

Utah State University

DigitalCommons@USU

Library Faculty & Staff Presentations

Libraries

11-2016

Recommending community best practices for metadata in the Digital Preservation Network: a progress report by the DPN Preservation Metadata Standards Working Group.

M Caruso

D Krewer

J Mullins

Liz Woolcott
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/lib_present



Part of the [Library and Information Science Commons](#)

Recommended Citation

Caruso, M., Krewer, D., Mullins, J., and Woolcott, L. (2016). Recommending community best practices for metadata in the Digital Preservation Network: a progress report by the DPN Preservation Metadata Standards Working Group. National Digital Stewardship Alliance Annual Conference, Milwaukee, Wisconsin, November, 2016.

This Presentation is brought to you for free and open access by the Libraries at DigitalCommons@USU. It has been accepted for inclusion in Library Faculty & Staff Presentations by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.



Recommending community best practices for metadata in the Digital Preservation Network

A progress report by the DPN Preservation Metadata Standards Working Group: Moriah Caruso, Liz Woolcott, Drew Krewer, Jennifer Mullins

Overview

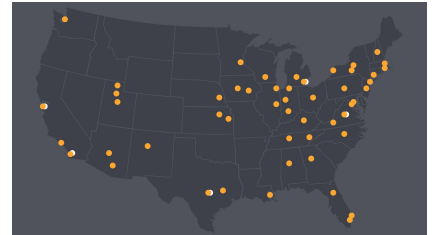
1. About DPN
2. Preservation Metadata Working Group - Who, What, Why, When, How
3. Current and Future Directions
4. Work Accomplished So Far
5. Challenges Encountered
6. Invitation for Feedback




About DPN

About DPN

- “Dark archive”
- “Redundant and varied technical and legal infrastructure to assure the survival, ownership and management of preserved digital content”
- 50+ charter members, including public and private universities and colleges, consortiums, and digital curation service providers
- Collaborative community
- Members are just starting to deposit content





Preservation Metadata Standards Working Group: Who, What, Why, When, How?

Preservation Metadata Standards Working Group

- Was created to outline the metadata standards that DPN will follow for preservation.
 - Deliverable 1: Define a set of fields that will remain constant and available with all deposits made into DPN.
 - Deliverable 2: Publish a paper documenting process and relevant work done at other institutions leading to the selection of these standards.
- Started its work in January 2016.



Preservation Metadata Working Group

- Members:
 - Moriah Caruso - University of Washington
 - Drew Krewer - University of Houston
 - Jenny Mullins - Dartmouth College
 - Simon O'Riordan - Emory University
 - Liz Woolcott - Utah State University
 - Dave Pcolar - DPN Staff Liaison




The background is a solid pink color. In the top right corner, there are several overlapping geometric shapes: a dark pink square, a medium pink square, and a light pink square, all partially cut off by the edge of the frame.

Work Accomplished
So Far

Use Case Development

To fulfill the goal of having geographically distributed copies of preservation master files, the Library decides to deposit copies of materials in an off-site dark storage environment, such as the Digital Preservation Network. Once deposited, materials cannot be changed or removed. **Files need to be packaged so that, when retrieved- whether in one, twenty or fifty years time- they can be understood, verified and used.**

Upon retrieval:

- The content and context of files must be discernable.
 - There must be proof that the files are identical to the ones initially deposited.
 - It must be possible to render the file's content (through current software or emulation), or relate the file to a migrated version along with documentation of the migration event
- 

Title
 Creator
 Date
 Description
 Publisher
 Subject
 Rights
 Access Rights
 Identifier
 Institutional Identifier
 Institution
 Content Type
 Formats
 Format-Extent
 Associated Items
 Filenames
 Checksum
 Relation-IsPartOf
 Significant Properties
 Relationship
 Event
 Environment

		UH	Penn State	Emory	USU	UW	Dartmouth
Descriptive	Title	dc:title	Title	Title	Title	dc.title	
Descriptive	Creator	dc:creator			Creator	dc.creator	
Descriptive	Date	dc:date		Date-Source	Date-Source; Date-Digitized; Date-Ba	dc.date; dc.date.accesioned	
Descriptive	Description	dc:description	Description		Description	dc.description	External-Descriptio
Descriptive	Publisher	dc:publisher					
Descriptive	Subject					dc.subject	
Administrative	Rights	dc:rights	Rights	Rights/Access statement	Rights	dc:rights	
Administrative	Access Rights	dcterms:accessRights	Access Rights	Rights/Access statement	Access Rights		
Administrative	Identifier		Identifier	Identifier	Identifier-bagLevel	dc.identifier.uri	External-Identifier
Administrative	Institutional Identifier		Institutional Identifier		Institutional Identifier		Source-Organizatio
Descriptive/Administrative	Institution		Institution	Institution	Institution		
Descriptive/Technical	Content Type			Content Type	Type		
Descriptive/Technical	Formats				Format	mimetype/format	
Descriptive/Technical	Format-Extent				Format-Extent		
Descriptive/Administrative	Associated Items				Identifier-associatedIdentifiers		
Technical	Filenames				Filenames		
Technical	Checksum				Checksum		
Descriptive/Administrative	Relation-IsPartOf				Relation-IsPartOf		
Administrative	Bagged by				Bagged by		
Administrative	Uploaded by				Uploaded by		
bag-info.txt	Size					Size	
bag-info.txt	Organization Address						Organization-Addre

DPN Bag-It Specification Amendment

- Member institution was creating metadata being stored as tag files.
- BagIt File Packaging Format Spec (V0.97) allows for “other tag files”:
 - 2.2.4. Other Tag Files: A bag MAY contain other tag files that are not defined by this specification. **Implementations SHOULD ignore the content of any unexpected tag files, except when they are listed in a tag manifest.** When unexpected tag files are listed in a tag manifest, implementations MUST only treat the content of those tag files as octet streams for the purpose of checksum verification.
- Decision that optional, member-created tag files must be listed in the tag manifest to ensure checksum is recorded.





Current & Future Directions

Current Directions

University of Washington:

*A tale of
"lost files and mixed blessings"*





University of Washington Special Collections Prosch Indian Albums. PH Coll 18 NA 1370

Metadata:

Preservation gathers:

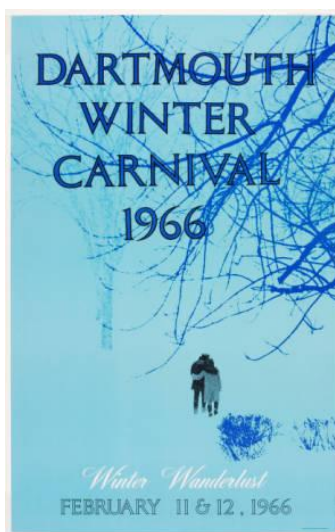
- Collection-level marc record
- EAD finding aid
- Item level metadata from access systems
- Rights data, donor agreements, supplementary info

Preservation creates: Archivematica-produced METS file with technical metadata

Current Directions

Dartmouth College

Working with what we've got




```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <mods xmlns:mods="http://www.loc.gov/mods/v3"
3   xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
4   xsi:schemaLocation="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/version="3.6">
5   <mods:titleInfo>
6     <mods:title>Seuss on the loose 2016</mods:title>
7     <mods:subTitle>Dartmouth Winter Carnival</mods:subTitle>
8   </mods:titleInfo>
9   <mods:titleInfo type="uniform">
10     <mods:title>Dartmouth Winter Carnival Posters</mods:title>
11   </mods:titleInfo>
12   <mods:name type="personal" usage="primary">
13     <mods:namePart>Miller, Anna C.</mods:namePart>
14     <mods:role>
15       <mods:roleTerm type="text" authority="marcrelator">artist</mods:roleTerm>
16     </mods:role>
17   </mods:name>
18   <mods:name type="corporate">
19     <mods:namePart>Dartmouth College</mods:namePart>
20     <mods:namePart>Winter Carnival</mods:namePart>
21   </mods:name>
22   <mods:typeOfResource>still image</mods:typeOfResource>
23   <mods:genre type="content" authority="rdacontent">still image</mods:genre>
24   <mods:genre authority="marcgt">picture</mods:genre>
25   <mods:genre authority="fast" valueURI="http://id.worldcat.org/fast/142384">
26     <mods:originInfo eventType="publication">
27       <mods:dateIssued encoding="w3cdtf" keyDate="yes">2016</mods:dateIssued>
28     </mods:originInfo>
29   <mods:language>
30     <mods:languageTerm type="text">English</mods:languageTerm>
31     <mods:languageTerm type="code" authority="iso639-2b">eng</mods:languageTerm>
32   </mods:language>
33   <mods:physicalDescription>
34     <mods:form type="media" authority="rdamedia">computer</mods:form>

```



```

<?xml version="1.0" encoding="UTF-8"?>
<fits xmlns="http://hul.harvard.edu/ois/xml/ns/fits/fits_output"
  <identification status="CONFLICT">
    <identity format="TIFF EXIF" mimeType="image/tiff" toolname="Jhove"
      <tool toolname="Jhove" toolversion="1.5" />
    </identity>
    <identity format="Tagged Image File Format" mimeType="image/tiff"
      <tool toolname="file utility" toolversion="5.18" />
      <tool toolname="Exiftool" toolversion="9.13" />
      <tool toolname="Droid" toolversion="6.1.3" />
      <tool toolname="ffident" toolversion="0.2" />
      <tool toolname="Tika" toolversion="1.3" />
      <externalIdentifier toolname="Droid" toolversion="6.1.3" />
    </identity>
  </identification>
  <fileinfo>
    <size toolname="Jhove" toolversion="1.5">123037932</size>
    <creatingApplicationName toolname="Jhove" toolversion="1.5">
      <lastmodified toolname="Exiftool" toolversion="9.13" status="CONFIDENTIAL">
      <lastmodified toolname="Tika" toolversion="1.3" status="CONFIDENTIAL">
      <created toolname="Exiftool" toolversion="9.13" status="CONFIDENTIAL">
      <filepath toolname="OIS File Information" toolversion="0.1">
      <filename toolname="OIS File Information" toolversion="0.1">
      <md5checksum toolname="OIS File Information" toolversion="0.1">
      <fslastmodified toolname="OIS File Information" toolversion="0.1">
    </fileinfo>
    <filestatus>
      <well-formed toolname="Jhove" toolversion="1.5" status="SINGLE">
      <valid toolname="Jhove" toolversion="1.5" status="SINGLE">
    </filestatus>
    <metadata>
      <image>
        <byteOrder toolname="Jhove" toolversion="1.5" status="SINGLE">
        <compressionScheme toolname="Jhove" toolversion="1.5" status="SINGLE">

```

Future Directions

- **Case Study**
 - Finalize core record
 - Evaluate if there are additional ways we can help the membership



Future Directions

- **Focus group**
 - Representative sample of institutions
 - Analyze metadata needs, uses, and workflows
 - Gather feedback



Future Directions

- **Final Deliverables :**
 - Recommendations/best practices for DPN metadata record
 - Toolkit for partners to use
 - Paper outlining the outcomes of the project



Feedback?

- jennifer.l.mullins@dartmouth.edu
- dave@dpn.org
- neilsm@uw.edu
- ajkrewer@uh.edu
- liz.woolcott@usu.edu