog record somewhere, as long as we can look and find somewhere, don't' have resources to catalog, have to have MARC record to be in BHL.

License record, documents, haven't figured out yet, persons passed away have field notes if donor agreement does not specify BHL then need explicit written

permission from next of kin, obstacle, have to do the leg work in order to get it up on BHL,

Bodega Marine Lab, collection, permission to digitize, digitization release signature. Make sure releases are done so don't end up with more problems.

Send in list of material if not sure is appropriate for BHL, do take institutional reports, single page or single documents can't really do, will try to make it work within model we have now.

BHL have page level records that can link back to our collection records and have more availability for students to do that, attractive for field notebooks not just whole document level, indexing of scientific names by OCR automatically indexed, let users add their own directly in form of tags and comments, all searchable, can pull up exact page very useful for researchers.

End result of where it's going to live will dictate, how's it going to be discovered and what does that discovery tool require, how it's formatted and what it needs will be dictated. See if it has guidelines on how to create the metadata. Harvesting OCLC catalog record in MARC, original cataloging.

Laurie, UCSB library, partner with librarian if you need expertise of MARC catalog. Different materials are being cataloged and described in different schemas and don't all play together yet, look at type of material, that information will be uploaded to certain system.

Field notebook page to specimen, can create field in record and link some of literature in specimen. Depending on system you use.

Almost a different standard for every type of object, different stages of development. Community practice around them, difference between standards.

Archival standards have been around for hundreds of years, profession lengthy, field notes are not treated as inactive record, used more frequently than regular archival materials, less likely to be deposited into institutional library. Room to create collaborative standards to try and make it available to broader community.

The field book project, wikispaces.com, Smithsonian institution archives.

Archives because amount of data being deposited movement is towards collection level, folder level, depart from normal to item level to capture geographic location, what level should we be capturing, not details that would normally be cataloged, collections do warrant, great conversation in this community, moving in more itemized level of description.

Set of filters, variable details, resource limitations, level of collection record, level to box of boulders to specimen, item level, then you go to digitization, different type of users going to use differently. Think about  it from both ways, archives perspective and user perspective.

Record level access, care about species, locality and date, if can't get there, won't find it, distinguish field notes from all kinds of uses.

If you have information about other resources that you use send to list serve.