

# Supercharged Digital Collections: Moving to the fast lane with scalable open source infrastructure



Photos from University of Utah, Multimedia Archives Photo Collection, Gary Dean Brown

# Presenters

Anna Neatrou

Brian McBride

Matt Brunsvik

Harish Maringanti

Other Presenters/contributors:



Jeremy Myntti



Alan Witkowski

# What we'll cover today

- ▶ History
- ▶ SIMP demo
- ▶ Migration
- ▶ Lessons learned

The screenshot displays the library's digital archive interface. At the top, there's a navigation bar with 'Home', 'Browse All', 'Ask a Question', 'Log In', 'Favorites', and 'Help'. A search bar is present with a 'Search' button and a 'within results' dropdown. A blue banner prompts users to 'Please take a moment to fill out our survey'.

The main content area is titled 'Browsing items in: Univ of Utah - Multimedia Archives Photographs' and shows 68009 results. It includes a 'Narrow your search by:' sidebar with filters for Type, Creator, Keywords, Date, and Title. The main list shows items like 'Geodesic Dome', 'Mister Walrus', and '#10 Springs gully, 1979'. A 'Limit your search' section on the right allows filtering by Type, Format, and Collection Name, with a list of 11 collection names and their respective item counts.

Below the search filters, a table displays search results. The table has columns for 'Title', 'Collection Number And Name', and 'Photo Number'. Three results are visible:

	Title	Collection Number And Name	Photo Number
1	"\$800 Reward" poster of capture of polygamists John Taylor and George O. Cannon.	P0056 Church of Jesus Christ of Latter-day Saints History Photograph Collection	P0056n01_01_07
2	"...one of the deepest cuts in the world" in Cataract Canyon, about ten miles below the Green River.	P1049 Harold Herbert Leich Photograph Collection	P1049n1_11_43
3	"1287 Crystal Avenue, the home that we built and moved into October 15, 1923, loved and lived together in for 33 years. Two years later I moved into the home Mel, Emma Lou, and I had built at 1965 St. Mary's Drive."	P1113 Grace Richards Warner Photograph Collection	P1113n192

# History of UU Marriott Digital Library

2001

Launched  
ContentDM



# History of UU Marriott Digital Library

2002

Launched Utah  
Digital Newspapers  
(UDN)



# History of UU Marriott Digital Library

2003

UDN receives IMLS  
NLG grant



# History of UU Marriott Digital Library

2010

UDN reaches 1  
Million pages  
milestone



# History of UU Marriott Digital Library

2001

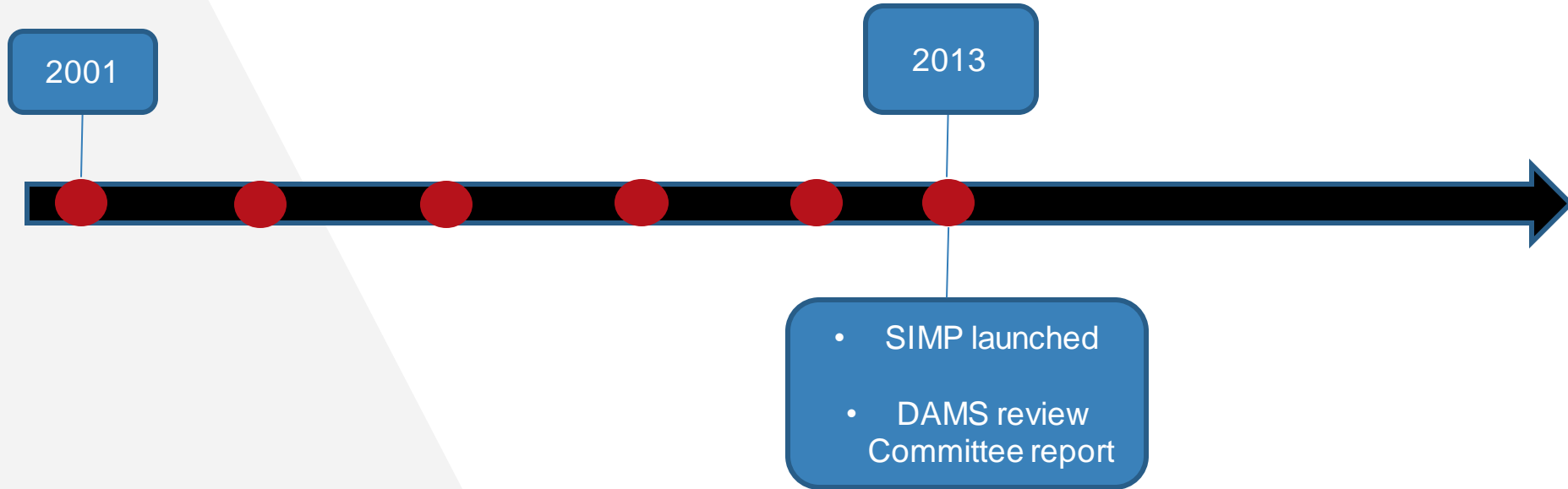
2012

- Purchased Rosetta for long term archiving
- SIMP development started

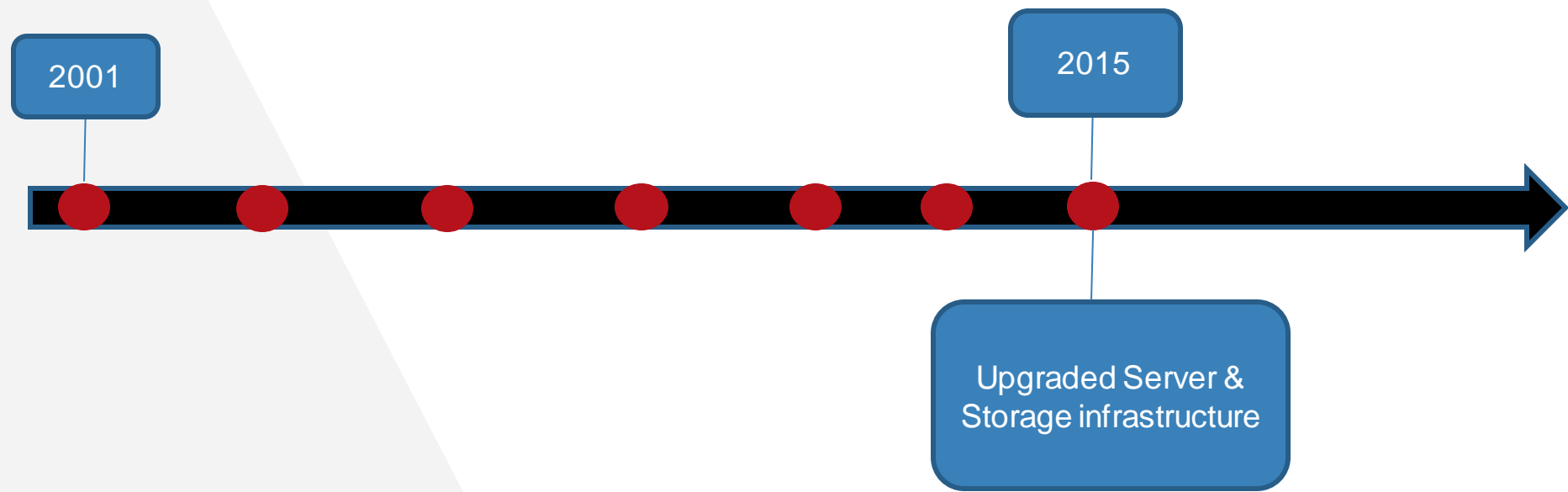




# History of UU Marriott Digital Library



# History of UU Marriott Digital Library



# History of UU Marriott Digital Library

2001

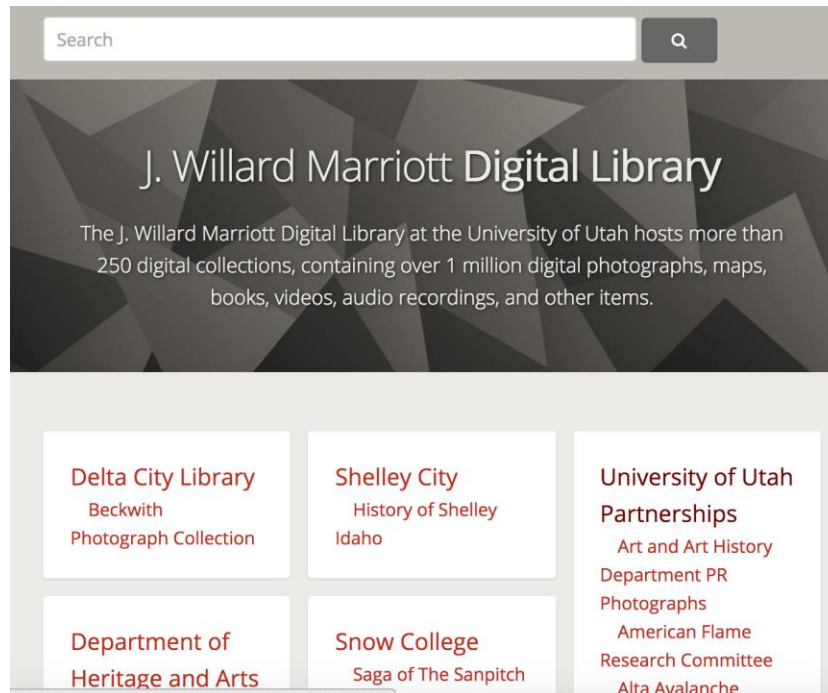
2016



UDN migrated (Jan – June)  
  
All Other digital Collections migrated (June – Dec)

# Marriott Library Digital Collections

- ▶ 269 collections
- ▶ 717,858 IEs (items)
- ▶ 139,583 compound objects
- ▶ 12 file types  
(images, videos, pdfs...)
- ▶ 3 TB storage
- ▶ Over 50 partners (internal and external) hosting collections with the Marriott Library



The screenshot shows the homepage of the J. Willard Marriott Digital Library. At the top, there is a search bar with the text "Search" and a magnifying glass icon. Below the search bar, the title "J. Willard Marriott Digital Library" is displayed in a large, white font against a dark background with a geometric pattern. Underneath the title, a paragraph of text reads: "The J. Willard Marriott Digital Library at the University of Utah hosts more than 250 digital collections, containing over 1 million digital photographs, maps, books, videos, audio recordings, and other items." Below this text, there are six white boxes arranged in a 2x3 grid, each containing the name of a partner collection and a brief description of its contents. The collections listed are: Delta City Library Beckwith Photograph Collection, Shelley City History of Shelley Idaho, University of Utah Partnerships (Art and Art History, Department PR Photographs, American Flame Research Committee, Alta Avalanche), Department of Heritage and Arts, Snow College Saga of The Sanpitch, and another collection partially visible at the bottom left.

<https://collections.lib.utah.edu/>

# Utah Digital Newspapers (UDN)

- ▶ ~1.926 Million pages
- ▶ ~21.64 Million articles
- ▶ ~138 Newspaper titles
- ▶ Article level newspaper data
- ▶ Many unique papers not in Chronicling America

The screenshot shows the Utah Digital Newspapers website. At the top, the title "Utah Digital Newspapers" is displayed in a large, serif font, with the subtitle "Creating Citizen Historians" below it. A navigation bar contains links for "Home", "Browse by County", "Newspaper Holdings", "Utah Newspaper Hall of Fame", and "Contact".

The main content area is divided into several sections:

- Search All Newspapers:** A search box with a "Go" button and a "Search/Browse a Newspaper:" section with a "Select Newspaper:" dropdown menu.
- Recent Additions:** A list of newspaper titles and their active years:
  - [The Beaver County News](#) 1957-1981
  - [The Milford News](#) 1947-1956
  - [Iron County record](#) 1979-1982
  - [The Springville Herald](#) 1964-1969
  - [Utah Daily Chronicle](#) 1963-1972
  - [Davis County Clipper](#) 1988-1992
- In the News:** A section featuring a "familytree magazine" logo.
- Featured Title:** A preview of a newspaper page titled "Manti Messenger" with the headline "Manti Kills Fatted Calf". Below the preview is a caption: "Click on the image to go to this issue."
  - Help Us Improve:** A request for feedback with a link to an "online survey".
  - Help:** Links for "About the Program" and "How to Use this Website".

At the bottom left of the page, there is a link: "75 Best Websites for US State Genealogy Research in 2015".

<https://digitalnewspapers.org/>

# SIMP Tool for metadata management

- ▶ Initially designed to facilitate management between CONTENTdm and Rosetta, our preservation system
- ▶ Platform agnostic and modular, provides ability to use other DAMs/Preservation Systems
- ▶ Options for automatically extracting some metadata (format, OCR text)
- ▶ Updated in 2016 to support Solphal

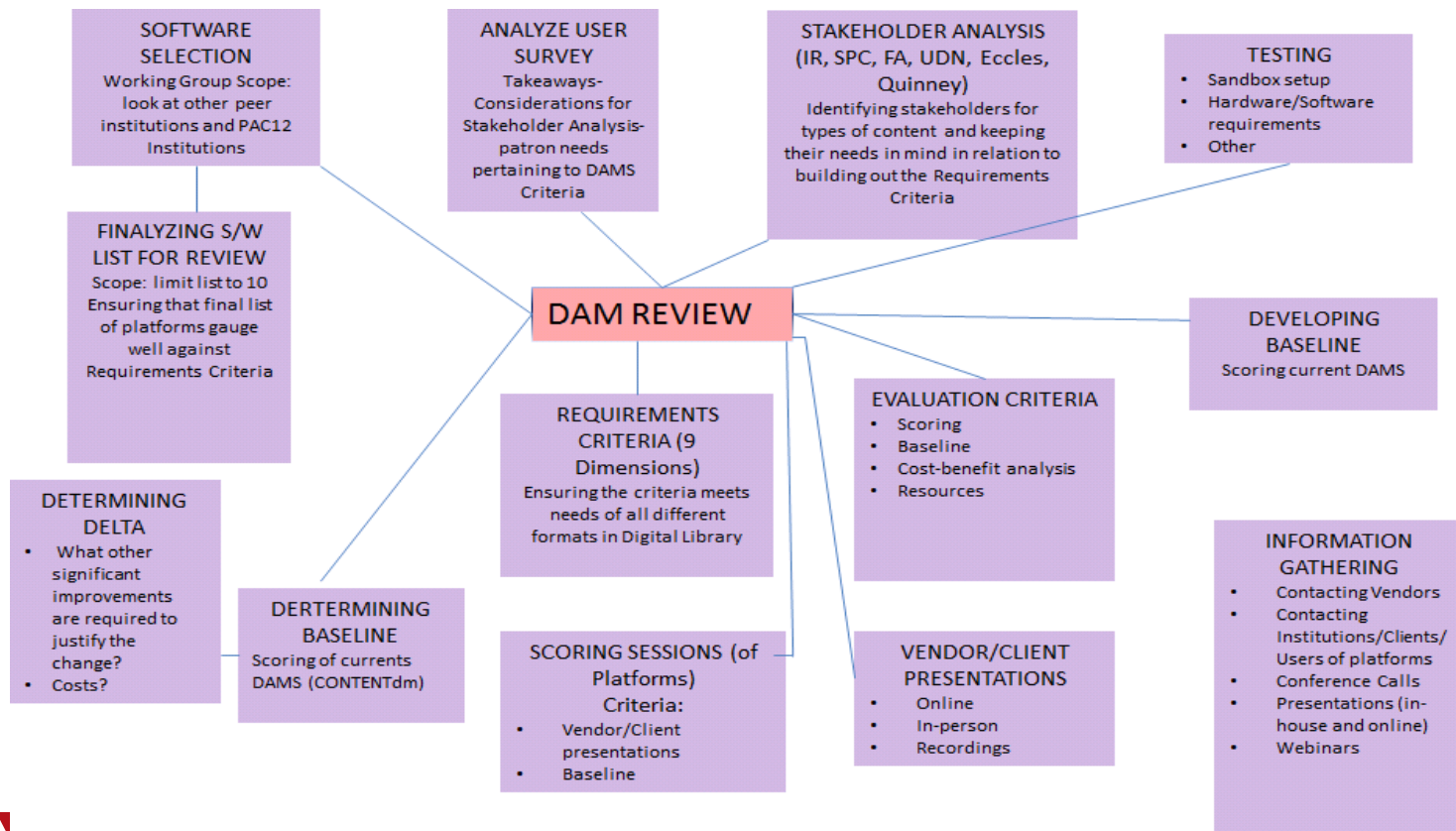
The screenshot shows the SIMP tool interface. At the top, there are navigation tabs: "Browse Servers", "Assess Packages", "Admin", "Queue (0)", and "Sandbox". Below these are action buttons: "Invert", "Select all", "Select without ARK", "Select unedited", "Select unsent", "Select unapproved", "Select cancelled", and "Select none". A dropdown menu shows "uu\_wss - Western Soundscape" with options: "Set Collection", "Create ARKs", "Edit Metadata", "Approve", "Download Metadata", "Select Archival Destination", and "Queue for Ar".

Package Name	Title	Collection	Last Editor	Approved By	ARK	Archive Status
<input type="checkbox"/> 252600103_00oi3hep	Watercolor Southwest fifth an...	uum_wfws	u0035221	u0751700	ark:/87278/s6nr5gff	Queued by u0751
<input type="checkbox"/> 252600105_01oi3hep	41st Annual exhibition of the ...	uum_wfws	u0035221	u0751700	ark:/87278/s6n04qrg	Queued by u0751
<input type="checkbox"/> 252600108_02oi3hep	Western Federation of Waterc...	uum_wfws	u0035221	u0751700	ark:/87278/s6h7401j	Queued by u0751
<input type="checkbox"/> 252600109_03oi3hep	17th annual exhibition of the ...	uum_wfws	u0035221	u0751700	ark:/87278/s6cg37b8	Queued by u0751

More details available in D-Lib article:

<http://www.dlib.org/dlib/july14/neatour/07neatour.html>

# DAMS Review in 2013



# DAMS Review in 2013

CONTENTdm®

CHRONICLING AMERICA  
Historic American Newspapers

MDID

ExLibris  
Rosetta

hydra

EQUELLA®  
PEARSON

XTF  
eXtensible Text Framework

omeka

Cumulus

bepress™

ARCHIVES WEST

ORBIS  
CASCADE  
ALLIANCE



# DAMS Review - requirements criteria

## Patron ease of use

- End-user experience
- Search options and accuracy - discoverability of the content
- Handle all formats in similar fashion for all browsers
- ADA compliance

## Types of content supported

- Major content categories (e.g. IR, newspapers, EAD)
- File types and datasets
- Tiers of content (both object and item level)
- Media streaming

## Issues related to ingest/conversion/exit

- Workflows for adding new content
- Converting existing collection data
- Barrier(s) to exit

## Collection administration

- Interface design
- User permissions
- Website configuration
- Statistics/Reporting/Logs

## Metadata administration

- Metadata creation, editing, and maintenance options
- Faceting
- Copyright/Embargo
- Search Engine Optimization (SEO)

## Integration with other library platforms

- Discovery layer (e.g. Primo)
- Digital preservation system (e.g. Rosetta)
- Locally developed tools (e.g. SIMP Tool)

## Technical infrastructure/administration

- Scalability of collections/content
- Hardware/Server(s) specs
- API support
- System and server installation, configuration, and upgrades
- Internal resources required to run the system

## Technical support

- Training
- Help desk

## Future and strategic directions

# DAMS Review Outcome - Solphal



# Hardware Requirements - CONTENTdm vs Solphal

## CONTENTdm

- 2 servers (UDN & Digital Collections)
- Digital Collections (2x12 Core 3.0 Ghz, 64 Gb mem) UDN (2x8 Core 2.4 Ghz, 96 Gb mem)
- Scalability issues
- Limited customizations

## Solphal

- 3 VMs (indexing services, UDN, and Digital Collections)
- VMs avg. 2 cores, 3 Gb mem
- Extremely small footprint (hardware and power savings)
- Highly scalable and sustainable
- Fully customizable
- Easy replication and cloud-centric

# Benefits of Solphal

- In-house expertise and support
- Student, patron, and partner convenience
  - faster response times
  - advanced search functions
  - front-end metadata editing
  - customized interface
- Substantial power, hardware, and licensing cost savings
- Improved workflows and QA/QC for all teams
- Staff time savings = ~1 less FTE supporting servers
- Demonstrating and sharing open source solutions with other institutions



# Indexing Improvements

## Utah Digital Newspapers

- 55 GB of metadata
- 21,648,462 records
- CONTENTdm
  - Full re-index ~ 1440 hrs (site is online)
  - Full re-index ~ 240 hrs (site is offline)
- Apache Solr indexing
  - Initial index ~ 2.75 hrs
  - Incremental index ~ instant

## Digital Collections

- 6 GB of metadata
- 2,248,922 records
- CONTENTdm
  - Full re-index ~ 144 hrs (site is online)
  - Full re-index ~ 72 hrs (site is offline)
- Apache Solr indexing
  - Initial index ~ 25 min
  - Incremental index ~ instant

# User Interface Improvements

## Utah Digital Newspapers

- CONTENTdm
  - Page load times (avg) 10s
  - Pdf display issues
- Solphal
  - Page load times (avg) under 500ms

Name	Status	Type ▲	I...	Size	Time	Waterfall	40.00 ms
<input type="checkbox"/> analytics.js	307		...	0 B	57 ms		
<input checked="" type="checkbox"/> search?page=3	200	docu...	...	9.0 KB	48 ms		
<input type="checkbox"/> cJZKeOuBrn4kERx...	200	font	...	20.8 KB	35 ms		
<input type="checkbox"/> MTP_ySUJH_bn48...	200	font	...	21.5 KB	31 ms		
<input type="checkbox"/> k3k702ZOKiLJc3W...	200	font	...	21.6 KB	41 ms		
<input type="checkbox"/> fontawesome-webfo...	200	font	...	65.4 KB	20 ms		
<input type="checkbox"/> 0d55f26291d87846...	200	jpeg	...	3.9 KB	8 ms		
<input type="checkbox"/> 0be4e1d772b4c77d...	200	js	...	1.3 KB	9 ms		
42 requests   303 KB transferred   <b>Finish: 405 ms</b>   DOMContentLoaded: 264 ms   Load: 382 ms							

## Digital Collections

- CONTENTdm
  - Page load times (avg) 5s
  - Pdf display issues
- Solphal
  - Custom HTML5 image viewer
  - HTML response times (avg) under 50ms
  - Complete page load times (avg) under 500ms

# Sample Solr Ingest File



```

<add>
<doc>
<field name="title_t">U Archives D. Art 1890-1939 fd 1
#2</field>
<field name="form_t">photograph</field>
<field name="source_t">University of Utah Archives</field>
<field name="work_id_t">101</field>
<field name="date_digital_t">4/20/2010</field>
<field name="rights_t">Digital Image Copyright 2010,
University of Utah. All Rights Reserved.</field>
<field name="type_t">Image</field>
<field name="imaging_technician_t">Amelia Walchli</field>
<field name="metadata_cataloger_t">Amelia Walchli</field>
<field name="file_name_t">101.tif</field>
<field name="ark_t">ark:/87278/s6sb7cp8</field>
<field name="archival_file_t">Volume1/U Archives D. Art
1890-1939 fd 1 #2.tif</field>
<field name="created_tdt">2010-04-19T00:00:00Z</field>
<field name="modified_tdt">2010-04-26T00:00:00Z</field>
<field name="oldid_t">Art2 0</field>
<field name="id">15729</field>
<field name="parent_i">0</field>
<field
name="thumb_s">/68/84/6884a28421c31d96d4d9ec466f47fecbb99af
700.jpg</field>
<field
name="file_s">/5c/bc/5cbc0793020d24aba7eb6e4443d0c6574970ac
57.jpg</field>
<field name="format_t">image/jpeg</field>
<field name="setname_s">uu_aah_artpr</field>
<field name="restricted_i">0</field>
</doc>
<doc>
<field name="title_t">U Archives D. Art 1890-1939 fd 1
#4</field>
<field name="form_t">photograph</field>
<field name="source_t">University of Utah Archives</field>
<field name="work_id_t">102</field>

```

# UDN migration vs. Digital Collections migration

## Utah Digital Newspapers

- Newspapers metadata more standardized, but not perfect
- Inconsistencies with titles, dates, paper names, and type fields
- Clean-up for nonstandard characters, newspaper names issues
- Previous architecture had one newspaper title split into multiple collections (scalability issues)
- Consistent field names and metadata meant that migration happened sooner and faster
- No external partners to contact

## Digital Collections

- More communication needed with both internal and external partners during migration
- Problem or missing metadata values in older collections not fixed (yet)
- Field label standardization
- Standardization of type and format
- Lots of clean-up work still to do in the future



# Metadata issues: Corrections

## Metadata Standardization

Establish core fields needed for faceting and queries in new DAM, and change varying fields to the following:

•Creator	Title	Mr. J. Holman Waters: Recollections of Samuel Newhouse, Salt Lake City, Utah: an interview by Hynda Rudd, July 16, 1977
•Date	Creator	Waters, J. Holman, 1912-1995
•Subject	Date	1977-07-16
•Rights	Spatial Coverage	Salt Lake City, Salt Lake County, Utah, United States
•Spatial Coverage	Subject	Newhouse, Samuel, 1853-1930; Businessmen--Utah--Salt Lake City--Interviews; Jews--Utah--Salt Lake City
•Type	Type	Text
•Title	Rights	Digital Image © 2016 University of Utah. All Rights Reserved. Contact the J. Willard Marriott Library Manuscripts Division for use and restrictions. 801-585-3082

# Metadata fields gone wild!

About the Art  
 About the Artist  
 Abstract  
 Accession number  
 Accession Number  
 Access restrictions  
 Access Restrictions  
 Access rights  
 Accompanying Material  
 Acoustic Zone  
 Added Title  
 Added title 2  
 Additional information  
 Additional Information  
 Additional Resources  
 Additonal notes  
 Address  
 Affiliation  
 Ages  
 Alt # 1  
 Alt # 2  
 Alternate Name(s)  
 Alternate title  
 AlternateTitle  
 Alternate Title  
 Alternative formats available  
 Alternative Formats Available  
 Alternative title  
 Amendments and updates  
 Anatomy  
 Annotation  
 Archival file  
 Archival Resolution  
 ark  
 ARK  
 Article Title  
 Article Title(s)  
 Artist  
 Artist/Manufacturer  
 Artist's Notation  
 Athlete's Name  
 Audience

Audio Clip Instructions  
 Author  
 Author/Editor  
 Author/Editor Bio  
 Author(s)  
 AV Item Number  
 Awards  
 A-Z assignment  
 Access Restrictions  
 Banner  
 Batch  
 Batch #  
 Batch and Box #  
 Batch number  
 bibliographicCitation  
 Bibliographic Citation  
 Binding  
 Birth Date  
 Birthplace  
 Bit depth  
 Bit-depth  
 Bit Depth  
 Box and Folder Number  
 browse order  
 Building  
 Business/building names  
 B/W Illustrations  
 Calendar  
 Calendar Month  
 Call Number  
 Camp  
 CA+P Accession No.  
 CA+P Course  
 Caption  
 Caption on slide  
 Cataloged by  
 Cataloger  
 Catalog Number  
 Catalog Transcription  
 Catalogue  
 Catalogued By  
 Category

CDM 5 URL  
 Century  
 Children of Deceased  
 citation\_issn  
 Citation  
 citation\_author  
 citation\_conference\_title  
 citation\_date  
 citation\_dissertation\_institution  
 citation\_dissertation\_name  
 citation\_doi  
 citation\_firstpage  
 citation\_inbook\_title  
 citation\_isbn  
 citation\_issn  
 citation\_issue  
 citation\_journal\_title  
 citation\_keywords  
 citation\_language  
 Citation\_lastpage  
 citation\_other\_author  
 citation\_publisher  
 citation\_technical\_report\_institution  
 citation\_technical\_report\_numbe  
 citation\_title  
 citation\_volume  
 City  
 Class  
 Classes  
 Classification  
 Click here for an immersive experience!  
 Click here for a 'page-flip' view  
 Click here for details  
 Click here to go to AFRC 2012 Winter Conference  
 Materials  
 Click here to go to link  
 Click here to view photo set!  
 Click to Play  
 Clinical  
 Clinical Signs

Cloth Discount  
 Cloth ISBN  
 Cloth Price  
 Cohort Number  
 Collection  
 Collection aliases  
 Collection compiler  
 Collection Compiler  
 Collection contact information  
 Collection information  
 Collection Information  
 Collection, Is part of  
 Collection Name  
 Collection Name and Number  
 Collection number  
 Collection Number  
 Collection number and name  
 Collection Number and Name  
 Collection Owning Institution  
 Collection type  
 Color Illustrations  
 Comment  
 Comments  
 Commissioned By  
 Common Name  
 Company  
 Company/Business  
 Complete Catalog  
 Completion Date  
 Compressed file editing softwar  
 Compressed File Extent  
 Compressed file operating syst  
 Compressed file quality  
 Compressed file size  
 Compression  
 Conference  
 Contact information  
 CONTENTdm file name  
 CONTENTdm file name v.5  
 CONTENTdm file path v.5  
 CONTENTdm number  
 CONTENTdm number v.5

Context  
 Context URL  
 Contributing Institution  
 Continent  
 Contrib. Institution  
 Contributing Institution  
 Contributing.Institution  
 Contributing Institution  
 Contributing Organization  
 Contributor  
 Contributor Primary  
 Contributor Primary  
 Contributor Secondary  
 Conversion specifications  
 ConversionSpecifications  
 Conversion Specifications  
 Copy ID  
 Copyright  
 Copyright Date  
 Copyright notice  
 Corporate name  
 Corporate Name  
 Country  
 County  
 County Name  
 Coverage  
 Coverage spatial  
 Coverage-spatial  
 Coverage Spatial  
 Coverage-Spatial  
 Coverage - Temporal  
 Coverage-Temporal  
 Covrage-spatial  
 Creation Date  
 Creation methodology  
 Creation Methodology  
 Creator  
 Creator Dates  
 Creator Nationality  
 Creator of Media Recording  
 Creator(s)/Author(s)  
 Credit Line  
 Culture  
 Current address  
 Current Address  
 Curriculum  
 Data Description  
 Date  
 Date created  
 Date-Created  
 Date created v.5

Date Delivered  
 Date digital  
 Date.digital  
 Date Digital  
 Date-Digital  
 Date.Digital  
 Date.Earliest  
 Date Entered  
 Date issued  
 Date.Latest  
 Date modified  
 Date modified v.5  
 Date of photograph  
 Date of Photograph  
 Date.original  
 Date Original  
 Date-Original  
 Date.Original  
 Date Original DF  
 Date or Period  
 Date Scanned  
 DC.citation.epage  
 DC.citation.issue  
 DC.citation.spage  
 DC.citation.volume  
 DC.relation.ispartof  
 Death Date  
 Deathplace  
 Deceased Age  
 Deceased NameDegree  
 Degree Granting Institution  
 Department  
 Department email  
 Department name  
 Depth (3D Objects)  
 Description Anecdotal  
 Description  
 Description2  
 Dewey Decimal Number  
 Digital Archive  
 Digital Collection  
 Digital File Location  
 Digital Format  
 Digitization Specification  
 Digitization specifications  
 Digitization Specifications  
 Digitization Specs  
 Digitized by  
 Digitizing Device  
 Digitizing Technician  
 Dimensions

Full NatureServe Report  
 Full resolution  
 Full Resolution 2  
 Full resolution folder path  
 Full resolution v.5  
 Full test  
 Full text  
 Full Text  
 Full Text field  
 Full Text Field  
 Full Text Search  
 Full Transcript  
 Function  
 Funding  
 Further InFunding/Fellowship  
 formation  
 Game Website  
 Gender  
 Genre  
 Genre (AAT)  
 Genre or Form  
 Genus  
 Geographical names  
 Geographical Names  
 Geographic coordinates  
 Geographic Coordinates  
 Geographic Location  
 Geo-locale  
 Geopolitical place  
 Get Media  
 Global Positioning System Coordina  
 Grade Level  
 Habitat Type  
 Has Part  
 Height, Framed  
 Height Unframed  
 Hidden Description  
 Historic address  
 Historic Address  
 Historic place name  
 Historic Place Name  
 History  
 Holding institution  
 Holding Institution  
 Holding.Institution  
 Holding Location  
 Holding institution  
 ID  
 Identifier  
 Image Capture

## Variations in title field name

<b>Collection Alias</b>	<b>Local Field Name</b>	<b>CONTENTdm nickname</b>
col/ehsl-wht	Title	title
col/epubs	Article Title	articl
col/fhwa	Document Title	title
col/fitz	Title	title
col/flwright-jp2	Alternative Title	latin
col/flwright-jp2	Title	title
col/GBA_Photos	Caption	title
col/GBM	Caption	title
col/gis	Title	title

# Controlled Vocabularies

## Type (DCMI Type vocab)

- Collection
- Dataset
- Event
- Image
- Image/MovingImage
- Image/StillImage
- InteractiveResource
- PhysicalObject
- Service
- Software
- Sound
- Text

## Format (Internet Media Type)

- application/msword
- application/pdf
- application/vnd.google-earth.kmz
- application/vnd.ms-powerpoint
- application/xml
- application/x-shockwave-flash
- application/zip
- audio/mpeg
- image/jpeg
- image/png
- text/html
- video/mp4

## Language (ISO 639-2)

- ara
- chi
- eng
- fre
- ger
- heb
- ita
- kor
- lat
- nav
- rus
- spa

# Delete Unneeded Legacy Collections



Home Browse All Ask a Question

Search  within results [Search](#) [Advanced Search](#)

Please take a moment to fill out our survey

Browsing items in: Univ of Utah - Career Services Library

Sort by: [Title](#) [Display Options](#) [Save to favorites](#)

Display: [100](#)

Narrow your search by:

- ▶ Creator
- ▶ Date Digital
- ▶ Subject
- ▶ Type

<input type="checkbox"/> Thumbnail	<a href="#">Title</a>	<a href="#">Description</a>	<a href="#">Subject</a>	<a href="#">Collection</a>
	- 90 THINGS TO DO BEFORE U JOIN THE REAL WORLD	How do you get a job without experience and get experience without a job? It's the question virtually every college student or recent graduate faces. In Getting from College to Career, Lindsey Pollak offers the first definitive guide to building...	JOB SEARCH	Univ of Utah - Career Services Library
	10 THINGS EMPLOYERS WANT YOU TO LEARN IN COLLEGE	A college diploma can open many doors, but it's just an expensive piece of paper if you don't have what employers seek. Long-time professor and student adviser Bill Coplin surveyed the nation's top companies and identified the essential skills...	CAREER PLANNING	Univ of Utah - Career Services Library
	100 BEST CAREERS FOR WRITERS AND ARTISTS	What and where the jobs are for people with creative talents! For all the aspiring writers and artists who want to put their creative talents to good use, here's the guide that can show them how! Compiled by one of America's leading career...	WRITING	Univ of Utah - Career Services Library

## Future metadata corrections - next steps

Review 441 fields manually, noted issues (field used in only one collection, field synonyms)

Develop spreadsheet to cluster possible fixes by partner, write up metadata recommendations - ongoing process

To do:

- Explore possibilities of assessment scripts against solr to review metadata values
- Explore visualization tools



# Future metadata corrections - Rights statements

40867	<u><a href="#">This material may be protected by copyright. Permission required for use in any form. For further information please contact the Multimedia Archives, Special Collections, J. Willard Marriott Library, University of Utah.</a></u>
20207	<u><a href="http://creativecommons.org/licenses/by-nc-sa/1.0/">http://creativecommons.org/licenses/by-nc-sa/1.0/</a></u>
5134	<u><a href="#">Digital Image Copyright 2008, University of Utah. All Rights Reserved.</a></u>
4325	<u><a href="#">Digital Image Copyright 2010, University of Utah. All Rights Reserved.</a></u>
3597	<u><a href="#">Copyright owner unknown</a></u>
3583	<u><a href="#">Digital image copyright 2005, University of Utah. All rights reserved.</a></u>
3488	<u><a href="#">Digital image, copyright 2003 Uintah County Library</a></u>
3405	<u><a href="#">This material is protected by copyright. Express permission is required for use in any form. For further information please contact the Multimedia Archivist, Special Collections, J. Willard Marriott Library, University of Utah.</a></u>
3033	<u><a href="#">This material may be protected by copyright. Permission required for use in any form. Please contact the Multimedia Archives, Special Collections, J. Willard Marriott Library, University of Utah.</a></u>
2975	<u><a href="#">This material is protected by copyright. Permission required for use in any form. For further information please contact the Multimedia Archives, Special Collections, J. Willard Marriott Library, University of Utah.</a></u>
2616	<u><a href="#">Digital Image Copyright 2007, University of Utah. All Rights Reserved.</a></u>

# Future metadata corrections - MWDL Application Profile

title	date	description	subject	type	format	rights
National Ski Association officers and directors	January 13 - 15, 1961	Photo shows National Ski Association	Melville, Alton Crane, 1961	Image	image/jpeg	This material may be protected by copyright.
Group Picture in the Orderville Ward				Image	image/jpeg	This material may be protected by copyright.
Saltair Beach Pavilion and bathers, Saltair Resort.		1893-1925 (destroyed by fire)		Image	image/jpeg	This material may be protected by copyright.
Interior of the Dancing Pavilion, Saltair Resort.		1893-1925 (destroyed by fire)		Image	image/jpeg	This material may be protected by copyright.
Saltair Beach Pavilion, Saltair Resort.		1893-1925 (destroyed by fire)		Image	image/jpeg	This material may be protected by copyright.
Group of ski tourers high in the Wasatch mountains		A group photo of ski tourers in the	Wasatch Mountains (Utah)	Image	image/jpeg	This material may be protected by copyright.
Painting of soldiers.					image/jpeg	This material may be protected by copyright.
George Thompson in England.				Image	image/jpeg	This material may be protected by copyright.
George Thompson with second wife Katherine.				Image	image/jpeg	This material may be protected by copyright.
William Jennings Bryan at the radio station.					image/jpeg	
Nels Just with Emma Bennett Just				Image	image/jpeg	This material may be protected by copyright.
Emma Bennett Just's son Fred Bennett.				Image	image/jpeg	This material may be protected by copyright.
Charley and James Just				Image	image/jpeg	This material may be protected by copyright.
Nels Just				Image	image/jpeg	This material may be protected by copyright.
Thomas Moran				Image	image/jpeg	This material may be protected by copyright.
Just family by their ranch house				Image	image/jpeg	This material may be protected by copyright.
Francis and Will Just				Image	image/jpeg	This material may be protected by copyright.
Emma Bennett Just at age 16.				Image	image/jpeg	This material may be protected by copyright.



# Future metadata corrections - Subject faceting

## Topic

<u>American newspapers</u>	1
<u>Antisemitism</u>	12
<u>Auschwitz (Concentration camp)</u>	1
<u>Bamberger Railroad Company</u>	2
<u>Bar mitzvah</u>	1
<u>Cantors (Judaism)</u>	1
<u>Conservative Judaism</u>	1
<u>Eastern Front (World War (1914-1918))</u>	1
<u>Emigration and immigration</u>	3
<u>Ethnic groups</u>	1
<u>Evacuation and relocation of Japanese Americans (United States : 1942-1945)</u>	1

# Roadmap for future development

## Solphal (Front end) development

- International Image Interoperability Framework (IIIF)
- RDFa and schema.org support

## SIMP Tool enhancements

- Workflow automation
- Large updates monthly, small updates as needed

## Large projects

- IMLS Grants
  - Newspapers in Hydra (LG-70-17-0043-17)
  - Western Name Authority File (LG-72-16-0002-16)
- Born digital content
- Partner projects
- Data Repository





People

Platforms

Processes

# Lessons Learned

## Wild wild west of metadata wrangling



Shooting of Patrick Coughlin

<https://collections.lib.utah.edu/details?id=962118>

Don't keep shoveling  
while it is still snowing!



Unidentified man holding snow shovel in front of his house, (family, friend or acquaintance of Norman D. Nevills).

<https://collections.lib.utah.edu/details?id=984034>

# Thank You!

Contact us at:

Jeremy Myntti ([jeremy.myntti@utah.edu](mailto:jeremy.myntti@utah.edu))

Anna Neatrour ([anna.neatrour@utah.edu](mailto:anna.neatrour@utah.edu))

Harish Maringanti ([harish.maringanti@utah.edu](mailto:harish.maringanti@utah.edu))

Brian McBride ([brian.mcbride@utah.edu](mailto:brian.mcbride@utah.edu))

Alan Witkowski ([alan.witkowski@utah.edu](mailto:alan.witkowski@utah.edu))

Matt Brunsvik ([matt.brunsvik@utah.edu](mailto:matt.brunsvik@utah.edu))

The screenshot shows the homepage of the J. Willard Marriott Digital Library. At the top, there is a search bar with the text "Search" and a magnifying glass icon. Below the search bar, the title "J. Willard Marriott Digital Library" is displayed in a large, white font against a dark background with a geometric pattern. Underneath the title, a paragraph states: "The J. Willard Marriott Digital Library at the University of Utah hosts more than 250 digital collections, containing over 1 million digital photographs, maps, books, videos, audio recordings, and other items." Below this text, there are six white boxes arranged in a 2x3 grid, each containing the name of a collection and its description:

- Delta City Library**: Beckwith Photograph Collection
- Shelley City**: History of Shelley Idaho
- University of Utah Partnerships**: Art and Art History Department PR Photographs American Flame Research Committee Alta Avalanche
- Department of Heritage and Arts**
- Snow College**: Saga of The Sanpitch

<https://collections.lib.utah.edu/>

<https://newspapers.lib.utah.edu/>