Projects to Collect Born-Digital Archival Material from Individuals and Community Groups

Howard Besser
Moving Image Archiving & Preservation
New York University
http://besser.tsoa.nyu.edu/howard/Talks/
http://www.nyu.edu/tisch/preservation/

Projects to Collect Born-Digital Archival Material from Individuals and Community Groups

- Background & The Problem of Personal Digital Archiving
- The PDA Conferences
- Interesting solutions and approaches to these problems; Lessons learned
  - InterPARES
  - Preserving Digital Public Television
  - Activist Archivists & the Occupy Movement

PDA Covers born-digital

- Correspondence/email
- Personal photos and group photo albums
- Manuscript drafts
- Personal documents
- Diaries
- Home movies

And has been extended to encompass:

- Family history
- Community/Ethnic history & Movements
- Genealogy
- Digital humanities.

In what environments do we find PDA material?

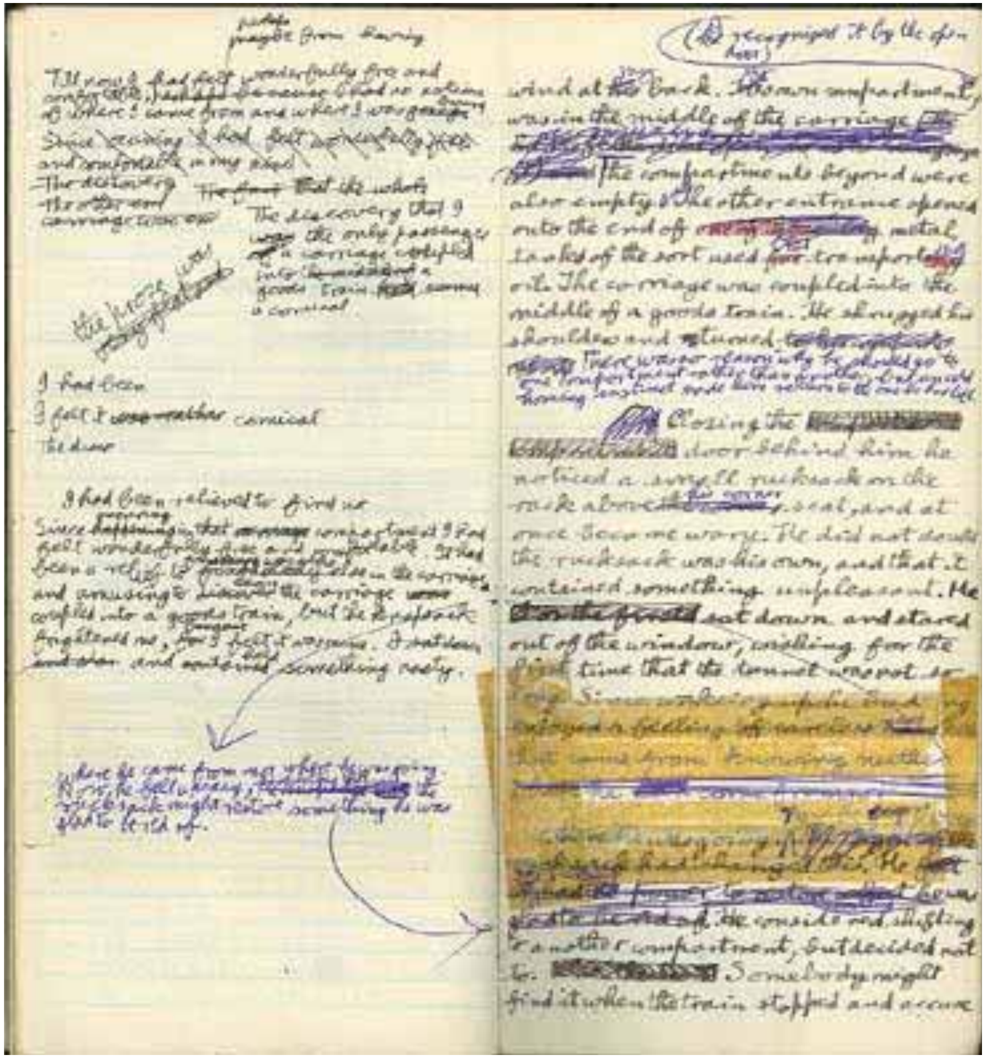
- Library Special Collections
- Collections documenting a community
- Collections documenting an ethnic group
- Collections documenting a social movement
- Collections documenting the work of any other type of group (a group of Architects, a set of law-makers, etc.)

GENERAL PROBLEMS OF BORN-DIGITAL PERSONAL CONTENT

In the analog world

- Traditionally, we have come to understand the work of writers and scientists by scholars studying their papers in Special Collections and Archives
- Their correspondence and progressively different drafts of papers reveal their changing thoughts and craft
- But how do we gather these in the Digital Age?

Alasdair Gray's Lanark (Glasgow U Library)



[illegible]
maybe from [illegible]

Till now I had felt wonderfully free and comfortable, [illegible] of where I came from and what I was going

~~Since waking I had felt wonderfully free and comfortable in my mind~~

~~The discovery~~
~~the other end~~
carriage [illegible]

~~The fact that the whole~~
The discovery that I was the only passenger of a carriage coupled into [illegible] a goods train [illegible] a comical

the [illegible]

I had been
I felt it ~~was rather~~ comical
the door

I had been relieved to find no
Since [illegible] in that compartment I had felt wonderfully free and comfortable. It had been a relief to [illegible] in the carriage, and amusing to discover the carriage was coupled into a goods train, but the rucksack frightened me, for I felt it was mine. I sat down and [illegible] and contained something nasty.

[illegible] he came from and what he was going [illegible] he felt a [illegible] the rucksack might contain something he was glad to be rid of.

(He recognized it by the [illegible])

wind at the back. His own compartment was in the middle of the carriage [illegible] The compartments beyond were also empty & the other entrance opened onto the end of [illegible] metal tanks of the sort used for transporting oil. The carriage was coupled into the middle of a goods train. He shrugged his shoulders and returned [illegible] There was no reason why he should go to one compartment rather than another, but a [illegible] homing instinct must have returned [illegible]

Closing the [illegible] door behind him he noticed a small rucksack on the rack above the [illegible] seat, and at once became wary. He did not doubt the rucksack was his own, and that it contained something unpleasant. He [illegible] sat down and stared out of the window, wishing for the first time that the tunnel was not so long. Since waking [illegible] he had enjoyed a feeling of [illegible] that came from knowing nothing [illegible]

[illegible] had changed this. He [illegible] the rucksack off [illegible] considered shifting to another compartment, but decided not to. [illegible] Somebody might find it when the train stopped and accuse

Correspondence

Correspondence

FILE COPY

STATE OF UTAH

Dear Governor Walker

CONGRATULATI

Where can we find these today?

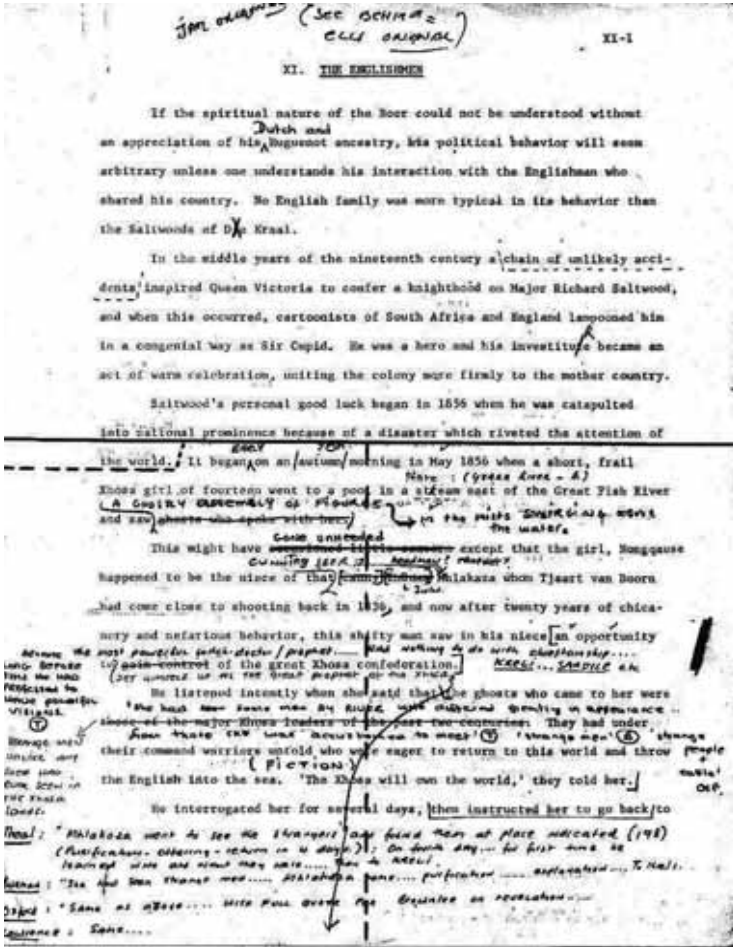
- Do people write letters on paper? Can we see the iterations of changes on manuscripts?
- Where can we find today's equivalent of these?

- This will require
  - new interventions (like changing authors' workflow, or intervening in email handling software)
  - New tools (like for analyzing email)
  - new approaches like digital archeology, forensics

Stages of the problem

- Stage #1: People write on computers instead of paper
- Stage #2: People no longer store their digital works in places over which they have absolute control
  - Email services (gmail, yahoo)
  - Cloud storage for documents (google docs)
  - Social network services (FLICKR, YouTube)

Our Changing Environment



Jan changes (See Schima's ell original)

XI-1

XI. THE ENGLISHMEN

If the spiritual nature of the Boer could not be understood without an appreciation of his Dutch and Huguenot ancestry, his political behavior will seem arbitrary unless one understands his interaction with the Englishman who shared his country. No English family was more typical in its behavior than the Saltwoods of De Kraal.

In the middle years of the nineteenth century a chain of unlikely accidents inspired Queen Victoria to confer a knighthood on Major Richard Saltwood, and when this occurred, cartoonists of South Africa and England lampooned him in a congenial way as Sir Cupid. He was a hero and his investiture became an act of warm celebration, uniting the colony more firmly to the mother country.

Saltwood's personal good luck began in 1856 when he was catapulted into national prominence because of a disaster which riveted the attention of the world. It began early on an ~~autumn~~ Sept morning in May 1856 when a short, frail Xhosa girl of fourteen went to a pool in a stream east of the Great Fish River and saw ~~ghosts who spoke with her~~ A ghostly assembly of figures in the mists shimmering above the water.

Name: (Great Kwee - A)

This might have ~~occasioned little concern~~ gone unheeded except that the girl, Nongqause happened to be the niece of that ~~cunning~~ Mhlakaza whom Tjaart van Doorn had come close to shooting back in 1836, and now after twenty years of chicanery and nefarious behavior, this shifty man saw in his niece an opportunity to ~~gain control~~ of the great Xhosa confederation.

cunning seer ... madman! prophet

among the most powerful witch-doctor / prophet ... had nothing to do with chieftainship ... Kreli ... Sarhili etc.

(set himself up as the great prophet of the Xhosa)

He listened intently when she said that the ghosts who came to her were ~~those of the major Xhosa leaders of the past two centuries~~. They had under their command warriors untold who were eager to return to this world and throw the English into the sea. 'The Xhosa will own the world,' they told her.

'she had seen some men by river, differing greatly in appearance from those she was accustomed to meet' (T) 'strange men' (B) 'strange people' (S)

(FICTION)

Among long before time she had [illegible] to whose powerful villages.

(T)

strange men unlike any [illegible] ever seen in the Xhosa lands.

He interrogated her for several days, then instructed her to go back to

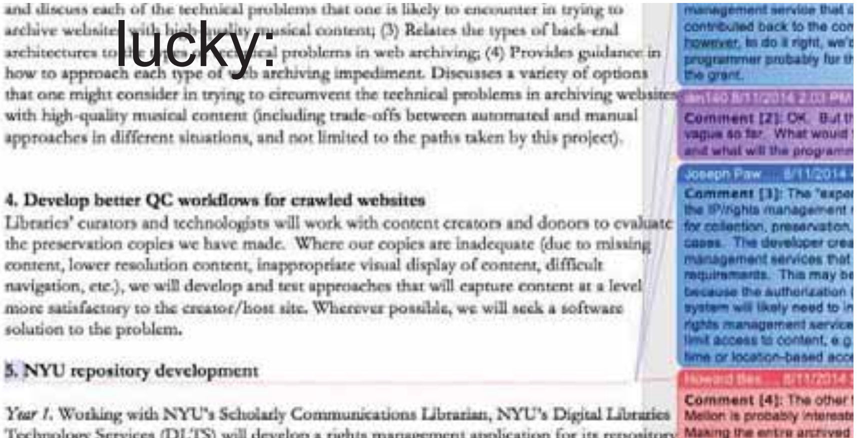
Real: "Mhlakaza went to see the strangers and found them at place indicated (198) (Purification - offering - return in 4 days); On fourth day ... for first time he learned who and what they were ..... then to Kreli ... T. Hall.

[illegible]: "Saw the two strange men ..... Mhlakaza [illegible] ..... purification ..... [illegible] ...

[illegible]: "Same as above .... with full [illegible] [illegible] Ergeulee on revelation ...

Lawrence: Same ....

If we're lucky:



and discuss each of the technical problems that one is likely to encounter in trying to archive websites with high-quality musical content; (3) Relates the types of back-end architectures to the types of technical problems in web archiving; (4) Provides guidance in how to approach each type of web archiving impediment. Discusses a variety of options that one might consider in trying to circumvent the technical problems in archiving websites with high-quality musical content (including trade-offs between automated and manual approaches in different situations, and not limited to the paths taken by this project).

4. Develop better QC workflows for crawled websites

Libraries' curators and technologists will work with content creators and donors to evaluate the preservation copies we have made. Where our copies are inadequate (due to missing content, lower resolution content, inappropriate visual display of content, difficult navigation, etc.), we will develop and test approaches that will capture content at a level more satisfactory to the creator/host site. Wherever possible, we will seek a software solution to the problem.

5. NYU repository development

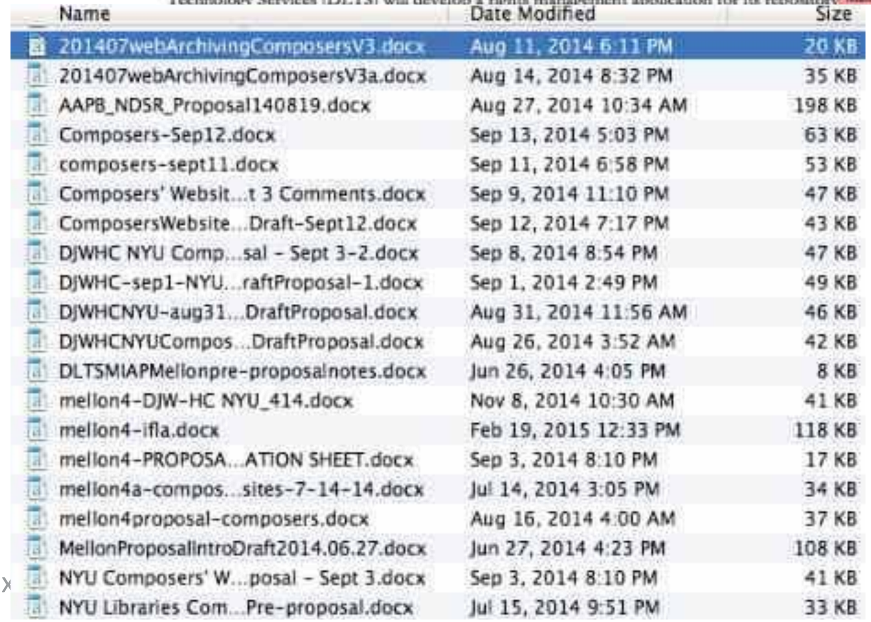
Year 1. Working with NYU's Scholarly Communications Librarian, NYU's Digital Libraries Technology Services (DLTS) will develop a rights management application for its repository

management service that
contributed back to the com
however, to do it right, we'd
programmer probably for th
the grant.

an140 8/11/2014 2:03 PM
Comment [2]: OK. But th
vague so far. What would
and what will the programm

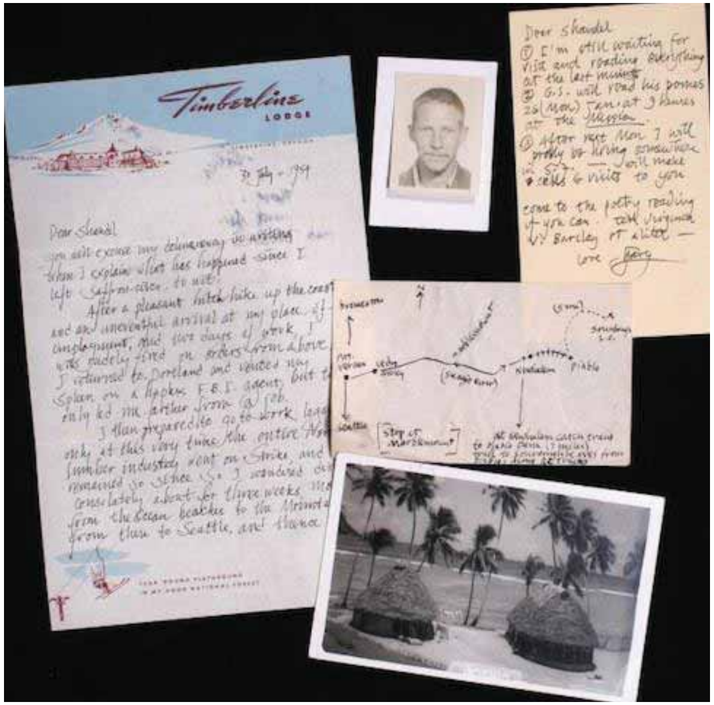
Joseph Paw 8/11/2014
Comment [3]: The "expe
the IP/rights management
for collection, preservation,
cases. The developer crea
management services that
requirements. This may be
because the authorization
system will likely need to in
rights management service
limit access to content, e.g.
time or location-based acc

Howard Bes 8/11/2014
Comment [4]: The other
Mellon is probably interest
Making the entire archived



Name	Date Modified	Size
201407webArchivingComposersV3.docx	Aug 11, 2014 6:11 PM	20 KB
201407webArchivingComposersV3a.docx	Aug 14, 2014 8:32 PM	35 KB
AAPB_NDSR_Proposal140819.docx	Aug 27, 2014 10:34 AM	198 KB
Composers-Sep12.docx	Sep 13, 2014 5:03 PM	63 KB
composers-sept11.docx	Sep 11, 2014 6:58 PM	53 KB
Composers' Websit...t 3 Comments.docx	Sep 9, 2014 11:10 PM	47 KB
ComposersWebsite...Draft-Sept12.docx	Sep 12, 2014 7:17 PM	43 KB
DJWHC NYU Comp...sal - Sept 3-2.docx	Sep 8, 2014 8:54 PM	47 KB
DJWHC-sep1-NYU...raftProposal-1.docx	Sep 1, 2014 2:49 PM	49 KB
DJWHCNYU-aug31...DraftProposal.docx	Aug 31, 2014 11:56 AM	46 KB
DJWHCNYUCompos...DraftProposal.docx	Aug 26, 2014 3:52 AM	42 KB
DLTSMIAPMellonpre-proposalnotes.docx	Jun 26, 2014 4:05 PM	8 KB
mellon4-DJW-HC NYU_414.docx	Nov 8, 2014 10:30 AM	41 KB
mellon4-ifla.docx	Feb 19, 2015 12:33 PM	118 KB
mellon4-PROPOSA...ATION SHEET.docx	Sep 3, 2014 8:10 PM	17 KB
mellon4a-compos...sites-7-14-14.docx	Jul 14, 2014 3:05 PM	34 KB
mellon4proposal-composers.docx	Aug 16, 2014 4:00 AM	37 KB
MellonProposalIntroDraft2014.06.27.docx	Jun 27, 2014 4:23 PM	108 KB
NYU Composers' W...posal - Sept 3.docx	Sep 3, 2014 8:10 PM	41 KB
NYU Libraries Com...Pre-proposal.docx	Jul 15, 2014 9:51 PM	33 KB

Our Changing Environment



Timberline Lodge

31 July - '59

Dear Shandel

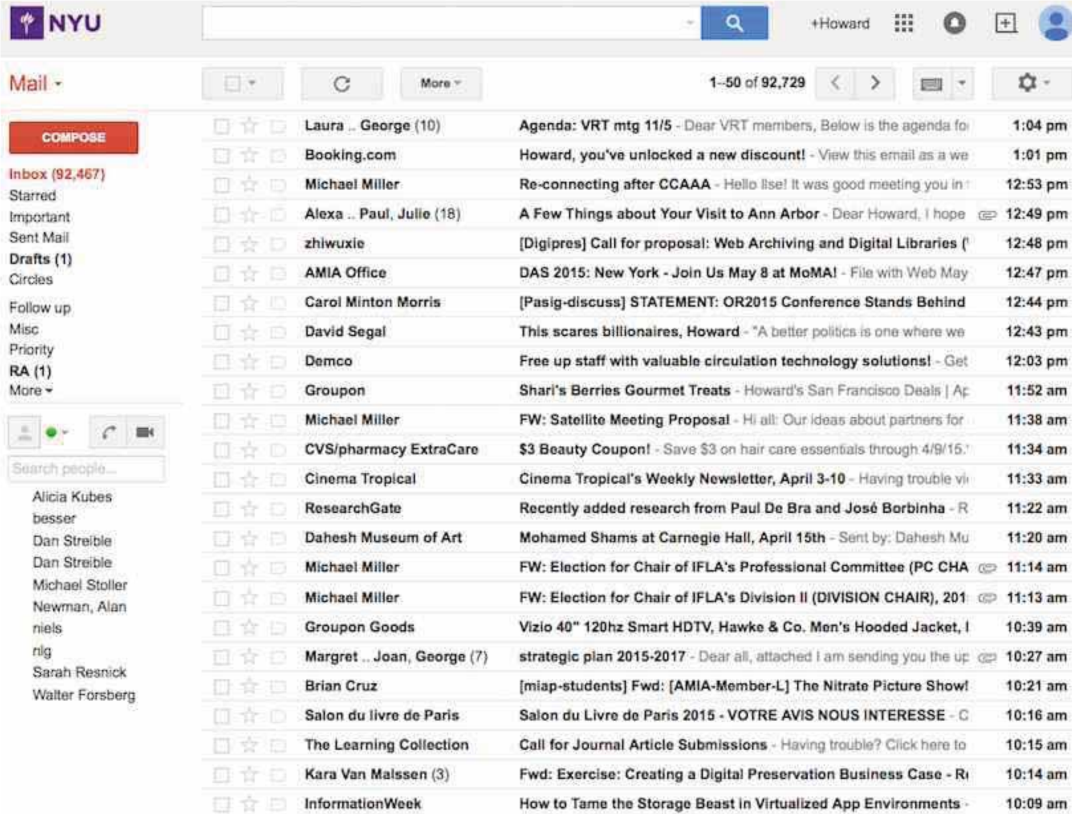
You will excuse my delinquency in writing when I explain what has happened since I left San Francisco, to wit:

After a pleasant hitch hike up the coast and an uneventful arrival at my place of employment, and five days of work, I was rudely fired on orders from above. I returned to Portland and visited my [illegible] on a hopeless F.B.I. agent, but this only led me farther from a job.

I then prepared to go to work logging only at this very time the entire Nor[illegible] lumber industry went on Strike, and remained so since. So I wandered disconsolately about for three weeks, [illegible] from the Ocean beaches to the Mount[illegible] from there to Seattle, and thence

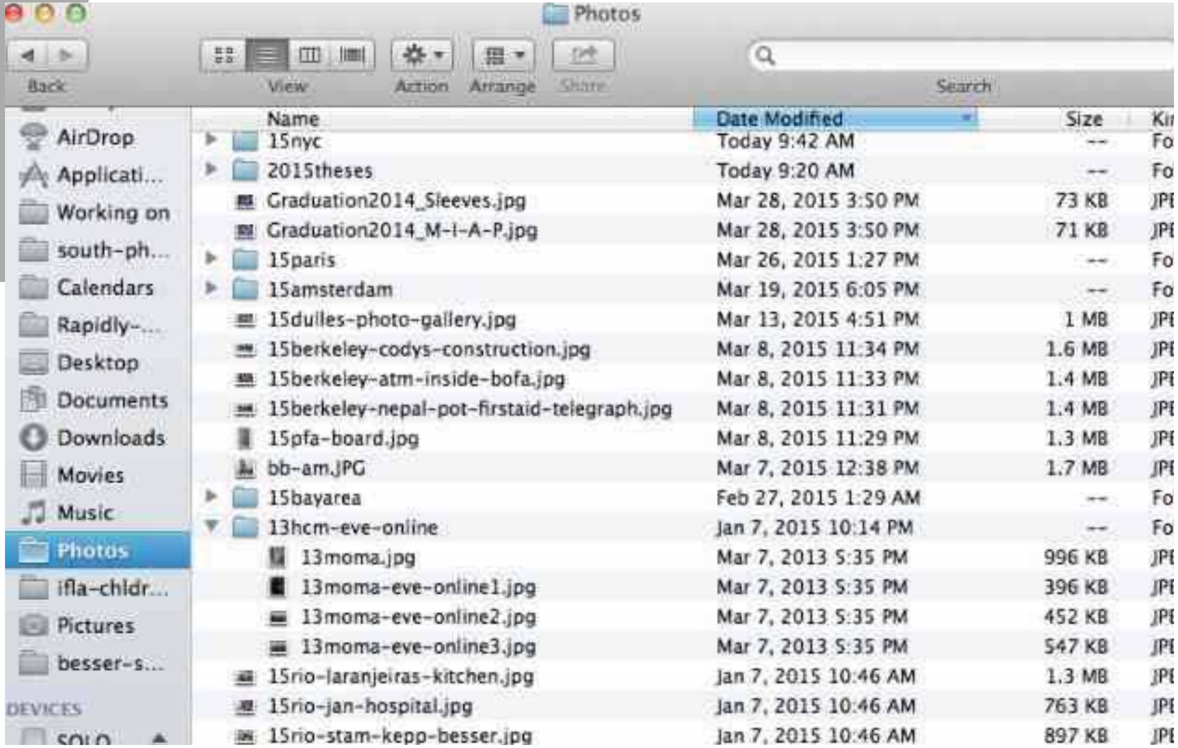
Dear Shandel
① I'm still waiting for visit and reading everything at the last minute
② G.S. will read his poems 26 (Mon) 7 pm. at [illegible] at the [illegible]
③ After next Mon I will probably be living somewhere — will make [illegible] calls & visits to you

Come to the poetry reading if you can. Tell Virginia W. Barclay et al. —
Love Gary



Laura .. George (10)	**Agenda: VRT mtg 11/5** - Dear VRT members, Below is the agenda fo	1:04 pm
Booking.com	**Howard, you've unlocked a new discount!** - View this email as a we	1:01 pm
Michael Miller	**Re-connecting after CCAAA** - Hello Ilse! It was good meeting you in	12:53 pm
Alexa .. Paul, Julie (18)	**A Few Things about Your Visit to Ann Arbor** - Dear Howard, I hope	12:49 pm
zhiwuxie	**[Digipres] Call for proposal: Web Archiving and Digital Libraries (**	12:48 pm
AMIA Office	**DAS 2015: New York - Join Us May 8 at MoMA!** - File with Web May	12:47 pm
Carol Minton Morris	**[Pasig-discuss] STATEMENT: OR2015 Conference Stands Behind**	12:44 pm
David Segal	**This scares billionaires, Howard** - "A better politics is one where we	12:43 pm
Demco	**Free up staff with valuable circulation technology solutions!** - Get	12:03 pm
Groupon	**Shari's Berries Gourmet Treats** - Howard's San Francisco Deals | Ap	11:52 am
Michael Miller	**FW: Satellite Meeting Proposal** - Hi all: Our ideas about partners for	11:38 am
CVS/pharmacy ExtraCare	**$3 Beauty Coupon!** - Save $3 on hair care essentials through 4/9/15.	11:34 am
Cinema Tropical	**Cinema Tropical's Weekly Newsletter, April 3-10** - Having trouble vi	11:33 am
ResearchGate	**Recently added research from Paul De Bra and José Borbinha** - R	11:22 am
Dahesh Museum of Art	**Mohamed Shams at Carnegie Hall, April 15th** - Sent by: Dahesh Mu	11:20 am
Michael Miller	**FW: Election for Chair of IFLA's Professional Committee (PC CHA**	11:14 am
Michael Miller	**FW: Election for Chair of IFLA's Division II (DIVISION CHAIR), 201**	11:13 am
Groupon Goods	**Vizio 40" 120hz Smart HDTV, Hawke & Co. Men's Hooded Jacket, I**	10:39 am
Margret .. Joan, George (7)	**strategic plan 2015-2017** - Dear all, attached I am sending you the up	10:27 am
Brian Cruz	**[miap-students] Fwd: [AMIA-Member-L] The Nitrate Picture Show!**	10:21 am
Salon du livre de Paris	**Salon du Livre de Paris 2015 - VOTRE AVIS NOUS INTERESSE** - C	10:16 am
The Learning Collection	**Call for Journal Article Submissions** - Having trouble? Click here to	10:15 am
Kara Van Malssen (3)	**Fwd: Exercise: Creating a Digital Preservation Business Case** - R	10:14 am
InformationWeek	**How to Tame the Storage Beast in Virtualized App Environments** -	10:09 am

Our Changing Environment



Name	Date Modified	Size	Ki
▶ 15nyc	Today 9:42 AM	--	Fo
▶ 2015theses	Today 9:20 AM	--	Fo
Graduation2014_Sleeves.jpg	Mar 28, 2015 3:50 PM	73 KB	JPE
Graduation2014_M-I-A-P.jpg	Mar 28, 2015 3:50 PM	71 KB	JPE
▶ 15paris	Mar 26, 2015 1:27 PM	--	Fo
▶ 15amsterdam	Mar 19, 2015 6:05 PM	--	Fo
15dulles-photo-gallery.jpg	Mar 13, 2015 4:51 PM	1 MB	JPE
15berkeley-codys-construction.jpg	Mar 8, 2015 11:34 PM	1.6 MB	JPE
15berkeley-atm-inside-bofa.jpg	Mar 8, 2015 11:33 PM	1.4 MB	JPE
15berkeley-nepal-pot-firstaid-telegraph.jpg	Mar 8, 2015 11:31 PM	1.4 MB	JPE
15pfa-board.jpg	Mar 8, 2015 11:29 PM	1.3 MB	JPE
bb-am.JPG	Mar 7, 2015 12:38 PM	1.7 MB	JPE
▶ 15bayarea	Feb 27, 2015 1:29 AM	--	Fo
▼ 13hcm-eve-online	Jan 7, 2015 10:14 PM	--	Fo
13moma.jpg	Mar 7, 2013 5:35 PM	996 KB	JPE
13moma-eve-online1.jpg	Mar 7, 2013 5:35 PM	396 KB	JPE
13moma-eve-online2.jpg	Mar 7, 2013 5:35 PM	452 KB	JPE
13moma-eve-online3.jpg	Mar 7, 2013 5:35 PM	547 KB	JPE
15rio-laranjeiras-kitchen.jpg	Jan 7, 2015 10:46 AM	1.3 MB	JPE
15rio-jan-hospital.jpg	Jan 7, 2015 10:46 AM	763 KB	JPE
15rio-stam-kepp-besser.jpg	Jan 7, 2015 10:46 AM	897 KB	JPE

Our Changing Environment

- Rise of Online Services and Social Media is changing where this content resides (and is imposing restrictions that go beyond the rightsholder)

Core Multi-location Problems

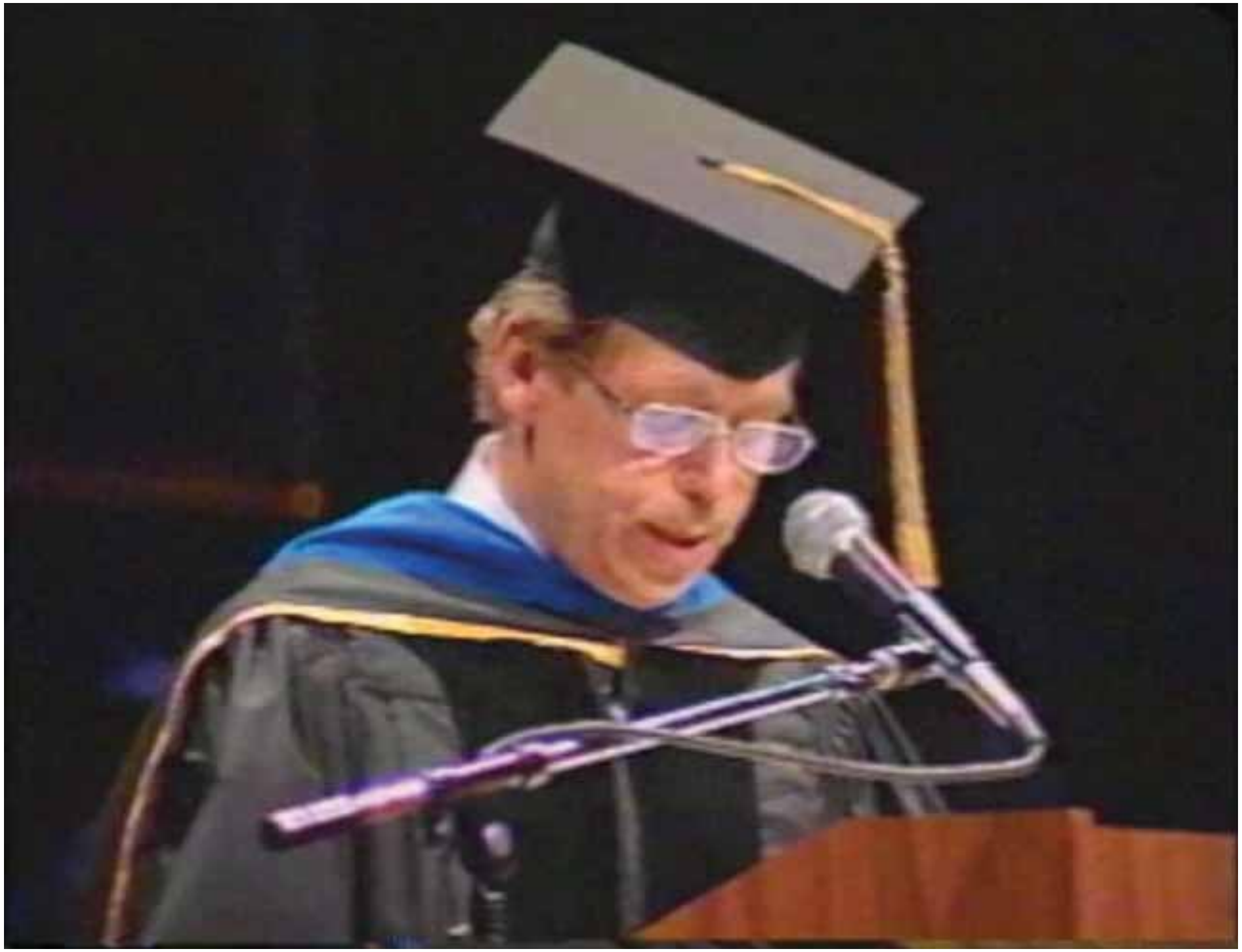
- It’s difficult enough when someone’s photos are spread throughout their hard disk. But today some photos there, but others on their phone(s), Instagram, Flickr, Facebook, in Tweets, etc.
- Similar problems plague email
- Most Social Network TOS policies prohibit the owner from giving their password to anyone else (even Library)

And how do we handle donations after an important person dies?

And these issues are also true for Community Grps & Assns

- w/Social Media, group activity is more important than ever
- But each person in the group is an individual collector. And frequently a set of individual collections forms the group collection.

Famous visitor to UCLA



Documenting Protests

-photo from Activists Guide to Archiving Video

When aggregated, many different personal collections form an important picture of:

- An ethnic group
- A community
- A social movement
- A set of architects
- A set of law-makers

- What is important to them, how they go about their business, …

And we know from past works that aggregations create new meanings

- Aggregating all the photos of the Digital Diaspora is hugely more meaningful than a single photo
- One tweet says very little, but thousands of tweets can show trends or depict a particular event or day

But in the PDA world, aggregating items causes significant problems

- Vast quantity of user-contributed material
- No easy way to control for quality, file format, metadata (not even any consistency for any of these)-

Every Image Collector has a Different Approach

- Different file-naming conventions
- Different file formats
- Different compression schemes
- Different metadata
- Stored in different arrangements/hierarchies
- Stored in different places (cellphone, personal hard disk, Instragram/Facebook, …)

THE PDA CONFERENCES

PDA Covers born-digital

- Correspondence/email
- Personal photos and group photo albums
- Manuscript drafts
- Personal documents
- Home movies

And has been extended to encompass:

- Family history
- Community/Ethnic history & Movements
- Genealogy
- Digital humanities.

PDA: Who Attends & Presents

- Citizen Archivists
  - People who want to step in and rescue content in peril
  - People who like to create software/Apps/Guidelines to help others facing similar problems
- Community or Ethnic groups and activists wanting to save portions of their heritage
- Professional librarians & archivists (and their programming support staff)
- Regular software developers
- Researchers (both academic and computer industry)

PDA Goals—Sharing knowledge

- What worked and what didn’t; what parts turned out to be more difficult than anticipated
- New and different types of content to collect
- Guidelines, procedures, workflows, methodologies
- Software

PDA History

Initially started by Internet Archive with co-sponsorship from Netherlands Sound & Vision, LC/NDIIPP and CNI

- 2010 Internet Archive
- 2011 Internet Archive
- 2012 Internet Archive
- 2013 Univ of Maryland
- 2014 Indiana State Library
- 2015 New York University

SOME KEY IDEAS FROM INTERPARES, PDPTV, ACTIVIST ARCHIVISTS (HTTP://ACTIVIST-ARCHIVISTS.ORG/)

What I know from my prior work with other types of Digital Content

- InterPARES—If we hope to preserve electronic records, archivists need to be involved early in the life-cycle of that record, long before the record enters the archive
- Preserving Digital Public Television—Pushing metadata gathering upstream into the production cycle

Preserving Digital Public Television Workflow in Production Process-

- Site Visits to productions
- Interview Production staff
- Diagrams of Workflow-

Pushing Metadata Gathering Upstream: The Problem

TRADITIONALLY…

- Very little metadata required for preservation accompanies an object to a repository.
- Archives, libraries and other repositories must create (or re-create) most of the necessary metadata.
- This requires many manual hours, and significant resources - both time and money.

IN THE DIGITAL WORLD…

- This doesn't scale up. Repositories will be unable to continue in this manner, as more metadata than ever is required.

But much of the necessary metadata has already been gathered during production

- For each element/clip, production team usually notes source, date, place, people, and other descriptive info
- But this is treated as internal information, and often various parts of the info are distributed among the personal notebooks of different production assistants
- There is seldom a central location for this info, and the info is seldom turned over to the archive (which later tries to recreate much of it)

When the Archive tries to re-create this info, it is seldom successful

Producers know much more about the content of their productions than the archivists do. Archivists wanting accurate info must go back to the production staff (often years later) to start brainstoriming over the info

"Once the (television) program is finished, it is passed on to the archive or library for safe keeping. Librarians will catalog and classify the content, possibly using a proxy copy, and enter the resulting informative metadata in their database so they can retrieve it in the future. However, rarely if ever is the metadata from the rest of the process passed onto them, except, perhaps, for the title, tape number, and basic technical information about recording formats. **It has to be re-created, with all the associated risk of errors and lack of accuracy--not to mention the work and time involved.**"

- Cox, Tadic, and Mulder, *Descriptive Metadata for Television* (2006)

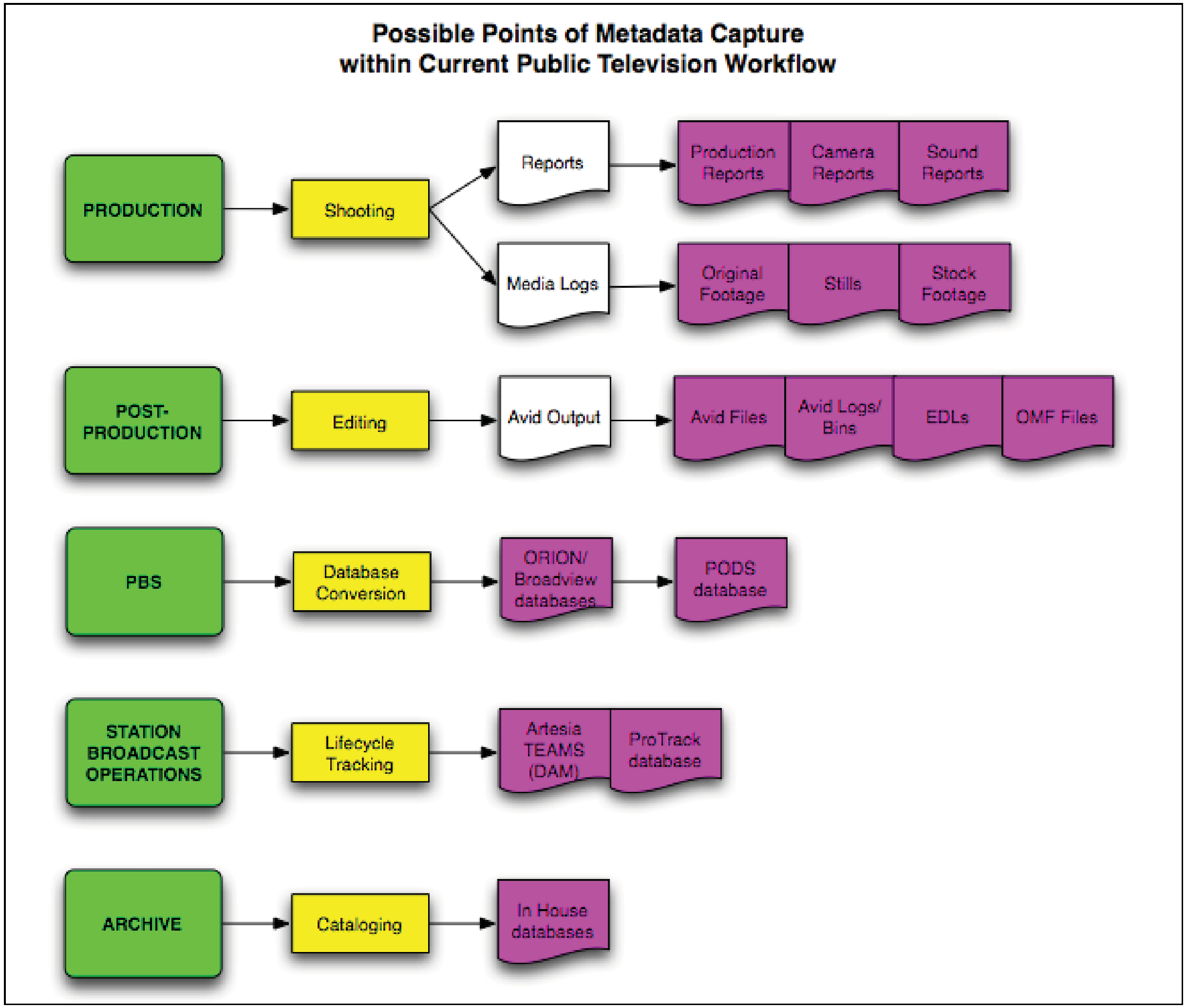
We need to find ways to push metadata access upstream

- Digital requires even more metadata than Analog
  - As the workflow becomes file-based, the need for robust and accurate metadata will become critical. File relationships, video codecs, bit rates, and rights information must be explicit, accurate, and immediately accessible. This will require a much deeper level of metadata than is currently captured in tape-based archives.

– We can't continue to supply this metadata at ingest; that won't scale

  - Obtaining the necessary metadata at the end of production and broadcast life cycle is not feasible. Metadata will need to be systematically gathered during the production lifecycle and submitted with the programs to the preservation repository.

Examined Potential Points of Metadata Capture



Examined Potential Points for Metadata Capture

- Much of the necessary metadata for preservation is already generated by the production unit, but discarded after their internal use. This needs to be captured throughout the workflow.

- "Those in the production unit are the creators and have first hand knowledge of who, what, where, when, and why the content was created." -- Mary Ide and Leah Weisse, WGBH Archivists.

Proposed Solutions…?

- Preservation becoming a shared responsibility between content creators, distributors, curators, and preservationists.
- Partnerships are needed to come to unified solutions.
- Preservationists seek reliable metadata back upstream in the production workflow...

WorldFocus

- Nightly news program begun Oct 2008
- We began working with Workflows six months before program began
- Had ability to engineer metadata gathering into the creation/production process

Activist Archivists

http://activist-archivists.org/

- MIAP students and grads originally working on archiving media from the Occupy movement
- Guidelines for recorders to make their works more easily preservable: make notes, turn on GPS, upload to service that doesn't strip out metadata, keep raw footage, don't compress
- For meeting recordings, have them read a script at start of the recording

How Occupy material resembles what we’ll be facing in the future

- Vast quantity of user-contributed material
- No easy way to control for quality, file format, metadata
  - no enforcing guidelines as with organizational records
  - no semi-consistency as in a single individual’s personal records
- Much of the material can most easily be found on Social Networks
- …we need to find smart ways to harvest metadata and analyze files, as well as to influence behavior of potential contributors

Make it easy for future contributors to create “archival-friendly” works-

- Low-hanging fruit
- Easy instructional material that appeals to what they think is important

Low-Hanging fruit

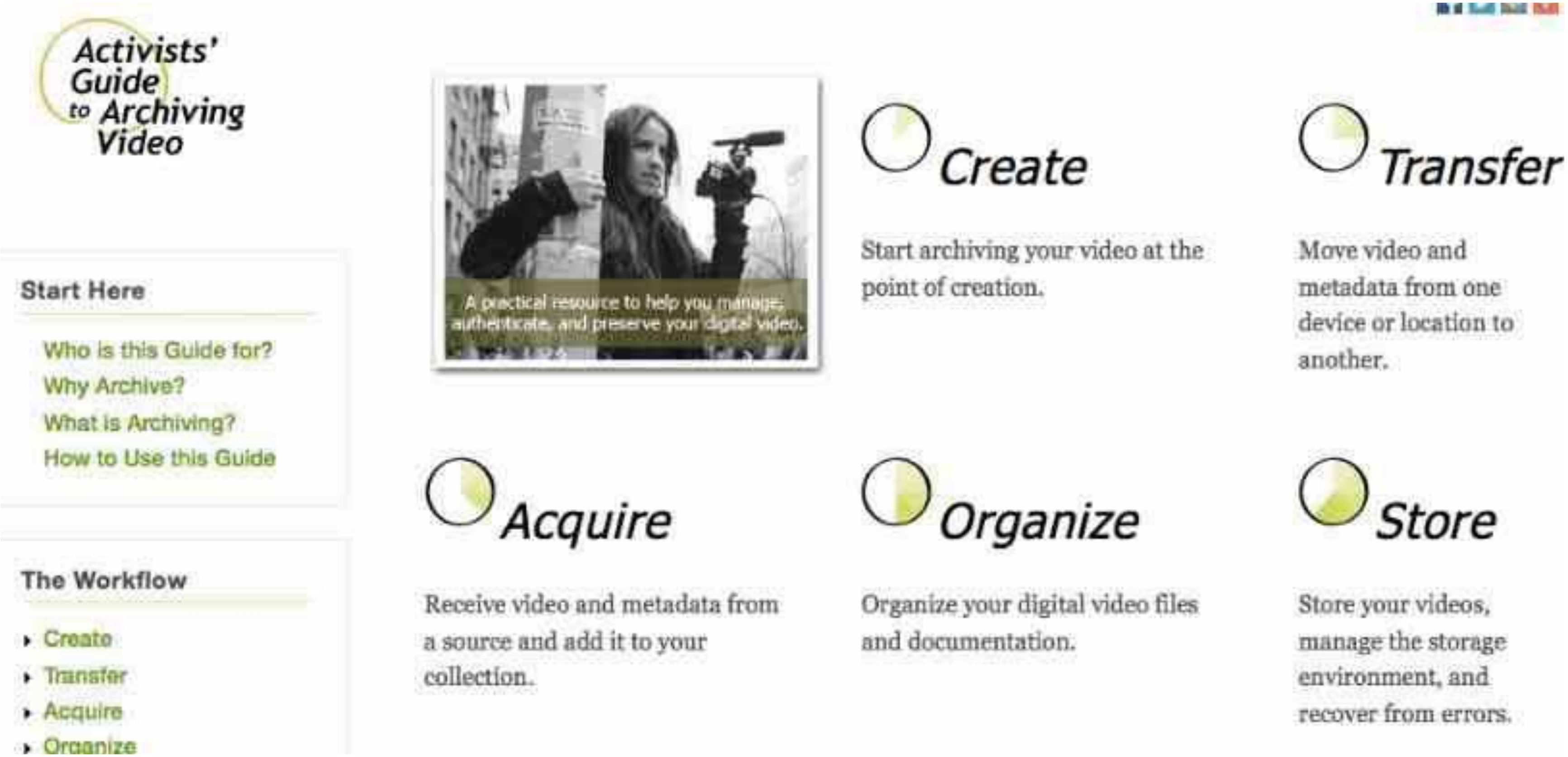
- Turn GPS on
- Develop strategies for automating a profile and uploads (our ideal App)

7 Tips to Ensure Your Video Is Usable in the Long Term

- Collect details while filming
- Keep your original raw footage, unaltered
- Make your video discoverable
- Contextualize it
- Make it verifiable
- Allow others to collect and archive
- Or archive it yourself

WITNESS: Activists' Guide to Archiving Video, Yvonne Ng

http://archiveguide.witness.org/



Activists' Guide to Archiving Video

Start Here

- Who is this Guide for?
- Why Archive?
- What is Archiving?
- How to Use this Guide

The Workflow

- Create
- Transfer
- Acquire
- Organize

A practical resource to help you manage, authenticate, and preserve your digital video.

Create

Start archiving your video at the point of creation.

Transfer

Move video and metadata from one device or location to another.

Acquire

Receive video and metadata from a source and add it to your collection.

Organize

Organize your digital video files and documentation.

Store

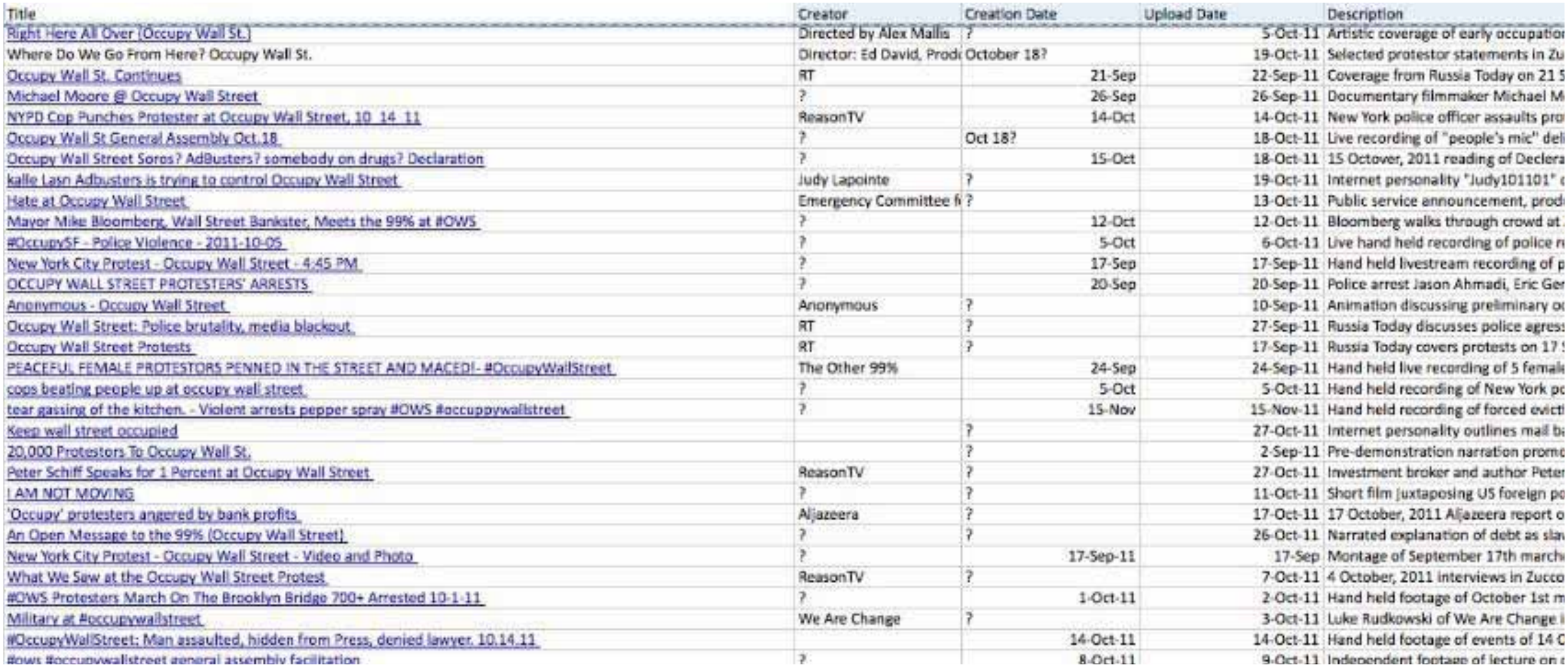
Store your videos, manage the storage environment, and recover from errors.

Find smart ways to deal with Scale-

Tamiment YouTube collecting

- plug-in for FireFox (downloadhelper.net)
- As of April, ca. 250 items, policy: large events
- Fair Use: 2012 ARL Code of Best Practices
  - "transformative" collecting with context
- Tamiment has been selectively browsing through YouTube Occupy videos, trying to choose which ones to keep, then cataloging them with
  - Title, Creator, Creation Date, Upload Date, Description, URL, Youtube Username, License, Format, Codec, Source Media, On Internet Archive, CC License type

Tamiment YouTube Cataloging



Title	Creator	Creation Date	Upload Date	Description
Right Here All Over (Occupy Wall St.)	Directed by Alex Mallis	?	5-Oct-11	Artistic coverage of early occupatio
Where Do We Go From Here? Occupy Wall St.	Director: Ed David, Prod	October 18?	19-Oct-11	Selected protestor statements in Zu
Occupy Wall St. Continues	RT	21-Sep	22-Sep-11	Coverage from Russia Today on 21 S
Michael Moore @ Occupy Wall Street	?	26-Sep	26-Sep-11	Documentary filmmaker Michael M
NYPD Cop Punches Protester at Occupy Wall Street, 10_14_11	ReasonTV	14-Oct	14-Oct-11	New York police officer assaults pro
Occupy Wall St General Assembly Oct.18	?	Oct 18?	18-Oct-11	Live recording of "people's mic" deli
Occupy Wall Street Soros? AdBusters? somebody on drugs? Declaration	?	15-Oct	18-Oct-11	15 Octover, 2011 reading of Declera
kalle Lasn Adbusters is trying to control Occupy Wall Street	Judy Lapointe	?	19-Oct-11	Internet personality "Judy101101" c
Hate at Occupy Wall Street	Emergency Committee fc	?	13-Oct-11	Public service announcement, prod
Mayor Mike Bloomberg, Wall Street Bankster, Meets the 99% at #OWS	?	12-Oct	12-Oct-11	Bloomberg walks through crowd at
#OccupySF - Police Violence - 2011-10-05	?	5-Oct	6-Oct-11	Live hand held recording of police r
New York City Protest - Occupy Wall Street - 4:45 PM	?	17-Sep	17-Sep-11	Hand held livestream recording of p
OCCUPY WALL STREET PROTESTERS' ARRESTS	?	20-Sep	20-Sep-11	Police arrest Jason Ahmadi, Eric Ger
Anonymous - Occupy Wall Street	Anonymous	?	10-Sep-11	Animation discussing preliminary oc
Occupy Wall Street: Police brutality, media blackout	RT	?	27-Sep-11	Russia Today discusses police agress
Occupy Wall Street Protests	RT	?	17-Sep-11	Russia Today covers protests on 17 S
PEACEFUL FEMALE PROTESTORS PENNED IN THE STREET AND MACED!- #OccupyWallStreet	The Other 99%	24-Sep	24-Sep-11	Hand held live recording of 5 female
cops beating people up at occupy wall street	?	5-Oct	5-Oct-11	Hand held recording of New York po
tear gassing of the kitchen. - Violent arrests pepper spray #OWS #occuppywallstreet	?	15-Nov	15-Nov-11	Hand held recording of forced evict
Keep wall street occupied		?	27-Oct-11	Internet personality outlines mail ba
20,000 Protestors To Occupy Wall St.		?	2-Sep-11	Pre-demonstration narration promo
Peter Schiff Speaks for 1 Percent at Occupy Wall Street	ReasonTV	?	27-Oct-11	Investment broker and author Peter
I AM NOT MOVING	?	?	11-Oct-11	Short film juxtaposing US foreign po
'Occupy' protesters angered by bank profits	Aljazeera	?	17-Oct-11	17 October, 2011 Aljazeera report o
An Open Message to the 99% (Occupy Wall Street)	?	?	26-Oct-11	Narrated explanation of debt as slav
New York City Protest - Occupy Wall Street - Video and Photo	?	17-Sep-11	17-Sep	Montage of September 17th march
What We Saw at the Occupy Wall Street Protest	ReasonTV	?	7-Oct-11	4 October, 2011 interviews in Zucco
#OWS Protesters March On The Brooklyn Bridge 700+ Arrested 10-1-11	?	1-Oct-11	2-Oct-11	Hand held footage of October 1st m
Military at #occupywallstreet	We Are Change	?	3-Oct-11	Luke Rudkowski of We Are Change i
#OccupyWallStreet: Man assaulted, hidden from Press, denied lawyer. 10.14.11		14-Oct-11	14-Oct-11	Hand held footage of events of 14 O
#ows #occupywallstreet general assembly facilitation	?	8-Oct-11	9-Oct-11	Independent footage of lecture on

But this won’ t scale!

March 24, 2012 YouTube stats

(just 6 months after start of movement)

- “#Occupy” 169,000
- “Occupy Wall Street” 98,400
- “Occupy Protest” 70,500
- “Occupy Movement” 54,800
- “#OWS” 50,300
- “Occupy Oakland” 13,400
- “Zucotti Park” 6,690

Alternative approach to YouTube Selection process

- Develop categories of important YouTube videos
  - Celebrity visits, Internal workings (library, kitchen, media), Confrontations with police, Labor, Housing, etc.
- Have Occupiers fill in an online form listing the 5 most important videos in each category

Advantages of YouTube Collaborative Filtering Selection Process

- Scalable and manageable
- Consistent with Occupy ideas of inclusiveness and of managing own story
- Tamiment can still choose to be selective in collecting only a portion of what is voted in, but the total set for review is a manageable scale

Tips on Outreach to Communities

- Build trust
- Speak in their language (not MARC-speak)
- Identify ways you can meet needs they already perceive
- Approach projects as collaboration whenever possible
- Don’t only focus on content and metadata, but also rights that can be an impediment to preservation

Projects to Collect Born-Digital Archival Material from Individuals and Community Groups

- Special Thanks to rest of the PDA 2015 Organizing Committee (especially Mike Ashenfelder for coherent summarization)
- http://blogs.loc.gov/digitalpreservation/2015/08/report-on-the-personal-digital-archiving-2015-conference/
- http://personaldigitalarchiving.com/
- http://besser.tsoa.nyu.edu/howard/Talks/