Overview & Updates

Andrea Goethals, Harvard Library
NDSR mission

... to develop the next generation of stewards to collect, manage, preserve, and make accessible our nation’s digital assets

... to provide residents with a combination of hands-on learning and expert guidance

... help define clear educational standards for digital stewardship education and training
Replicating & testing the model

IMLS-funded grants to test model in New York (METRO) and Boston areas (Harvard / MIT)
Replicating what exactly?

- Residency structure
- Resident & host cohort model
- Resident eligibility & application requirements
- Single institution administering program
- Distributed hosts within the same metropolitan area
- Core curriculum
Since we last met...

May 2013
1st NDSA NE meeting

Oct. 2014
2nd NDSA NE meeting
A lot has happened!

- NDSR DC residents selected
- NDSR Boston Project Manager hired
- 1st Round NDSR Boston & NYC residencies started
- LC/IMLS announce call for hosts for 2nd round of NDSR DC

May. 2013
1st NDSA NE meeting

NDSR DC 1st Round

NDSR Boston hosts selected

Oct. 2014
2nd NDSA NE meeting

NDSR Boston/NYC 1st Round
In DC

- NDSR DC residents selected
- May. 2013 1st NDSA NE meeting
- Oct. 2014 2nd NDSA NE meeting
- LC/IMLS announce call for hosts for 2nd round of NDSR DC
- NDSR DC 1st Round
DC pilot: positive benefits

100% of the residents said the program:

• Increased their knowledge and skills in digital stewardship
• Strengthened their professional networks
• Made them more competitive in the job market
DC pilot: areas to improve

• Program logistics (communication, management, etc.)
• Host engagement
• Curriculum beyond the immersion week
• Building a national program out of the regional NDSR programs
In Boston (and NYC)

- May 2013: 1st NDSA NE meeting
- NDSR Boston Project Manager hired
- NDSR Boston hosts selected
- NDSR Boston residents selected
- Oct. 2014: 2nd NDSA NE meeting
- 1st Round NDSR Boston & NYC residencies started
- NDSR Boston/NYC 1st Round
Boston model refinement

• Extended curriculum
• Formal development plans, residency requirements, progress tracking
• More host involvement
• Preservation-specific focus for resident projects
• Additional roles for instructors and community
Timeline

- Sept. 2014: 5 residents in NY, 5 residents in Boston
- June 2015: 5 residents in NY, 5 residents in Boston
- Sept. 2015: 5 residents in NY, 5 residents in Boston
- June 2016: 5 residents in NY, 5 residents in Boston
Boston hosts 2014-15
Resident applicants - state

MA (8)
NY (3)
PA (3)
IN (2)
NJ (2)
NY (3)
PA (3)
MA (8)
NJ (2)
Boston applicants - school

- Simmons
- NYU
- IU
- Rutgers
- UA
- UW
- UMich
- UDel
- UCLA
- Pitt
- DU
- UI
- LSU
- Pratt
- Drexel

INSTITUTE of Museum and Library SERVICES

NDSR Boston
Boston applicants - degrees

- MS LIS (11) 41%
- MLIS (9) 33%
- MLS (2) 7%
- MA MIAP (3) 11%
- MA Other (1) 4%
- MSI LIS (1) 4%
Boston Residents 2014-15

Jen LaBarbera
Host: Northeastern
MLIS

Joey Heinen
Host: Harvard
MA MIAP

Rebecca Fraimow
Host: WGBH
MA MIAP

Tricia Patterson
Host: MIT
MS LIS

Samantha DeWitt
Host: Tufts
MS LIS

Jen LaBarbera
Host: Northeastern
MLIS

Joey Heinen
Host: Harvard
MA MIAP

Rebecca Fraimow
Host: WGBH
MA MIAP

Tricia Patterson
Host: MIT
MS LIS

Samantha DeWitt
Host: Tufts
MS LIS
Nancy McGovern, MIT Libraries

NDSR Curriculum
Curriculum

• Immersive week to provide a landscape view of digital stewardship
• Scheduled events that build on pre-existing community events whenever possible
• Group activities that extend experiences and build community
• Individual residency requirements tailored to projects, skills, and interests
Immersion week

• Launched the residencies
• Morning instruction
  – DPOE*: Identify, Select, Store, Protect, Manage, Provide
  – Interactive with examples pertaining to projects
• Afternoon demonstration and hands-on
  – Demo relevant tools for projects
  – Review workflow scenarios and practical examples
• Discussion and review
• Participation by hosts, instructors

*DPOE = Digital Preservation Outreach and Education
Resident development plans

• Developed by the resident in consultancy with their mentor

• Documents:
  – **Project-specific** topics, skills & tools to learn so that the project goals can be achieved
  – **Professional development** goals – topics & tools to advance resident’s knowledge and skills
Residency events

• Scheduled curriculum events
  – In-person presentations, demonstrations, etc.
  – On-line activities, e.g. webinars plus discussion

• Group activities
  – Events at host institutions
  – Participation in regional events
  – Tours, visits and networking
  – Collaborate on organizing events

• Capstone event
September residency events

• Optical Media, NDSA Standards & Practices Working Group Call

• Launch of residents’ blog (ndsrboston2014.wordpress.com)

• The Information: A History, A Theory, A Flood by James Gleick, NEASIST Program Committee

• MIT Fellows event

• Open house, MIT Digital Sustainability Lab
Some upcoming events

• Site visits to NEDCC, JFK library
• NYC & Boston Google Hangout
• Jan. 2015 – Open mid-year event
• Monthly events at NDSR Boston hosts
• Mar. 2015: NDSR DC/Boston/NYC presentation at NEA/MARAC
• May 2015: Open NDSR Boston capstone event
Samantha DeWitt, Host: Tufts University

Institutional Knowledge of Research Data at Tufts University
Samantha DeWitt

Tufts University
Tisch Library
Medford, MA

Institutional Knowledge of Research Data
Tufts University

10 Schools:

- School of Arts and Sciences
- School of Engineering
- School of Medicine
- School of Dental Medicine
- Sackler School of Graduate Biomedical Sciences
- Friedman School of Nutrition Science and Policy
- Cummings School of Veterinary Medicine
- Fletcher School of Law and Diplomacy
- College of Special Studies
- Tisch College of Citizenship and Public Service

- Tufts was first named a Research Category I university by the Carnegie Foundation in 1989

- Today Tufts is rated as a university with "very high research activity"
Tufts University Project Goals:

• Help Tufts gain a more complete understanding of the research data produced by its faculty, research staff, post docs and graduate students.

• Investigate strategies for producing metadata objects for the Fedora-based digital repository. These objects would represent Tufts-created datasets, which would be either referenced (more likely) or stored in the repository (less likely).
Q: Why is it important for Tufts to understand the kind of research data it produces and how does this project relate to the missions of the Tufts libraries?

A: Understanding the scope of present and future datasets enables Tufts to better understand its data management and stewardship obligations.

_Tisch Library Mission Statement_

“...The library acts as a gateway and a contributor to networked information in and outside of Tufts.”
In providing assistance in research data management, Tufts libraries augment the work of the university’s research offices by helping to:

• Assure that Tufts is in compliance with funders’ data management mandates.

• Create an environment that encourages and facilitates research.
The collections of the Tufts Digital Library support teaching and research at Tufts and feature resources of enduring value, created by the community’s members.

http://dl.tufts.edu/
Fedora Object Model

Control Groups:

- Internal XML Content
- Managed Content
- Externally Referenced Content
- Redirect Referenced Content
WGBH Digital Media Preservation Project

Rebecca Fraimow, Host: WGBH
Major initiatives…
...and daily operations

- Stock sales
- Production assistance
- Researcher assistance
- Administrative metadata management
- Descriptive metadata management
- Rights metadata management
- Ingest into DAM
- LTO backup and maintenance
Phase One
Phase Two
Phase Four

AAPB Digital Preservation 101
Joey Heinen, Host: Harvard Library

Format Migration Plans and Framework for Harvard Library
Joey Heinen

National Digital Stewardship Resident
Format Migration Plans and Framework for Harvard Library - Kodak PhotoCD, SMIL playlists, RealAudio
File Formats Migration – What are the real hazards

Research – A Manic Process

Compiled Bibliography, reviewed, and culled to 30+ materials

- Data Migration Projects (large data centers)
- Format-specific case studies
- Workflow design and integrated systems (some tools)
- Development of new tools
- Understanding formats
Tools vs. Experiments – A few red herrings

Examples – AONS (Automatic Obsolescence Notification System)
GREAT concept, but major pieces missing…

- No authoritative/controlled index to reference
- Lack of robustness to characterize objects and identify inherent risks

http://www.willanswer.com/kbinfo/Main.aspx?Section=KBTourStart
Characterization – More Missing Pieces

Even within formats there can be derivations in how that object is delivered and represented:

Formats can vary in their supported fonts, types of embedded content, how technical information is formatted and where it is located in the file, etc.

Data “chunks” in Broadcast WAVE

Attribute Overlap in PDF


From Source to Target – Not a One-Hop Journey!

Important metadata from both!!!
Conclusions Drawn:

- Characterization is an essential step, content models should be tested to determine adequacy. All aspects of the workflow need to have a shared understanding of format complexity.

- Digital “objects” may be multivalent – external relationships are part of characterization and need to be considered in tandem with migration.

Next Steps:

- Map research/bibliography to migration workflow drafts: overall framework, format and deliverables (artifacts), context/stakeholders.

- Deep exploration of the Digital Repository Services (DRS) and how the workflow interfaces with this system.

- Into the Belly of the Beast: Intensive research on the three formats.

Joey Heinen
joseph_heinen@harvard.edu
ndsrboston2014.wordpress.com
@joey_heinen  #ndsr
Jen LaBarbera, Host: Northeastern University

Channeling Streams of Archival Records: Northeastern University
Project Goals

1) Born digital materials: Create a workflow for ingesting born-digital content to new digital repository service *(in progress)*

2) Develop digital preservation plan for NU Libraries *(beginning November/December 2014)*

3) Legacy digital materials: Ingest obsolete legacy digital materials (e.g. floppy disks) to DRS *(beginning early spring 2015)*
Ingesting born-digital materials

**Goal:** Transfer all materials from current states / platforms (Omeka, external hard drive, local server, Google drive, etc) to new Fedora-based repository

**Challenges:**
- metadata currently in Dublin Core, needs to be in MODS
- managed by doctoral student who has minimal time to dedicate
- agreements are scattered
- new DRS is still in pre soft launch stage
- others TBD
Progress - Literature Review

Literature Review - list of resources

Digital Preservation and the Digital Humanities

Born Digital: Guidance for Donors, Dealers, and Archival Repositories
http://www.cirr.org/pubs/reports/pub159/pub159.pdf

Digital Forensics and Born-Digital Content in Cultural Heritage Collections
http://www.cirr.org/pubs/reports/reports/pub149/pub149.pdf

A Survey of Digital Cultural Heritage Initiatives
http://www.cirr.org/pubs/reports/reports/pub118/pub118.pdf

Digital Humanities and the Library (link roundup)
http://miniamposner.com/blog/digital-humanities-and-the-library/#datapers

Digital Humanities SPEC Kit (ARL)

Checking your Digital Content: What is Fixity, and When Should I Be Checking It?

Repository to Repository Transfer of Enriched Archival Information Packages
Progress - Workflow
Progress - Workflow

1. Select materials for ingest
2. Select metadata fields for each item
3. Determine storage plan for physical media
4. Identify rights and agreements for each item
5. Crosswalk metadata (Dublin Core to MDCS)
6. Create manifest (checksums, index, etc.)
7. Package all materials for submission
8. Upload to DRS
9. Audit/quality control
Progress - Workflow
Next Steps

- Complete workflow for born-digital materials
- Test born-digital workflow on Our Marathon, revise, test again, revise…
  - (in the background) DRS soft launch (November), user testing (Nov-Dec), DRS full launch (January)
- Next (concurrent) project: Digital Preservation Plan
Making Music Last: Preservation Planning for ‘Music at MIT’ Digital Audio Content

Tricia Patterson, MIT Libraries
THE CONTEXT

- Music is the 2nd most popular minor at MIT
- Lewis Music Library is a subject-specific library (obviously)
- Increased support for digital content management at MIT libraries
- Ongoing project to inventory, digitize, preserve, and facilitate access to music content
- Workflow from this can be applied to other types of digital content
THE PROJECT

- Gap analysis of workflow documentation
- Test and enhance the life cycle workflow
- Create general digital content management workflow
- Create digital audio use case
- Help evaluate dissemination options for audio
- Consider extensibility of access platform
WHAT IS A WORKFLOW?
THE WORK!
MORE OF THE WORK!

- Avalon evaluation
- DAMP workflow subgroup
- Digital audio music project (DAMP)
- Digital content management workflow/gap analysis
- Workflow/gap analysis literature review
- Digital audio scan
THE END
Thank You!

Questions?