2014-09-11 Supporting research with Digital Asset Management systems

Background reading/preparation for this meeting (from email of 4 Sept 2014 to Reading Group listserv):

Please review the following brief articles / report sections prior to the meeting:

DAMS Overview
=================================================
Systems for Managing Digital Media Collections (JISC)

Geared to researchers' and librarians' use of DAMS
=================================================

==> Towards a Virtual Environment for Supporting Research Activities at the Smithsonian (Thornton Staples)
==> Choosing a System for Managing Your Image Collection and Case Study: History of Art, University College Cork (JISC)

About DAMS technologies
=================================

==> A brief overview of Shared Shelf (History of Art VRC's management system); and an article discussing the use of Shared Shelf
==> UC Libraries Digital Collection (UCLDC) Overview
==> Executive Summary and Report Background sections of Proposed Model for Systemwide Digital Asset Management System (DAMS) with Discovery and Display

Optional reading/viewing -- some of it lengthier than the pieces above -- includes the rest of the UC Systemwide/CDL report (Proposed Model...); a 4 min. UCLDC Year In Review video; and UCB's Digital Collection Development Plan Task Force Final Report (2009)

Warren Hall access: For those who do not have keycard access to the building, please take the elevator to the second floor (stairwell door requires keycard). Before noon, let the receptionist know you're joining the Reading Group in 200C and s/he will let you in and show you the way. After noon, look for a sign next to the receptionist window to the right as you exit the elevators. We'll post a note with a phone number that you can call or text, and someone will come out to open the locked doors.

Please join us Thurs 11 Sept for a Research IT Reading Group on Digital Asset Management Systems (DAMS). We'll meet in 200C Warren Hall from noon until about 1pm (building access info below my signature).

We'll explore how DAMS functionality supports research (and teaching) via the library and otherwise; and some approaches and solutions to DAMS implementation. Discussion will be facilitated by: Sherri Berger (CDL); Lynn Cunningham (History of Art Visual Resources Center); and Mary Elings (Bancroft Library).

1) Aspects of DAMS that are important for researchers?
2) UC DAMS - systemwide; interest in providing larger services that aren't localized to units, but overreaching
What are opportunities and risks with larger shared systems?

Lynn Cunningham—
VRC --- largely support pedagogical visual needs, images used in classroom
15 permanent faculty, 16+ adjuncts, visitors, GSI
Image intensive, 20-80 images per class, couple dozen classes per week
Build digital database; 70k images, adding 8k assets per year
One of 9 VRCs across UC system
DAMS as bridges: reaching larger audiences, can leverage work across different campuses
Coordinated all VRCs in VRC Shared Images project
Picked Shared Shelf, hosted by ArtStor
Web-based media management software
Manage, store, use and publish faculty personal collections as well
Sharable campus-wide, system-wide
Integrates with ArtStor; publishes to ArtStor so there's a meta-search for faculty
Level 1 remote digital preservation
Support images, audio, video, document
VRE Core schema; integrated Getty vocabularies
Also customizable, allows you to use local fields as well
Shared work records: if someone at Princeton catalogs something, can link to that directly
No OCLC for images, this is first step for sharing records
Artstor — 2M images
Options for publishing to Omeka, open API, OAI harvesting
Has pedagogical tools built in to ARTstor interface
Can create a folder with images, then download them automatically to a Powerpoint
Create and share folders with students
Challenge/opportunity: collection awareness campus-wide
Integrating with other campus search interfaces, like library OPACs
Opportunity for integrating LOD content for researchers
Sherri Berger, CDL
Product manager for UC Libraries Digital Collection, 2-year project on shared infrastructure for managing and exposing unique digital collections
“unique resources” — libraries have 24k unique collections, things held in special collections, but also maps, newspapers, oral history, researcher data
Lots of materials where we are the only libraries in the world that hold these resources
JISC Digital Media “Creating digital media is the easy part of building a collection. What is often harder is to establish an effective way of storing and providing access to your collection.”
Infrastructure has been uneven across campuses
6 libraries have had no robust way of managing collections — been making do with random systems
Hard to manage current collections — hard to edit metadata, expose online
Hindrance to building new ones
UCLDC: suite of services, system-wide, three components: manage, aggregate, share
Management is discrete part of technical stack
Aggregation, exposing to the web - different pieces that sit on top of it
Using Nuxeo — content management platform, basis for CollectionSpace
Six campuses will be using the system, in production now, released in July
Currently working with those campuses to try it out, put collections in, providing feedback
Lots of work went into improving default interface — making it a pleasure for libraries to do their work
Happy to talk about other components if interested in stack for aggregating and exposing collections
Lynn: Relationship between ARTstor and Shared Shelf? Made by the same organization
Also provide Shared Shelf Commons, public collection (e.g. if you have something archival)
We do a lot of surrogate things, don’t have copyright
But if you have a unique collection, could publish to Shared Shelf Commons
End-user won’t know distinction between what’s in ARTstor and what Berkeley has specifically contributed to it
Shared Shelf is backend / cataloger end; researcher only needs to know about ARTstor
Sherri: Different sites; would have to go to ARTstor to access visual resources of UC system, and CDL site for other resources
Lynn: For things that could publish widely, will go route of dissemination, but need a wall for most content
Always the issue of not wanting users to go to multiple portals
DPLA, Europeana trying to solve that, mediating it at national level
Sherri: Aggregate part of stack, that’s where those materials are coming together
Sending materials up to DPLA from that stack
Patrick: Most of stuff CDL is ingesting doesn’t have copyright constraints?
Sherri: No, do have copyright, but not as extreme as Lynn’s cases
Initial phase of project; new public interface coming next year
Focusing on public materials (not necessarily public domain, but could at least be shared widely, used under fair use)
Technical constraint— will take longer time to build a system managing user restrictions and access levels
Wanted to launch big, why put a lot of time upfront in materials you can’t expose at the beginning
Will be able to accommodate access restrictions down the line
Jenn: Lots of discussions around DAMS for educational material
ETLG discussing learning analytics, how people use materials
What materials do people use?
Sharing materials — this for us is different distribution methods (multiple learning management systems, etc.)
Having same conversations with people talking about digital asset management for educational content
What happens when someone wants to use a piece of content in multiple different environments
Lots of people talking about this in the education space
Could we use the VRC DAMS as one of the DAMS connected to an LMS?
Seems like there aren’t discussions between librarians and groups working on standards for learning objects in education space
Chris: Have worked with VRC for a long time, but also represent users in Research Hub, which is being retired, and they need a digital asset management system
Been using Research Hub as a stop-gap
Libraries have coordinated, VRCs have coordinated, museums are struggling
David: Shared Shelf — teaching/research distinction is artificial, depending on you and your personal workflow
Good thing is that we’re at a moment of investing in multiple layers of infrastructure/tools that can be used for multiple purposes
When you try to connect LMS in a deep way with a tool that’s designed for rich management of collections, this is hard

Jenn: Used Shared Shelf at a previous institution; can put anything in there, suddenly why don’t we just use this instead of video channel
Where are the synergies for having these conversations?

Lynn: Shared Shelf could function like Research Hub
Campus license — college of environmental design, VRC — paying for campus wide license, though
Extra fee in terms of storage space
If someone else wants to use it, could have those discussions
Did talk to Lynne in early stage when investigating options; went with this because it could be shared with other campuses
Served them better to have an ARTstor portal, rather than being integrated into library

Mary Elings, Bancroft
Library is in development of Hydra/Fedora stack that will be DAMS for our resources
Bancroft is largest contributor, digitizing for a long time
Commercial solution — not sustainable due to size of collections
Budget is the issue
24 TB right now, 2M assets
Incoming, born-digital collections coming in, could be any manner of material
An “item” is loosely defined
Looking at system that can be used to just manage that quantity of material within institution, also feed in to UCLDC where appropriate, up to DPLA
Ultimately think that DPLA/Europeana will be hubs people go to

Sherri: Nuxeo DAMS meant as tool for libraries at this stage, handling function that UCB is using Fedora

Mary: Everyone’s contributing to UCLDC at some level, some larger institutions are using larger DAMS to feed in to UCLDC

Sherri: Issue of scale— Berkeley has so many existing collections
Riverside, Davis, Merced, digitized 15 collections — great stuff
Goal is to get those into a system where they can manage it, jumpstart more work
Different situation; could never ingest 2M things into Nuxeo tomorrow
How to make it optimal for system, and for campus

David: Is frontend for Nuxeo something that could be made available for scholars to do what they might do in Shared Shelf

Sherri: Maybe, eventually — currently an access platform
Could be a richer experience, thinking about faculty content intermingled

Mary: Smithsonian a lot like UC, similar situations
Diverse institutions doing lots of research, lots of content, no central way to manage it, different needs
Important that it’s built in Fedora (what he’s doing in Islandora)
Have to collect research data sets coming from researchers; faculty papers aren’t papers, but laptops and hard drives and spreadsheets
Bringing those in and managing them
Way to streamline donation to the library

Patrick: Lots of talk about feeding things up to DPLA/Europeana
But faculty need to get access to stuff that’s available only under fair use
For them, not reasonable to go to DPLA, need to go to the one that has all the stuff they can use
Dual track in thinking about where people will go to get access to this — that’s where the APIs or potential integration points have to happen for learning management systems

Sherri: Not that all content surfaced in these systems are in public domain
Libraries are claiming fair use in putting them on-line
Up to the end-user to decide
There will be restricted materials not available at DPLA level

Lynne: Content in DPLA is links back to original source

Mary: They show you a thumbnail, metadata record
Just harvesting, not posting

Patrick: Wrong that faculty will need a private portal?

Lynne: For faculty, ArtStor is their first stop — informed them of other options, but they go to ArtStor first
DPLA is overwhelming, too much content, not what they want
Lots of unique material that may not relate to what they want for lecture
Shared Shelf is like a customized database
Chris: College of Environmental Design — their scholarly workflows are similar to art history, but a little different

Mary: Services like DPLA/Europeana — breadcrumb that takes you back to that institution
See it as a pathway, not giving you enough to do research
Just thumbnail, quick view

David: As you see people growing up with tools like Evernote, nice tools for collecting stuff, do you expect to see more faculty/grad students who want various ways to access parts of materials, then track them as those tools evolve?

Sherri: Seeing this at CDL
Talked for years about book bag on public site— can grab various things, put it all together
But that represents selfish view of world; researcher is only going to want to collect stuff on CDL site
Thinking about features of public site; realization that researcher comes with tools
Building those tools isn’t where it’s at, build hooks for their own tools
Lots of confidence in infrastructure behind platform to move in that direction, but haven’t had those conversations yet to figure out how it’ll work

Mary: People want to download 100k images we digitized in one shot
Curators aren’t sure if it’s ok — well, they can do it one at a time, it’ll just take longer
Student wants to have Bancroft re-digitize 30-40 volumes because he can’t get access to content
People come into reading room, used to sit there and read, getting boxes of folders and snapping as many pictures as they can in the day
Digital version is their research, it’s all on their laptop
Giving them one-at-a-time access

Sherri: Need something with more nuance than “available” or “not”

Patrick: If someone did want to apply computational analysis to content in collections, are APIs there for them to get things out at an API level, and if they produce interesting computational derivatives — where do those belong?
Could someone build tools against this stuff?

Mary: APIs are answer, but if researcher creates data set, wants to access it— can put it in DASH at UCSF, build on it, stick it in as a new version
Products that we would like to be collecting
Finding a place to put them, getting the faculty to use a tool where they upload and describe it

David: This happens within trusted domains— contribution model in art history, people contribute derived data sets

Patrick: But these are silos, no connection between corpus and analytic resouts

Sherri: Thought about this a lot, have Calisphere (place where you get material) and eScholarship (where research goes)
No connection between the two of them
As systemwide service, hard to say when we provide something vs. campus or particular library providing it
Becomes more complicated landscape
More diverse institutions, more difficult it is to hook up on other side
Who provides support for faculty member

Jenn: Faculty-generated content: is this a collection, do they want to share with other faculty members?
Lessons, videos they’ve created, other content
Want to use same piece of media in campus course, EdX course, Extension course
DAMS for learning assets — e.g. faculty-generated learning assets
Faculty no longer so happy about uploading things 10 times

Lynn: Instructor can make a folder private, share with students, or public folder so anyone at institution can see it
Shared Shelf is very image-centric, but accepts all formats — video, documents, audio

Jenn: Indiana University, Michigan, Oregon State — having these conversations in ed tech space
Big pitch around learning assets
Involving Michigan libraries; has open-source repository
Having same conversations among ed tech folks; how to share learning assets
If we’re already doing this at library level, should talk about UC joint something, building on what CDL is doing instead of building a separate ed tech thing

Steve: Question about research derivatives in DAMS, and putting faculty products related to teaching into DAMS are functionally the same thing
Easier line to trace in ed space to the question about problem of giving faculty incentives to put appropriate metadata around artifacts so they can be meaningfully contributed
Maybe learning contexts are a way to grease the path towards future where it’s more apparent to research why they should contribute products of research to central DAMS

Lynn: IP issues— some profs don’t want to share
“are other people going to be able to use syllabus? don’t want people to copy it”

Sharon: Some really want to share, it depends
Patrick: Lots of related themes, motivating quality metadata, what are work products/derivatives
Doesn’t matter on functional level, more around getting workflows right
Do we have to build something to show the utility?