OVERVIEW

- ANALYSIS OF DPLA RIGHTS STATEMENTS
- DEVELOPMENT AND COLLABORATION
- DELIVERABLES AND IMPLEMENTATION
- REFLECTIONS

https://catalog.archives.gov/id/549322 — No Copyright — United States
ANALYSIS OF DPLA RIGHTS STATEMENTS

https://catalog.archives.gov/id/553798 — No Copyright - United States
**PRIOR WORK: DPLA RIGHTS ANALYSIS (2014)**

- Based on October 2014 dataset (8.1 million records)
- Aggregate reports exported containing rights statement and counts, Hub, and contributing institution
- Reports transformed into CSV and imported into OpenRefine for further processing
  - Associated institutions removed
  - Normalization to reduce and aggregate near-duplicates (whitespace removal; multi-pass clustering process)
DPLA RIGHTS ANALYSIS (2014): FINDINGS

- Incredible diversity and absence of rights statements in 8.1 million DPLA records
- 87,610 “unique” rights statements after normalization
- ~1.01M (~12.5%) records missing rights statements
- ~2.4M (~29.9%) “in copyright/©/all rights reserved”
- ~1.6% under a Creative Commons license
- Rights statements matching multiple “categories” that could be confusing to end users
Copyright [Redacted]. In the public domain use of this object governed by a CC-BY-NC-ND license.
Audio file copyright [Redacted]; Creative Commons Attribution-NonCommercial-NoDerivs 2.5 License; http://creativecommons.org/licenses/by-nc-nd/2.5/; Public Domain; Photo Courtesy of [Redacted]; Image retrieved from [Redacted]
DEVELOPMENT AND COLLABORATION
TECHNICAL WORKING GROUP

- Mark A. Matienzo (Stanford; co-chair)
- Antoine Isaac (Europeana; co-chair)
- Valentine Charles (Europeana)
- Esmé Cowles (Princeton)
- Karen Estlund (Penn State)
- Tom Johnson (University of Washington)
- Patrick Peiffer (Bibliothèque nationale de Luxembourg)
- Richard J. Urban (Florida State University)
- Maarten Zeinstra (Kennisland)
- Implementation contractors
  - Felix Ostrowski (graphthinking GmbH)
  - Adrian Pohl
CONSENSUS AND CONVERGENCE WITH COMMUNITY

- Not everyone in final roster for technical working group was a member from the start.
- Discovery of related work within Hydra Project open source framework community (http://projecthydra.org/) provided opportunity for working group to expand.
LOGISTICS AND NORMS OF COLLABORATION

- Regular standing meetings and asynchronous communication
- Transparency with larger working group and outside community even if there is minimal engagement
- Ensure adequate feedback periods with clear deadlines
- Open RFP process for technical development
- Occasional need to fall back on “formal channels” (e.g. technical memoranda; see http://bit.ly/rs-http-behavior)
DELIVERABLES & IMPLEMENTATION
DELIVERABLES

- White papers
- RDF-based vocabulary and data model for expressing rights statements, published as linked data
- Web application to serve vocabulary/rights statements
- Associated CMS for content and informational resources (e.g. implementation guidelines)
WHITE PAPERS

- Technical working group produced a series of white papers complementary to the larger working group


- *Requirements for the Technical Infrastructure for Standardized International Rights Statements* (October 2015 and March 2016): finalized modeling concerns, HTTP behavior, versioning, interpretation of linked data publication norms and patterns

DESIGN PRINCIPLES FOR RIGHT STATEMENTS

- Simple
- Flexible
- Descriptive
- Accurate
- Transparent
TECHNICAL PRINCIPLES FOR IMPLEMENTATION

- Reflect the rights statements accurately with minimal ontological commitments
- Publish according to best practices for linked data
- Design for extensibility
- Focus on ease of implementation
TECHNICAL IMPLEMENTATION: DATA MODEL

- skos:ConceptScheme
- skos:Collection
  - skos:inScheme
  - skos:member
  - dcterms:rights (or subproperty)
    - rdf:Resource (e.g. edm:ProvidedCHO, dpla:SourceResource)
- skos:Concept
  - dcterms:RightsStatement
## TECHNICAL IMPLEMENTATION: VOCABULARY REUSE & ALIGNMENT

<table>
<thead>
<tr>
<th>Source</th>
<th>Prefix abbreviation</th>
<th>Namespace</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creative Commons Rights</td>
<td>cc:</td>
<td><a href="http://creativecommons.org/ns#">http://creativecommons.org/ns#</a></td>
</tr>
<tr>
<td>Expression Language (ccREL)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dublin Core Elements 1.1</td>
<td>dc:</td>
<td><a href="http://purl.org/dc/elements/1.1/">http://purl.org/dc/elements/1.1/</a></td>
</tr>
<tr>
<td>DCMI Type Vocabulary</td>
<td>dcmitype:</td>
<td><a href="http://purl.org/dc/dcmitype/">http://purl.org/dc/dcmitype/</a></td>
</tr>
<tr>
<td>DCMI Metadata Terms</td>
<td>dcterms:</td>
<td><a href="http://purl.org/dc/terms/">http://purl.org/dc/terms/</a></td>
</tr>
<tr>
<td>Europeana Data Model</td>
<td>edm:</td>
<td><a href="http://www.europeana.eu/schemas/edm/">http://www.europeana.eu/schemas/edm/</a></td>
</tr>
<tr>
<td>ODRL</td>
<td>odrl:</td>
<td><a href="http://www.w3c.org/ns/odrl/2/">http://www.w3c.org/ns/odrl/2/</a></td>
</tr>
<tr>
<td>PREMIS Copyright Status</td>
<td>premiscopy:</td>
<td><a href="http://id.loc.gov/vocabulary/preservation/">http://id.loc.gov/vocabulary/preservation/</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td>copyrightStatus/</td>
</tr>
<tr>
<td>SKOS</td>
<td>skos:</td>
<td><a href="http://www.w3.org/2004/02/skos/core#">http://www.w3.org/2004/02/skos/core#</a></td>
</tr>
<tr>
<td>OWL</td>
<td>owl:</td>
<td><a href="http://www.w3.org/2002/07/owl#">http://www.w3.org/2002/07/owl#</a></td>
</tr>
<tr>
<td>ODRS</td>
<td>odrs:</td>
<td><a href="http://schema.theodi.org/odrs#">http://schema.theodi.org/odrs#</a></td>
</tr>
</tbody>
</table>
## TECHNICAL IMPLEMENTATION: MODEL DEFINITION

<table>
<thead>
<tr>
<th>Short name of RS</th>
<th>Name of Rights Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>dc:identifier</td>
<td>skos:prefLabel</td>
</tr>
</tbody>
</table>

**URI for Rights Statement**

**One sentence description of the Rights Statement.** This will not be displayed as part of the Rights Statement. Intended for use in documents or on websites describing the Rights Statements.

**Text of the Rights Statement**

**Notices:** One or more notices related to the Rights Statement

**Disclaimer regarding this being a Rights Statement and not a legally operative License summary.**

**Generic selection criteria for the Rights Statement.** Short text that describes when this Rights Statement should be used, aimed primarily at data providers. This text will not be displayed as parts of the Rights Statement.

**Extra metadata**

For some statements it is possible to provide additional metadata which triggers the display of optional information at the text of the Rights Statement (and above the notices). If this is the case this will be noted here. Specific behavior is specified by keywords in bold as described by RFC2119.\(^{31}\)
ADDITIONAL METADATA AND HTTP BEHAVIOR

- Requirement for certain statements to allow for extra properties to be presented in human-readable representations
- Led to development of specifications for additional HTTP behaviors that followed spirit of *Best Practice Recipes for Publishing RDF Vocabularies*
- However, requirement and solution led to development of custom application rather than using existing vocabulary publishing applications
- Additional guidelines under development by DPLA and Europeana to address other additional metadata needs
TECHNICAL IMPLEMENTATION: WEB APPLICATION & CMS

- Custom web application for publishing rights statements, built in using Play framework (Java); built to follow Linked Data best practices and to address custom HTTP behavior requirements

- Static content management system built with Jekyll; support for translation of content/navigation

- Licensing: applications/code released under European Union Public License, version 1.1; content published as CC0

- All available online at https://github.com/rightsstatements
REFLECTIONS

- Transparency during active development created expectations for ongoing community process
- Desire for simplicity allowed us to make reasonable progress
- Ongoing concern we underspecified some longer-term needs (revision workflow, versioning); intertwined with governance concerns
- Need to identify tools needed to ease adoption
THANK YOU!

MARK A. MATIENZO
STANFORD UNIVERSITY LIBRARIES
@ANARCHIVIST
HTTP://MATIENZO.ORG/

All images used in this presentation, unless otherwise specified, are in the public domain in the United States, and are provided by courtesy of the US National Archives.