Digital Curation in Action

Matt Schultz
Texas Library Association, Wednesday, April 9th, 2014
Curators are traditionally experts in cultural selection, collection maintenance, and exhibition...but also human perception and learning: sight, sound, attention, and interaction...
What is Curation?

Curation is traditionally the practice of informing & educating through a carefully crafted immersive experience.

Photo courtesy Flickr user Kwong Yee Cheng: https://flic.kr/p/dwSynF
What is Digital Curation?

- Digital Curation can tend to mean different things to different libraries
  - Public and school libraries tend to focus on information architecture and Web 2.0
  - Academic/Research libraries tend to focus on scholarly services and repositories
What is Digital Curation?

• Digital curation as a process of maturation

• Digital Libraries and Archives - moving our catalogs and collections online

• Digital Preservation - ensuring that our digital collections remain accessible & usable

• Digital Curation is the on-going act of bridging digital preservation with our digital libraries and archives and enriching access to our digital collections
Digital Curation in Action

http://www.dcc.ac.uk/resources/curation-lifecycle-model
Digital Curation in Action

- University of Texas Digital Repository
  - Submission Services to Established Collections
  - Community Creation Services
  - Access Services
  - Preservation Services
    - Secure backup
    - Storage media refreshment
    - File format migration (including possible migration to preferred formats during submission)

http://repositories.lib.utexas.edu
Digital Curation in Action

- Texas Digital Library
- Provides DSpace hosting services (TACC)
- Connects to DuraCloud
- Charter member of the Digital Preservation Network
- Data Management Services
- Data discovery

http://www.tdl.org
Digital Curation in Action

- University of North Texas Libraries

  ![Portal to Texas History](http://www.library.unt.edu)

- Aubrey (access) & Coda (preservation)

- Curation micro-services
Digital Curation in Action

- The trend toward Curation Micro-Services
  - California Digital Library
  - Other examples
    - Archivematica
  - iRODS

https://wiki.ucop.edu/display/Curation/Microservices
https://www.archivematica.org/wiki/Main_Page
https://irods.org
Digital Curation in Action

- MetaArchive & use of micro services
  - UNT PREMIS Event Service
    - Ingest & Replication completion
  - Checksum Validator
  - BagIt

https://github.com/metaarchive
https://github.com/unt-libraries
Digital Curation in Action

- NEH Guidelines for Digital Newspaper Preservation Readiness

- BagIt Case Study

  - Packaged member’s digital newspaper collections (150GB) using BagIt

  - Used a Bag Split Utility to break the 150GB Bag into 5 smaller 30GB Bags - made it easier to transfer by network

  - Rebuilt the 5 smaller Bags into the original Bag with the help of a Bag Unsplit Utility

  - BagIt gave us the tools to validate that no content was damaged
Digital Curation in Action

- NEH Guidelines for Digital Newspaper Preservation Readiness
  - Inventorying
  - Organization
  - Format Management
  - Metadata Management
  - Checksum Management
  - Packaging

http://www.educopia.org/publishing/gdnpr
Digital Curation in Action

- IMLS Guidance Documents for Lifecycle Management of ETDs
  - Guidelines for Implementing ETD Programs – Roles and Responsibilities
  - Guide to Access Levels and Embargoes of ETDs
  - Briefing on Copyright and Fair Use Issues in ETDs
  - Guidelines for Collecting Usage Metrics and Demonstrations of Value for ETD Programs
  - Managing the Lifecycle of ETDs: Curatorial Decisions & Practices
  - Metadata for ETD Lifecycle Management
  - Guide to ETD Program Planning and Cost Estimation
  - Guide to Options for ETD Programs

http://www.educopia.org/publishing/gdlmetd
Digital Curation in Action

Lifecycle Curation Tools

- Conceptualise
- Virus Checking
- Dispose
- Unix
- JHOVE2
- DROID
- FITS

PREMIS Event Record-Keeping

Preservation Metadata
Digital Curation in Action

- Curation - What Libraries Are Up To?
  - File naming conventions
  - Selection (appraisal)
  - Organizing data for deposit & preservation
  - Virus checking
  - Fixity checking
  - Format identification
Digital Curation in Action

• Curation at Home - Getting Experience
  • ClamXav (virus checking)
  • DROID (format identification)
  • Bagger (inventory & checksums)
  • TAR (packaging)
  • Glacier (long-term preservation)
Digital Curation in Action

• Resources

  • Digital Curation Centre (UK) - http://www.dcc.ac.uk

  • CDL Curation Microservices - https://wiki.ucop.edu/display/Curation/Microservices

  • Digital Preservation Outreach & Education Program (DPOE) - http://www.digitalpreservation.gov/education/

  • National Digital Stewardship Alliance (NDSA) - http://www.digitalpreservation.gov/ndsa/

  • Digital-POWRR Project - http://digitalpowrr.niu.edu

  • COPTR - http://coptr.digipres.org/Main_Page

  • Open Planets Foundation (OPF) - http://www.openplanetsfoundation.org

  • ArchiveTeam: http://archiveteam.org
Digital Curation in Action

• Upcoming Events
  • 2014 ICPSR Curation Workshop, July 28-Aug 1: https://www.icpsr.umich.edu/icpsrweb/sumprog/courses/0149

• Ongoing Events
  • Code4Lib Annual & Regional Meet-Ups: http://code4lib.org
  • CurateCamp: http://curatecamp.org
  • CurateGear: http://ils.unc.edu/digccurr/index.html
Digital Curation in Action

- Important Listservs
  - Digital Curation Google Group: https://groups.google.com/forum/#!forum/digital-curation
  - Code4Lib: https://listserv.nd.edu/cgi-bin/wa?SUBED1=CODE4LIB&A=1
  - digipres: http://lists.ala.org/wws/info/digipres
Digital Curation in Action

- Digital curation is as much about us as it is about the technologies we use

- We are all becoming digital curators

- Digital curation is a maturing process for libraries - bridging preservation & access

- Digital curation is about knowing where and when to take action - targeted technologies

- Dive in and start curating - lots of resources!
Thank You!

Questions?
Contact

Matt Schultz
Program Manager
MetaArchive Cooperative
www.metaarchive.org
matt.schultz@metaarchive.org