

Digital Projects/Public Viewer Digital Preservation Storage Update National Preservation Projects

Robert Cartolano, Stephen Davis, Jim Neal

April 23, 2014

Theme: From Local to National

- Local Effort
 - Digital Collections
 - Public Digital Collections Viewer (DCV)
- Local and National
 - Digital Preservation Storage
 - Sustainable Funding
 - Sustainable Technology
- National Preservation Strategy



Digital Collections Projects

Columbia University Libraries / Information Services has made commitments ...

- to granting agencies to provide long-term digital archiving for digital content created with *grant funds*
- to third-party content creators to provide permanent access to *born-digital content* acquired from them
- to continuing to collect and preserve *archival collections*, now partly or wholly born-digital content
- to permanently preserve *University-generated archival and research content*



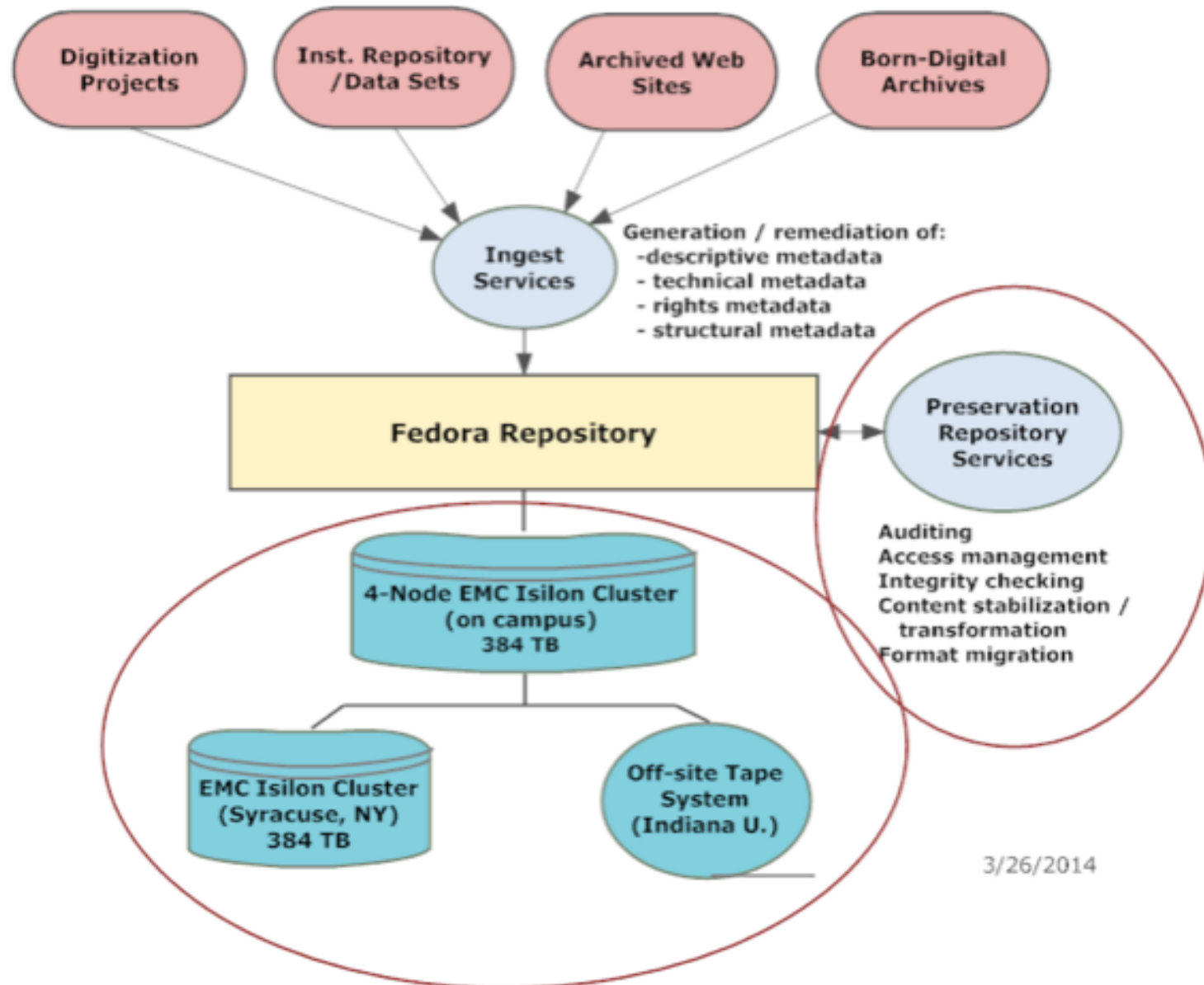
Digital Collections

We must preserve and provide access to ...

- Local Digitization Projects
- Preservation-Related Digitization
- Institutional Repository / Data Sets
- Born Digital Archival Content
- Archived Web Sites
- Super Dark Archives – *highly secure*



Current Digital Asset/Preservation Architecture



Example: Digital Collection Project

COLUMBIA UNIVERSITY LIBRARIES LINDQUIST Start a new search... Search Q

G.E.E. LINDQUIST NATIVE AMERICAN PHOTOGRAPHS

THE BURKE LIBRARY ARCHIVES @ COLUMBIA UNIVERSITY

ABOUT EXPLORE FACES CONTACT SELECTED ITEMS (0) SEARCH HISTORY

G.E.E. Lindquist Native American Photographs is an online presentation of the 1322 photographs, 124 postcards, 388 negatives, and 34 glass plate negatives/lantern slides, which derive from the G.E.E. Lindquist Papers archival collection at The Burke Library.

They depict the people, places, and practices of Native Americans and their communities from at least 34 States, plus Canada and Mexico in the period from 1909-1953. The majority of the images were taken by G. E. E. Lindquist (1886-1967), an itinerant representative of the ecumenical Home Missions Council of the Federal Council of Churches.

FEATURED TOPICS PLACES TOPICS NAMES CREATORS

A selection of items relating to the topic Lakes.

Select 

Lake Superior Near Grand Port...
Item Number: 453
[View Item](#)

Select 

Lake Superior Near Grand Port...
Item Number: 452
[View Item](#)

Select 

Shore and View of Lake Superio...
Item Number: 462
[View Item](#)

Featured Topics



Academic Commons

 COLUMBIA UNIVERSITY
ACADEMIC COMMONS



Search



Academic Commons is Columbia University's digital repository where faculty, students, and staff of Columbia and its affiliate institutions can deposit the results of their scholarly work and research. Content in Academic Commons is freely available to the public.



[Login](#) | [History](#)

BROWSE AC CONTENT

[DEPARTMENTS](#)
[SUBJECTS](#)

HELPFUL LINKS

[DEPOSIT YOUR RESEARCH](#)
[ABOUT ACADEMIC COMMONS](#)
[FAQ/ASK A QUESTION](#)

*New in
Academic
Commons:*

**Built-in Stabilizers and Risk
Literacy: Protecting the
Sustainability of the Insurance
Industry**

[Goto, Shigeyuki](#)

**As I sow, so shall you reap: The
different roles of different
gestures in knowledge
construction**

[Kang, Seokmin](#)

**Recovering Leisure: Otium as
the Basis of Education**

Symmetry and Probability

**Development of Systematic
Knowledge Management for
Public Health & Public Health**



COLUMBIA UNIVERSITY LIBRARIES / INFORMATION SERVICES

Older Projects - Data Migration, Refresh



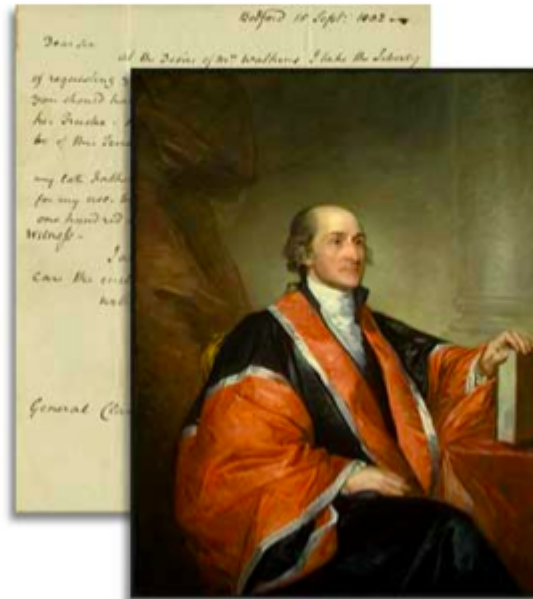
The Papers of John Jay is an image database and indexing tool comprising some 13,000 documents (more than 30,000 page images) scanned chiefly from photocopies of original documents. Most of the source material was assembled by Columbia University's John Jay publication project staff during the 1960s and 1970s under the direction of the late Professor Richard B. Morris.

Search

[More about the project >>](#)

Background Material

- ◆ [Biographical Essay](#)
- ◆ [Jay and the Constitution](#)
- ◆ [Jay and New York](#)
- ◆ [The Jay Treaty](#)
- ◆ [Jay and France](#)
- ◆ [Jay and Slavery](#)
- ◆ [Jay Print Project](#)
- ◆ [Legacy Conference](#)



About the Project

- ◆ [History & Description](#)
- ◆ [Source Collections](#)
- ◆ [Bibliography](#)
- ◆ [Rights and Permissions](#)
- ◆ [Participating Institutions](#)
- ◆ [News & Announcements](#)

Portrait courtesy of [The National Gallery of Art](#)

[Search](#) ◆ [Help](#) ◆ [Comments](#)



Web Archiving

HRWA Advanced Sites Archive Menu

HUMAN RIGHTS WEB ARCHIVE

@ COLUMBIA UNIVERSITY

COLUMBIA UNIVERSITY LIBRARIES
CENTER FOR HUMAN RIGHTS DOCUMENTATION & RESEARCH

ABOUT THE ARCHIVE FAQ NOMINATE A SITE

The Columbia University Human Rights Web Archive & Portal provides search access to archival copies of over 450 web sites in the area of human rights. Collection began in 2008 and is ongoing for active sites...[more](#)

Site URL Subject Geographic Focus Language

[Search for Sites](#)

[Search the Web Archive](#)

[Browse for Sites](#)

First Thumbnail label caption 1

İnsan Hüküqları XXI əsr- Azərbaycan Fondu	Community Appraisal & Motivation Program	Law & Society Trust
5-18 Kinyòm Chaedan	Comunidad de Derechos Humanos	The Lawyers Centre for Legal Assistance
ABONG--Associação Brasileira de Organizações Não Governamentais	CONECTAS Direitos Humanos	Legal and Human Rights Centre
Abuelas de Plaza de Mayo	Congolese Women's Campaign Against Sexual Violence in the Democratic Republic of the Congo (DRC)	Lembaga Studi dan Advokasi Masyarakat
Access to Justice	Conselho Indigenista Missionário	Ligue algerienne pour la defense des droits de l'homme
Action Group for Health, Human Rights and HIV/AIDS (AGHA)	Control arms	Ligue burundaise des droits de l'homme
Ação Brasileira pela Nutrição e Direitos Humanos	Coordinadora de Derechos Humanos del Paraguay	Ligue djiboutienne des droits humains
Adalah	Coordinadora Nacional de Derechos Humanos	Madre
Advocacy Forum--Nepal	Corporación de Promoción y Defensa de los Derechos del Pueblo	al-Majlis al-Istishārī li-Huqūq al- Insān
The Advocates for Human Rights	CRLDHT--Comité pour le respect des libertés et des droits de l'homme en Tunisie	al-Majlis al-Waṭani li-Humīyāt bi-Tūnis
AFAPREDESA	CSV, Centre for the Study of Violence and Reconciliation	Malawi Human Rights Commission
Afghanistan Independent Human Rights Commission	Danish Institute of Human	Malawi Human Rights Resource Centre
Africa Internally Displaced Persons Voice		MAP Foundation
African Centre for Democracy and Human Rights Studies		Mardu iravunk'nerē Hayastanum



Public Digital Collections Viewer (DCV)

- A system to provide access to traditional cultural heritage digital collections, audio/video collections, and born digital archival collections
- A platform for integrating existing collections and publishing new collections
- A set of tools for basic viewing / playing content in most media formats
- A framework for bringing to the surface relationships between content in Columbia's collections and content in other related collections at other locations



Who is making this happen?

- Advisory Group: Jane Gorjevsky, Chris Sala, Kate Harcourt, Melanie Wacker, Robbie Blitz, Matthew Baker, Chengzhi Wang, Carolyn Yerkes
- Project Lead: Stephen Davis
- Project Management Support: Leo Stezano
- Technical Staff: Eric O'Hanlon, Ben Armintor, Erik Ryerson



When will this happen?

Phase I – Tentative Timetable

- April-May 2014: Internal prototypes
- May 2014: CUL internal beta release
- June-August 2014: Further development, additional internal releases; user testing
- September 2014: Public beta release
- October-November: User testing, further development
- Winter 2014/2015: Public launch



What will Phase 1 content include?

200,000+ Digital Objects

- <https://wiki.cul.columbia.edu/display/PDCV/DCV+Potential+Digital+Content>
(Potential content)

30+ Digital Collections

- <https://wiki.cul.columbia.edu/display/PDCV/Public+Digital+Collections+Viewer+Home> (Project wiki with more info and links)



What will it not include?

- Web archives
- Finding aids
(but finding aids can link to the DCV)
- Collections / content with no 'suitable' metadata



What features will be in Phase I?

- Browsable list of collections / projects
- Keyword searching within individual projects or across all projects
- Standard facets across all collections: names, material types, collection names, repositories
- Custom facets where needed: e.g., geographic or chronological content
- Flexible displays for images; zoomable, multiresolution image functionality



Features in Phase I, cont.

- Persistent identifiers
- Function to generate a standard citation for individual objects
- Session-specific “bookbags” / “shopping carts”
- Search engine optimization features
- Full-text searching where supported by the data (e. g., IFP, University Seminars, Harrison)
- Ability to report errors
- Ability to correct errors (in Hyacinth metadata editor)



What about restricted content?

There will actually be two separate digital collection viewers:

- One available publicly, world-readable
- One restricted to on-site access in a specific custodial location (for example, for parts of archival collections, for content that is under copyright but can be made available locally as preservation copies)

Having two separate, parallel viewers will help ensure the privacy and security of restricted content.



DCV Benefits

- Supports three CUL/IS strategic goals
 - Develop Global and Special Collections
 - Digital Collections
 - Effective Interfaces and Improved Access
- Provides integrated access to CUL specialized and archival digital collections
- Aggregates search and discovery
- Reduces need for multiple separate systems
- Enables potential “bento box” discovery via CLIO
- Facilitates making CUL content available externally, e. g., to DPLA, through APIs, consistent SEO
- Ease of improvements, upgrades, sustainability



Local Digital Preservation Storage

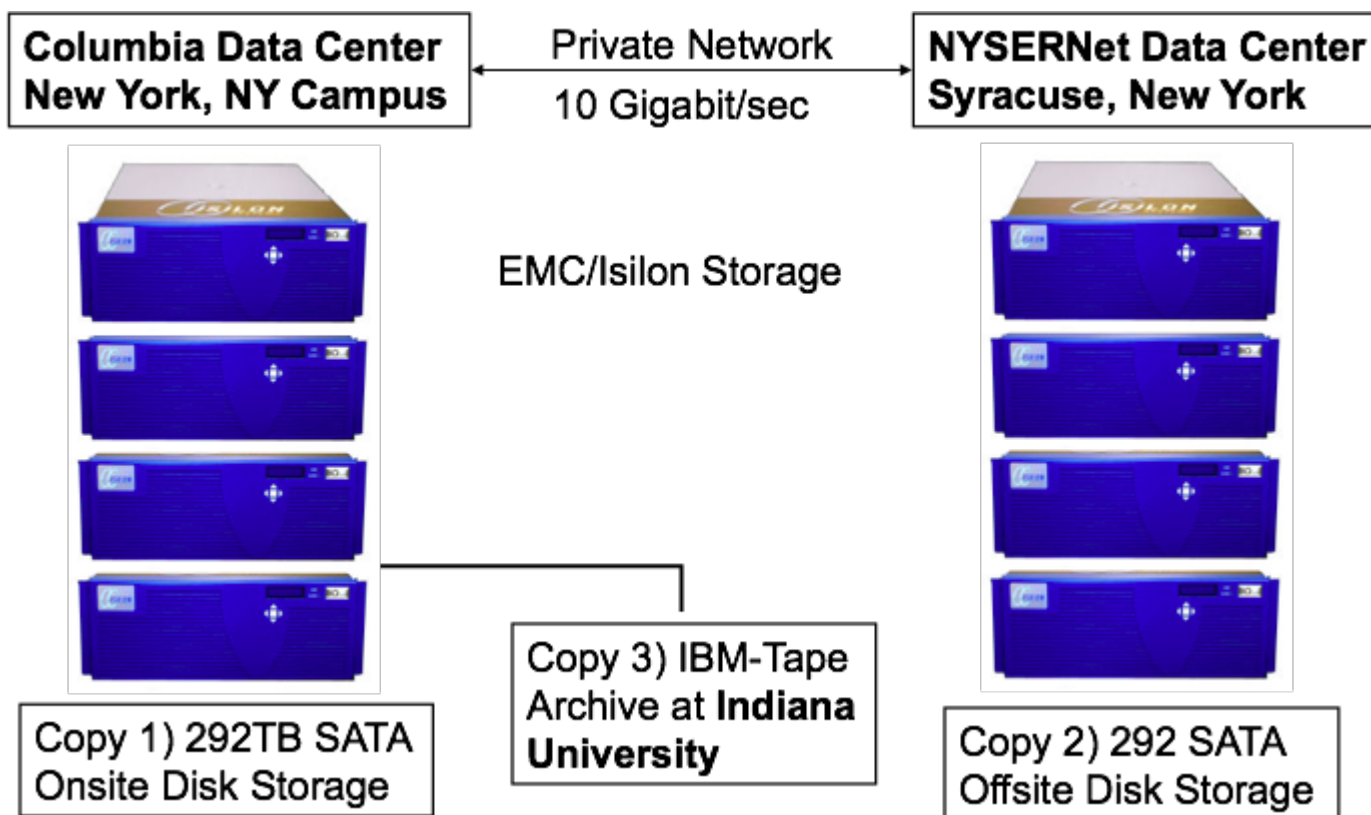
- Goals
 - Risk-averse technology approach
 - Different storage for different needs
 - Move towards using national systems
- Local Preservation - 2013 thru 2017
 - 300 Terabytes total capacity
 - ca. 140 Terabytes of current, committed projects
- Beyond 2017
 - Local systems - primarily for services, performance
 - National systems - primarily for certified preservation storage (TRAC)



Columbia Preservation Storage

2007 – Purchased and built by Columbia

2013 – Mostly at Columbia, some at Indiana



National Preservation Strategy

- Digital Preservation Network
 - Strategy and implementation plan for national preservation infrastructure
 - 65 member institutions
 - Coordinated with AAU Presidents, Internet 2
- Five DPN Nodes
 - APTTrust (University of Virginia) - member
 - Chronopolis
 - HathiTrust (University of Michigan) - member
 - Stanford Digital Repository
 - University of Texas Digital Repository

<http://www.dpn.org/>



National Preservation Strategy (cont'd)

- DuraSpace
 - Infrastructure for durable, persistent access to data
 - Sustainable software, communities, funding
 - 90+ sponsors

<http://duraspace.org/>

- Hydra Partners
 - Building sustainable software tools
 - 20+ partner institutions

<http://projecthydra.org/>

- Research Data Alliance
 - Enable open sharing of data
 - 35+ countries at 3rd plenary conference

<https://rd-alliance.org/>



Academic Preservation Trust (APTrust)

- Led by University of Virginia, 17 partner institutions
- APTrust is working toward TRAC certification
- First ever repository to be built from the ground up taking TRAC into account.
- Certification Working Group established and will be advising and consulting with the APTrust staff and partners on TRAC objectives.
- Initial development work is proceeding at the level of Digital Object Management and Infrastructure.

<http://aptrust.org/>



Examples of TRAC Requirements

- “The repository shall have an appropriate succession plan, contingency plans, and/or escrow arrangements in place in case the repository ceases to operate or the governing or funding institution substantially changes its scope.”
- “The repository shall have short- and long-term business planning processes in place to sustain the repository over time.”
- “The repository shall have contracts or deposit agreements which specify and transfer all necessary preservation rights, and those rights transferred shall be documented.”
- “The repository shall have the appropriate number of staff to support all functions and services.”
- “The repository shall have and use a convention that generates persistent, unique identifiers.”



Emerging National Digital Preservation Strategy

- Institution-level funding required
- Local efforts integrated with national systems
- Succession rights planning
- Governance
 - HathiTrust - Bob Wolven
 - DPN - Jim Neal - DPN



Questions

