

# Metadata in Shared Digital Libraries

## *Metadata Interest Group Program*

CORE VIRTUAL INTEREST GROUP WEEK

MONDAY, MARCH 4, 2024

8AM PST / 9AM MST / 10 AM CST / 11AM EST

# Metadata Interest Group Leaders

- **Catherine Smith**, Chair
- **Joan Wang**, Vice-Chair
- **Elisa Naquin**, Program Co-Chair
- **Kate James**, Program Co-Chair
- **Melissa Zilic**, Communications Co-Chair
- **Katherine Manifold**, Communications Co-Chair

# Reminders

- This session is being recorded
- Enter questions in chat to “Everyone” at any time during the presentation
- Core Statement of Conduct:  
<http://www.ala.org/core/about/statement-of-conduct>

# Presentations

- Describing Our Members, Ourselves: Maintaining a Shared Controlled Vocabulary in Explore Chicago Collections
  - **Presenters:** Jessica BrodeFrank (University of Illinois at Chicago), Kate Flynn (University of Illinois at Chicago) Erin Matson (Chicago History Museum), and Gretchen Neidhardt, (Northwestern University)
- Creating and Sharing Date Metadata
  - **Presenter:** Annamarie C. Klose (The Ohio State University Libraries)
- Developing a Metadata Application Profile Supporting Interoperability, Research, and Data Harvesting for a Digital Library of Music Sources
  - **Presenter:** Jesse P. Karlsberg (Emory Center for Digital Scholarship and Emory University Department of Music)

# Describing Our Members, Ourselves: Maintaining a Shared Controlled Vocabulary in Explore Chicago Collections

Jessica BrodeFrank, Kate Flynn, Erin Matson & Gretchen Neidhardt

Chicago Collections Consortium Controlled Vocabulary Committee Members



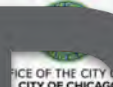
DEPAUL



LOYOLA  
UNIVERSITY CHICAGO



THE NEW

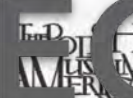


Northwestern

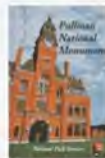


Northwestern

Northwestern



JASC Legacy Center





# Why create CCC?

- The resources that Chicago researchers need are distributed in archives across the city
- Researchers may not know which archives or libraries to consult
- They need to be able to search their topic and learn which archives hold collections on that topic

# What made this difficult?

Chicago's major libraries already had the resources to create and publish finding aids and digital collections independently

***BUT they all did it differently.***



# How do you express a subject?

Balloon framing

Burlesque (Theater) –

Burlesque (Theater) –

Hack writers

Murgatroyd, Everett F., 1880 or 1-1946.

one way streets

Painting.

Paintings

Radio -- Illinois -- Chicago Metropolitan Area -- Receivers and reception -- 1980-1989. -- lctgm tree

Trees -- Chicago Metropolitan Area -- 1900-1919. – lctgm

Weather experiments

Xylophone -- Instruction and study

## FEATURED EVENTS & MOVEMENTS



1933 World's Fair



Settlement House Movement



Demonstrations



Disasters



Fairs & Conventions

## ALL EVENTS & MOVEMENTS

1893 World's Fair (681)

1919 race riot (1)

1933 World's Fair (1415)

1968 Democratic Convention (18)

Before Chicago (until 1837) (102)

Celebrations (1826)

Civil War (117)

Demonstrations (485)

Disasters (299)

Early Chicago (1837-1871) (325)

Eastland Disaster (121)

Fairs & Conventions (1819)

Great Chicago Fire (319)

Haymarket Affair (30)

Iroquois Theater Fire (92)

Prohibition (37)

Pullman Strike (3)

Riots (43)

Settlement House Movement (836)

Temperance (39)

United States Civil War (1)

Wars (300)

World War I (189)

World War II (190)

# Controlled vocabularies overview

Chicago Collections Tags	Sources
Topics: Events, Government, Daily Life, Creativity, Environments, Work	LCSH, FAST, local terms
Topics: Communities	LCDGT, Census, National Center on Disability and Journalism style guide, Chicago History Museum research on official tribal language, Homosaurus, local terms
Neighborhoods	LCNAF, Geonames, local sources/terms
Cities	LCNAF, Geonames, Census, local sources/terms
Names	Institutional records (100/600/700 fields, creator fields)

# Example of controlled vocabularies in Explore

American Terra Cotta Company photograph collection, ca. 1890-1915.

See Also: [Fine Arts](#) | [Architecture](#) | [Industries](#) | [Design](#) | [Chicago](#) | [Elmslie, George G](#) | [View All >](#)



## NAMES:

- American Terra Cotta Company (Chicago, Ill.) Photographs.
- Elmslie, George G.
- Sullivan, Louis H., 1856-1924.

## SUBJECTS:

- Architects United States 1890-1919. Ictgm
- Buildings United States 1890-1919. Ictgm
- Decoration and ornament, Architectural 1890-1919. Ictgm
- Decoration and ornament, Architectural Illinois Chicago 1890-1919. Ictgm
- Designers Illinois Chicago 1890-1919. Ictgm
- Terra-cotta Illinois Chicago 1890-1919. Ictgm
- Terra-cotta industry United States 1890-1919. Ictgm
- Terra-cotta sculpture 1890-1919. Ictgm

GEOGRAPHIC COVERAGE: Chicago (Ill.) Buildings, structures, etc. 1890-1919.

# CCC Committees

## ADVISORY COMMITTEES

Communications Committee

Controlled Vocabulary Committee

Cooperative Reference Committee

Development Committee

Digital Resources and Collections Committee

Membership Committee

Programming Committee

Systems Development and Analysis Committee

User Support and Assessment Committee

### Communications Committee

The group will help make certain that individuals and groups in the Chicago Collections community feel represented and informed about consortium matters relevant to their roles, needs, and interests.



### Digital Resources & Collections Committee

This committee's charge is to guide and develop online resources that facilitate learning through materials found in Chicago Collections Consortium (CCC) member collections. Additionally, the DRC also serves to identify subject gaps inclusive of underrepresented historical records in the Explore Chicago Collections (ECC) portal and serves as a resource for outside agencies seeking guidance on archival preservation best practices and/or identifying appropriate repositories for their collections.



### Cooperative Reference Committee

The Cooperative Reference Committee works to advance access, knowledge, and use of collections held at Chicago Collections' member institutions. Participants include generalist and subject-specific reference librarians and archivists from member institutions.



### Membership Committee

Building on the work and recommendations of the Organization & Governance Work Group, the Membership Committee identifies potential members for Chicago Collections and recommends policies to the Board of Directors relating to membership.



### Programming Committee

The Programming Committee is responsible for long-range planning and general oversight of Chicago Collections' public programming and services. It identifies, develops, recommends, and supports the implementation of activities such as virtual and physical exhibitions, seminars, workshops, and the development of other educational programs and resources that build upon Explore Chicago Collections, highlight member collections, and further facilitate the discovery of Chicago's history and culture.



### Controlled Vocabulary Committee

The Controlled Vocabulary Committee works to integrate perspectives from a number of different areas including archives, cataloging, public access and services, and user experience to support the creation and maintenance of these controlled vocabularies.



### Development Committee

The Development Committee develops strategies and tactics to generate the required funding to sustain Chicago Collections. It closely follows the work of the Board of Directors and other committees to find activities that present opportunities for funding, identifies, researches, and evaluates potential funding agencies and other revenue streams, and assists in drafting grant proposals and other fundraising materials.



### Systems Development and Analysis Committee

The Systems Development and Analysis Committee assists with the maintenance of Explore Chicago Collections and Metadata Hopper as well as supporting the workflow of content into Explore Chicago Collections.



### User Support and Assessment Committee

The User Support and Assessment Committee acts as a liaison between the Systems Development and Analysis Committee and users of systems and tools developed by Chicago Collections Consortium (CCC). The committee is also responsible for collecting feedback on CCC-developed systems and tools. This feedback may come from users and other CCC committees and is intended to anticipate evolving needs and what new features are needed.



# Who we are

- The Controlled Vocabulary Committee's charge includes:
  - Keeping up to date with changes in how things are described, including updates in other controlled vocabularies
  - Reviewing content and search trends in EXPLORE Chicago Collections to determine if new tags are needed or if current tags need to be changed
  - Soliciting and reviewing proposals for new tags
  - Maintaining documentation on the controlled vocabularies, including information on how to apply tags and how they map to other controlled vocabularies
  - Providing training and support on the controlled vocabularies
- Committee started as Task Force that worked periodically on issues relating to the controlled vocabularies but became a full-time advisory committee in 2023



# Our current members

- Jessica BrodeFrank, Chicago Collections (co-chair)
- Jeremy Bucher, National Hellenic Museum
- Brigid Crawley, Chicago History Museum
- Stephanie Fletcher, Illinois Institute of Technology
- Kate Flynn, University of Illinois Chicago
- Mingyan Li, University of Illinois Chicago
- Greer Martin, Loyola University Chicago
- Erin Matson, Chicago History Museum (co-chair)
- Gretchen Neidhardt, Northwestern University, Feinberg School of Medicine



# 2023 Goals

- Maintenance
  - **Goal 1: Review new tag submissions**
    - What does the form to submit feedback or new terms look like?
    - How often do we review tag submissions?
    - What are the systems that need to be updated for new tags?
    - How do we notify the community about new tags?
  - **Goal 2: Review internal tag data**
    - What information do we have to find about what people are looking for and what tags they're using?
    - Uncover gaps and tags that aren't being used
    - Google analytics and internal tag assignment data

# 2023 Goals

- Maintenance
  - **Goal 3: Review pending Communities questions**
  - **Goal 4: Catch up with base vocab changes**
    - Some terms were based on other controlled vocabularies, have these terms changed?
    - Can we build a automated workflow for tracking changes?

# 2023 Goals

- Training
  - **Goal 5: Develop training and guidelines**
    - What do we need in our data dictionary specifically?
    - Builds on work done by User Support & Assessment Committee (conducts current Metadata Hopper training)
  - **Goal 6: Connect users with the data dictionary and guidelines**
    - Goal was completed early on in year, added links in Metadata Hopper (site used for upload to Explore) and links in training documentation

# 2023 Goals

- Outreach
  - **Goal 7: Buddy system administration**
    - How can we reach out to people who are applying tags to see their questions?
    - Possible collaboration with User Support & Assessment Committee
  - **Goal 8: Public communication about tags**
    - What information do we convey to the internal community at large (not just people who are applying tags)?
    - What information do we convey to the external community?
    - What is the workflow for doing this? Can we use a template?
    - Collaboration with Communications Committee

# 2023 Goals

- Outreach
  - **Goal 9: Decide openness of Data Dictionary**
    - We had an internal Data Dictionary but do we want to create one that is openly available?
    - Who would be our target audience if we do this?
    - What changes would we need to make to a data dictionary targeted for use outside of the CCC community?
    - What other information would we need to convey?

# Guide for new term proposals:

<https://guides.chicagocollections.org/publictaglist>

- Public facing
- Includes FAQ for a general audience
- Goals:
  - Transparency in our process
  - Information sharing

The screenshot shows the 'Public Guide to CCC Controlled Vocabularies' webpage. At the top, there is a navigation bar with the 'EXPLORE CHICAGO COLLECTIONS' logo on the left and links for 'SUBJECT GUIDES' and 'OUR ORGANIZATION' on the right. Below the logo, there are links for 'HOME', 'BROWSE', 'LIBRARIES & ARCHIVES', 'HELP', and 'ASK US'. The main heading is 'Public Guide to CCC Controlled Vocabularies'. Below this, there is a search bar and a 'Print Page' button. A sidebar on the left contains links for 'Home', 'Current tag list', 'Feedback and tag submissions', and 'FAQs'. The main content area has a 'Send us feedback!' section with a 'Feedback Form' link and a 'Controlled Vocabularies Reference List' link. Below this is a 'Proposals to national vocabularies' section with instructions for new tag submissions and a list of resources for getting terms added or changed in national vocabularies. The resources include links to the Cataloging Lab, the Library of Congress Subject Headings (LCSH) list, the Library of Congress Name Authority File (LCNAF), the Geonames database, and the FAST funnel form.

EXPLORE  
CHICAGO COLLECTIONS

SUBJECT GUIDES OUR ORGANIZATION

HOME BROWSE LIBRARIES & ARCHIVES HELP ASK US

Public Guide to CCC Controlled Vocabularies

Last Updated: Dec 5, 2023 10:41 AM URL: <https://guides.chicagocollections.org/publictaglist> Print Page Log in to LibApps

Home

Current tag list

Feedback and tag submissions

FAQs

Send us feedback!

Feedback Form

Use the above link to send us feedback or propose a new tag. Please feel free to use the following controlled vocabularies as a reference when proposing a tag:

Controlled Vocabularies Reference List

Proposals to national vocabularies

For new tag submissions, use the Feedback Form above and note the following:

Is the term you would like to use not available in an existing controlled vocabulary? Here are some resources to getting a term added or changed in national vocabularies that Chicago Collections sources terms from:

- <https://cataloginglab.org/> - A website that offers crowd-sourcing help for editing or adding terms to the Library of Congress Subject Headings (LCSH) list, or Library of Congress Name Authority File (LCNAF).
- <https://www.loc.gov/aba/pcc/saco/funnel/saco.html> - Contact information for themed groups working on subject heading edits. If one exists for your term(s), they can be a great source of help in submitting proposals.
- <https://www.geonames.org/manual.html#create> - Edit or add places in the Geonames database.
- <https://www.loc.gov/aba/pcc/saco/FASTfun.html> - Propose a term via the FAST funnel form at the link.

Questions and comments about Library of Congress Demographic Group Terms (LCDGT) may be sent to: Policy, Training, and


# The feedback process

- Form is open ended and public facing
- Committee process
  - Meet monthly to discuss
  - Often decisions are made at that meeting or we decide if more research is needed
- Steps after approval
  - Set up workflow to ensure changes are made, the submitter is informed, and the wider community is made aware

## Feedback Form - CCC Controlled Vocabularies

Please use this form to:

- Propose a new term to Chicago Collections Consortium's Controlled Vocabulary Committee
- Provide feedback on any of our existing terms
- Ask a question about how to apply a term, the process for selecting terms for inclusion, or anything else relating to the controlled vocabularies in Explore Chicago Collections.

matson@chicagohistory.org [Switch account](#) 

\* Indicates required question

Email \*

Your email

Question or comment for the Controlled Vocabulary Committee \*

Your answer

**New Term Submission**

If you would like to suggest a new term, please answer the questions below.



# Communication about vocabularies

## **With membership**

- Collaborated with the Communications Committee on outline for current workflow for communication
- Announcements are stacked as appropriate for the next email that goes out to the community

## **With the public**

- 2024 goals include more outreach outside the organization, including with our users and our professional community

# Maintenance of a dynamic vocabulary

- Changes to existing terms
  - Manual updates to mapped terms (for now)
  - Working on OpenRefine workflows, but LCSH reconciliation is a moving target
  - Next steps in testing are to use Wikidata as a proxy
- Evaluating existing terms
  - Review search queries from Google Analytics - helps us evaluate whether to add or change a tag
  - Review pageviews for tags (browsing data) - are any of our tags not used/needed?
  - Review how tags are applied to items by our institutional members when they upload resources

# Thank you!

Contact:

Jessica BrodeFrank - [jbrodefrank@chicagocollections.org](mailto:jbrodefrank@chicagocollections.org)

Kate Flynn - [kef@uic.edu](mailto:kef@uic.edu)

Erin Matson - [matson@chicagohistory.org](mailto:matson@chicagohistory.org)

Gretchen Neidhardt - [gretchen.neidhardt@northwestern.edu](mailto:gretchen.neidhardt@northwestern.edu)

A decorative graphic on the left side of the slide consisting of two overlapping parallelograms. The front one is blue and the back one is a light green. They are positioned diagonally, with the blue one partially covering the green one.

# Creating and Sharing Date Metadata

Annamarie C. Klose, The Ohio State University Libraries

# REALITY CHECK



Photo by belig at Morguefile.com,  
<https://morguefile.com/p/738291>

No matter what metadata standard you use to create records, sharing it usually involves Dublin Core, e.g. OAI-PMH standard.

The end result will involve many compromises.

# How do you represent today's date?

March 4, 2024

04.03.2024

4th of March 2024

4. März 2024

03/04/2024

Ngày 4 tháng 3 năm 2024

3/4/24

4 Μαρτίου 2024

2024-03-04

2024 年 3 月 4 日  
2024-Nen 3 tsuki 4-nichi

4 de marzo de 2024

4 березня 2024 р



John Tenniel illustration in Lewis Carroll's  
*Alice's Adventures in Wonderland* (1865)


## PUBLIC SERVICE ANNOUNCEMENT:

OUR DIFFERENT WAYS OF WRITING DATES AS NUMBERS CAN LEAD TO ONLINE CONFUSION. THAT'S WHY IN 1988 ISO SET A GLOBAL STANDARD NUMERIC DATE FORMAT.

THIS IS *THE* CORRECT WAY TO WRITE NUMERIC DATES:

2013-02-27

THE FOLLOWING FORMATS ARE THEREFORE DISCOURAGED:

02/27/2013 02/27/13 27/02/2013 27/02/13  
20130227 2013.02.27 27.02.13 27-02-13  
27.2.13 2013.II.27.  $27\frac{1}{2}$ -13 2013.158904109  
MMXIII-II-XXVII MMXIII <sup>LVII</sup>CCCLXV 1330300800  
 $((3+3)\times(111+1)-1)\times 3/3-1/3^3$  2013 miss  
10/11011/1101 02/27/20/13  $\begin{matrix} 2 & 3 & 1 & 4 \\ 0 & 1 & 2 & 3 & 7 \\ 5 & 6 & 7 & 8 \end{matrix}$  

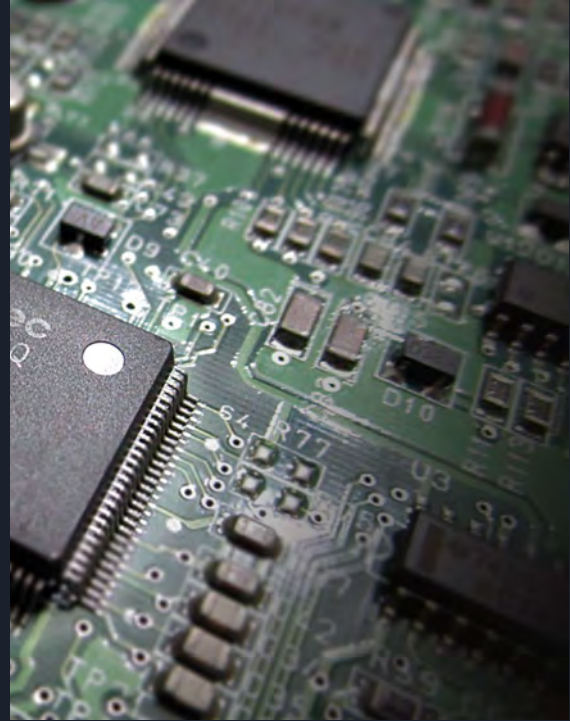
Randall Munroe, "ISO 8601," <https://m.xkcd.com/1179/>





Standards matter.

Encode your date  
values if you  
possibly can.





# One date standard to rule them all

20240304	ISO 8601 Basic Format, ANSI INCITS 30
2024-03-04	ISO 8601 Extended Format, W3CDTF, LOC's EDTF

Note: My research refers to ISO 8601-1:2019 and ANSI INCITS 30-1997 (R2008). ISO 8601 have since had a 2022 amendment and ANSI INCIT 30-1997 has a R2013 update.



# W3C Date and Time Formats

Year	YYYY	2024
Year and month	YYYY-MM	2024-03
Complete date	YYYY-MM-DD	2024-03-04
Complete date plus hours and minutes	YYYY-MM-DDThh:mmTZD	1997-07-16T19:20+01:00
Complete date plus hours, minutes, and seconds	YYYY-MM-DDThh:mm:ssTZD	1997-07-16T19:20:30+01:00
Complete date plus hours, minutes, seconds, and a decimal fraction of a second	YYYY-MM-DDThh:mm:ss.sTZD	1997-07-16T19:20:30.45+01:00

W3C Date and Time Formats, 15 September 1997, <https://www.w3.org/TR/NOTE-datetime>



# LOC's Extended Date/Time Format

- Draft published in 2012.
- Revised and formalized in 2019. It was integrated into ISO 8601-2019!
- Can be used to express ambiguous dates, approximate dates, date ranges, seasons, and much more!

But how much is it actually used?

Zavalina, et al. (2015) studied the use of EDTF in DPLA. My research partners and I are looking for how things have changed since EDTF was formalized.

From personal experience, some communities use a subset of EDTF for digital collections.

# Avoid garbage values



Photo by DodgertonSkillhouse at Morguefile.com, <https://morguefile.com/p/898683>

Examples: Undated, Unknown, No date, n.d.

If you want to include that information, you can put that in the Description field.

# Excel isn't your friend but OpenRefine is



<https://twitter.com/jxxf>



When formatting date values in Excel, be careful about its tendency to format them according to local standards, e.g. American dates. There are workarounds, e.g. load data set as .csv file and assign text values.

However, OpenRefine is a better tool when working with date metadata.



## Consider the quirks and limitations with your platform

- Dublin Core's temporal field is “aboutness.” However, many repositories use this field to group items by Date Created and/or Date Issued. It's a workaround that has consequences.\*
- When possible, create custom metadata fields to address these artificial groups for search instead.
- Other workaround include having Display Date Fields and Searchable Date fields.

\*Caveat: I'm hoping to correct that issue in my own repository with legacy metadata.





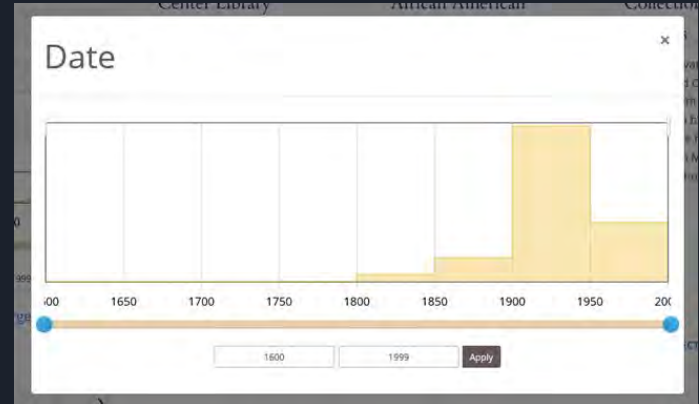
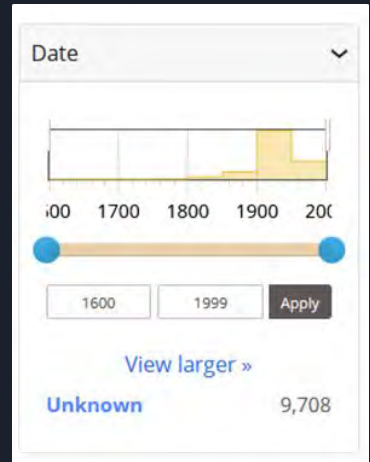
# Begin with the end in mind

- Well-known crosswalks, e.g. LOC's MARC to Dublin Core and Getty's, should be seen as inspiration. However, you need to consider local metadata considerations. Create your own crosswalk before sharing your metadata.
- Harvesting records will magnify issues. If possible, review and/or sample your data set after mapping but before sharing your harvest. You may find that you need to tweak your vision to make records more findable.
- The "Created Date" label in DPLA is awkward semantically. If you see the DPLA documentation, it's just dc:date which is broader. Consider more than just created dates if you want your records to be findable. However, there may be dates you choose to omit.

# Hindsight is 20/20

- When starting a new digital collection/project, it's useful to consider the impact of harvesting before you create records. What works in your system may require substantial changes when harvesting records.
- Even if you can't support encoded dates now, consider what you might be able to do down the road.
- Encoded dates can serve researchers in other ways, e.g. collections as data, in the future if LAMs make a greater investment.

*Emory Digital Collections' Date slider leveraged EDTF encoded dates to search and browse records by date.*



# Samvera community



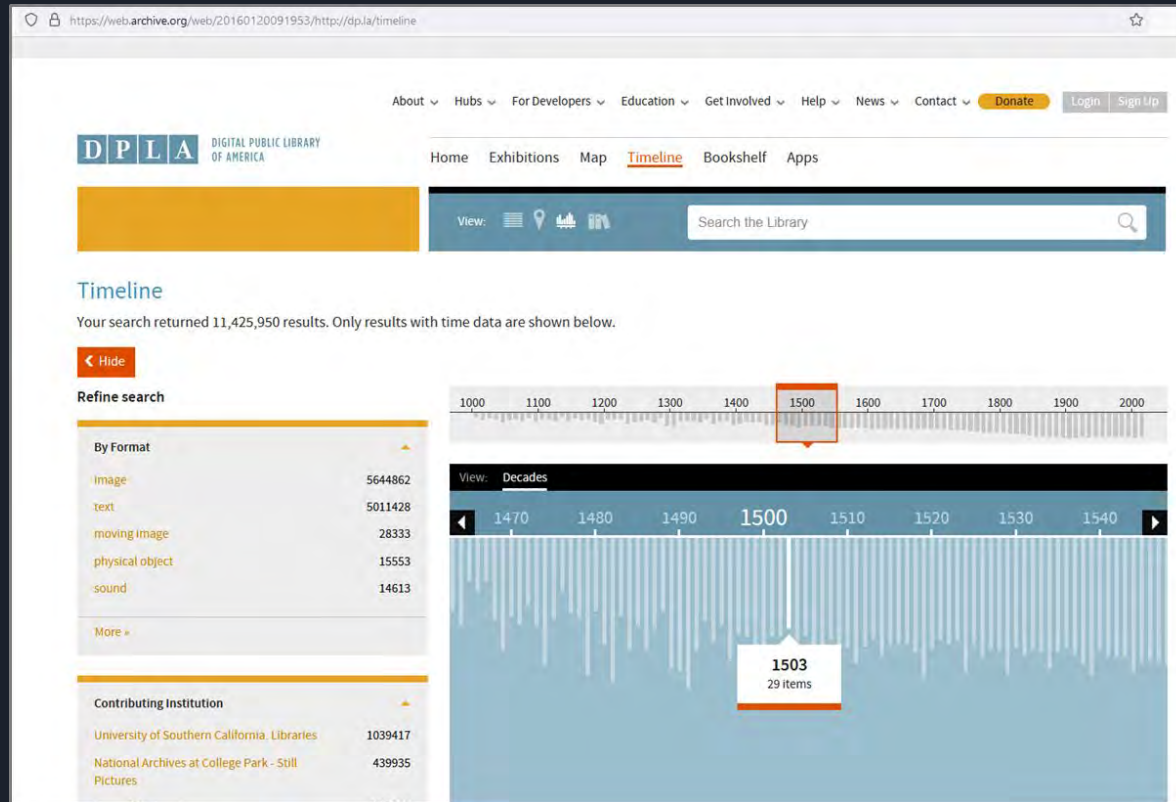
Several Samvera Hyrax institutions have forked and implemented the EDTF Humanize Ruby gem (<https://github.com/inukshuk/edtf-ruby>) with their repositories to encode a limited set of EDTF patterns.

Implementers include:

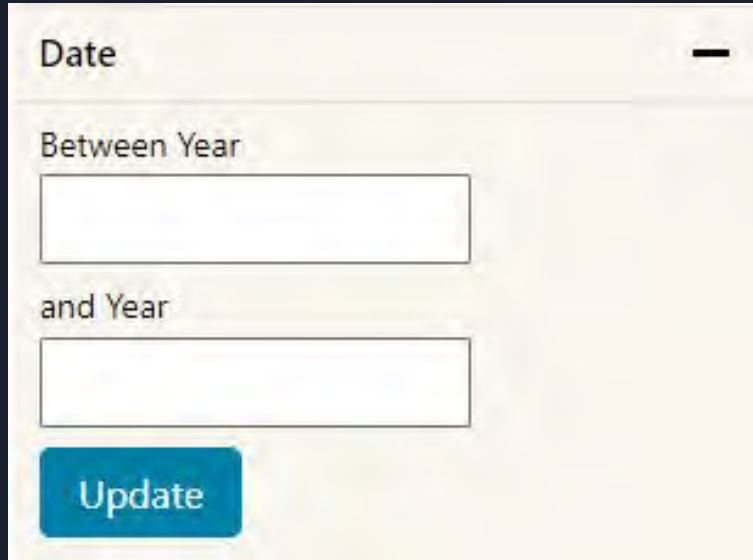
- Duke University Libraries
- Emory University Library
- The Ohio State University Library

CAVEAT: Coding converts EDTF dates to English only?

# Old DPLA interface for date searching



# Don't count on the temporal enrichment process in DPLA



A screenshot of the DPLA Date filter interface. It features a title bar with the word "Date" and a minus sign. Below the title bar, the text "Between Year" is followed by a text input field. Below that, the text "and Year" is followed by another text input field. At the bottom left, there is a blue button with the word "Update" in white text.

GET [https://api.dp.la/v2/items?id=\[DPLA ID\]&api\\_key=\[API Key\]](https://api.dp.la/v2/items?id=[DPLA ID]&api_key=[API Key])

If you look at DPLA records via the API, you can see the records that were sent for harvest and the resulting metadata.

While there's various ways to interact with the API, Postman is a free app with a GUI that can make the process easier.

Unfortunately, using the API has shown me that the temporal enrichment process isn't necessarily happening as expected.

# Example 1



DPLA record:

<https://dp.la/item/680dde2bf560a8d62fca37e5932b329d>

API view of record:

```
"date": [  
  {  
    "displayDate": "1500~"  
  }  
]
```

Ingest Date: 2023-12-20T20:22:06.822Z

JSON manifest:

[https://library.osu.edu/dc/concern/generic\\_works/js956h40v.json](https://library.osu.edu/dc/concern/generic_works/js956h40v.json)

date\_created:["1500~"]

"Book of Hours,"

<https://hdl.handle.net/1811/87086df9-ff2f-48fb-81ee-90c3608b09e2>

## Example 2



(1399). *Book of Hours*. Retrieved from  
<https://libwww.freelibrary.org/digital/item/3836>

DPLA record:

<https://dp.la/item/ad90056eea3921712b7995fdbcc3c32d>

API view of record:

```
<dcterms:date>1375 - 1399</dcterms:date>  
"date": [  
  {  
    "begin": "1375",  
    "displayDate": "1375 - 1399",  
    "end": "1399"  
  }  
]
```

**Ingest Date: 2024-01-30T18:54:35.861Z**

---

**Creation Year (Single Year or Range Begin): 1399**



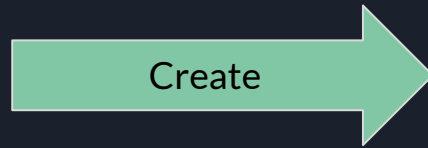
If all else fails...

Fix your metadata after harvest and then reharvest. Lots of institutions have corrected or otherwise changed date metadata after sharing it.

We're all human. :)



# Iterative metadata workflows





Shout-out to research partners:

- Scott Goldstein, McGill University
- Morris Levy, OCLC

Thank you!

Annamarie Klose, [klose.16@osu.edu](mailto:klose.16@osu.edu)



# Resources (part 1)

- American National Standard. (n.d.). *ANSI INCITS 30-1997 (R2008)*. ANSI webstore.  
<https://webstore.ansi.org/standards/incits/ansiincits301997r2008>
- Banerjee, K., & Reese, T. (2019). *Building digital libraries* (2nd ed.). ALA Neal-Schuman.
- Chalabi, M. (2013, December 16). *Why do Americans write the month before the day?* The Guardian.  
<https://www.theguardian.com/news/datablog/2013/dec/16/why-do-americans-write-the-month-before-the-day>
- Digital Public Library of America, Metadata application profile,  
<https://pro.dp.la/hubs/metadata-application-profile>
- International Organization for Standards. (2019). *ISO 8601-1:2019: Date and time: Representations for information exchange: Part 1: Basic rules*. ISO.  
<https://www.iso.org/standard/70907.html>
- International Organization for Standards. (2019). *ISO 8601-1:2019: Date and time: Representations for information exchange: Part 2: Extensions*. ISO.  
<https://www.iso.org/standard/70908.html>



## Resources (part 2)

- Klose, A., Goldstein, S., & Levy, M. (2021). Out of the the past: The challenges of B.C.E. dates in metadata. *ASIS&T Proceedings*, 58(1), 756-758. <https://doi.org/10.1002/pr2.552>
- Library of Congress. (2019, February 4). *Extended Date/Time Format (EDTF) specification*. Retrieved June 6, 2021 from <http://www.loc.gov/standards/datetime/>
- Miller, S. J. (2022). *Metadata for digital collections* (2nd ed.). ALA Neal-Schuman.
- MIT International Students Office. (n.d.). *Americanisms, date format in the United States*. MIT. <https://iso.mit.edu/americanisms/date-format-in-the-united-states/>
- World Wide Web Consortium. (1998, August 27). *Date and Time Formats*. W3C. <https://www.w3.org/TR/NOTE-datetime>
- World Wide Web Consortium. (2020, January 13). *Quality Assurance: Tips for Webmasters*. W3C. <https://www.w3.org/OA/Tips/iso-date>
- Zavalina, O. L., Alemneh, D. G., Kizhakkethil, P., Phillips, M. E., & Tarver, H. (2015). "Extended Date/Time Format (EDTF) in the Digital Public Library of America's metadata: Exploratory analysis. *Proceedings of the Association for Information Science and Technology*, 52(1), 1–5. <https://doi.org/10.1002/pr2.2015.145052010066>
- Zeng, M. L., & Qin, J. (2022). *Metadata* (3rd ed.). ALA Neal-Schuman.



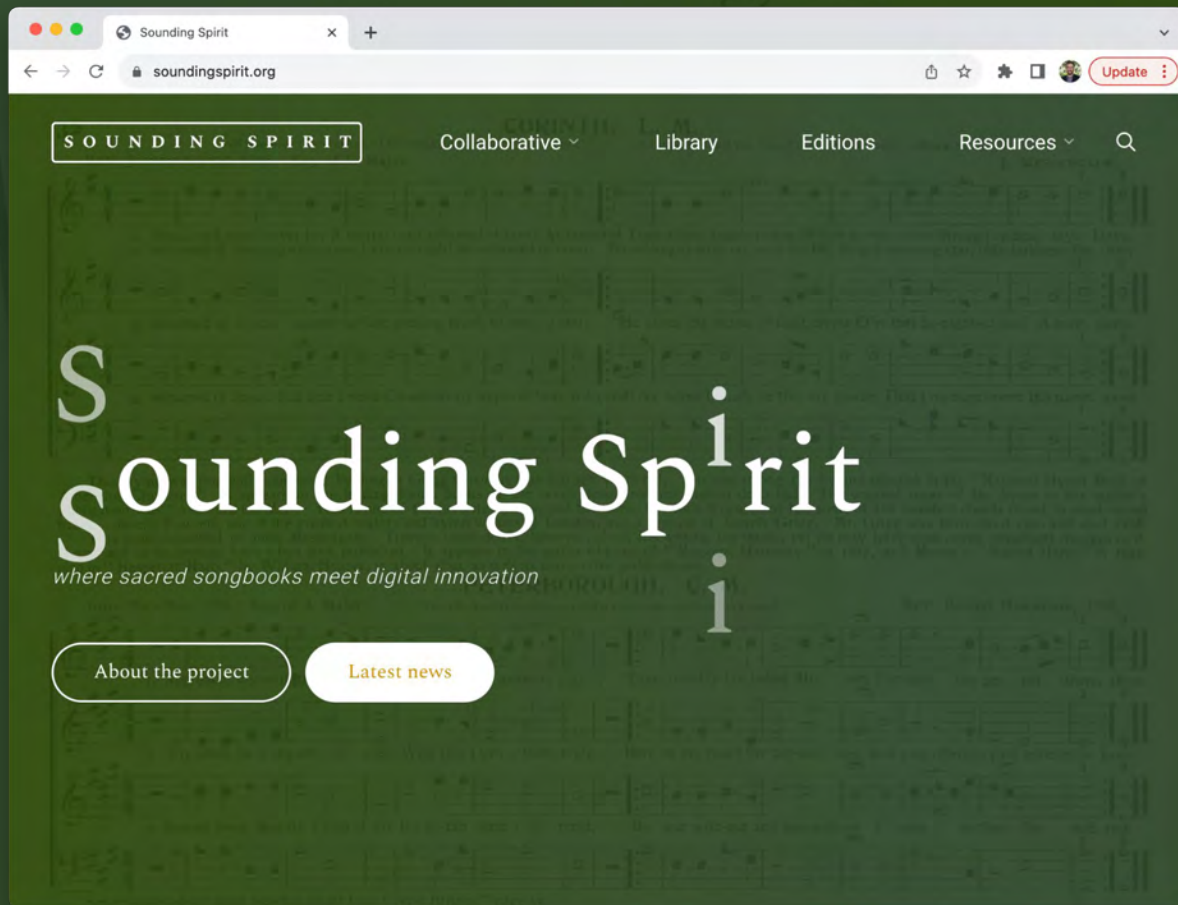
# Tools & Tips

- DPLA, API Codex, <https://pro.dp.la/developers/api-codex>
- DPLA Intro to OpenRefine, <https://dp.la/news/dpla-hosts-introduction-to-openrefine>
- Library Carpentry, OpenRefine, <https://librarycarpentry.org/lc-open-refine/>
- OpenRefine, <https://openrefine.org/>
- Postman, <https://www.postman.com/>



# Developing a Metadata Application Profile Supporting Interoperability, Research, and Data Harvesting for a Digital Library of Music Sources

Jesse P. Karlsberg, Emory University  
ALA Core Interest Group Week, March 4, 2024



# Sounding Spirit Collaborative

Promotes research, teaching, and public engagement with vernacular sacred songbooks.

- Scholarly editions
- Digital library



Sounding Spirit Digital Library x +

library.soundingspirit.org

ABOUT VOLUMES COLLECTIONS RESEARCH LAB

readux

# Sounding Spirit Digital Library

where sacred songbooks meet digital innovation

## Featured Collections

**Denomination: Restorationist**  
3 volumes

The Restorationist movement materialized in the US South in response to growing strife between fractious religious groups. Often traced to the...

**Geography: Tennessee**  
8 volumes

By the early twentieth century, Tennessee was becoming an important center for the production and consumption of gospel music. In 1903,...

**Sabbath School**  
8 volumes

Sunday or Sabbath schools were first established in the United States in the 1790s. Though initially confined to the urban Northeast, they later...

Checklist of Southern Sacred Music Imprints, 1850-1925

Title	Responsibility
Collection of hymns and tunes for public, social and domestic worship.	1
Hymns in the Cherokee Language	Chamberlin, A. N.
26 hymns for Jewish worship, for solo and chorus, with accompaniment of the organ	Kitzger, Frederick E.
200 popular songs and hymns	Shawalter, A. J.
223 Melodien deutscher Kirchengesänge meist aus dem 16. und 17. Jahrhundert, in ihren ursprünglichen Rhythmen und Tönen	Layitz, Fr.
225 Melodien deutscher Kirchengesänge meist aus dem 16. und 17. Jahrhundert, in ihren ursprünglichen Rhythmen und Tönen	Layitz, Fr.

SHOW ALL FEATURED COLLECTIONS >

# Sounding Spirit Digital Library

Expanding a pilot launched in 2020

1,300 songbooks from US South and its diasporas

Published 1850–1925

Forthcoming fall 2024





# Sounding Spirit Digital Library

Includes gospel  
songbooks, spiritual  
collections, tunebooks,  
hymnals, Sunday  
school books

NEH-funded

Archives at seven  
contributing institutions

# Sounding Spirit Digital Library Contributing Archives

- The John Hay Library at Brown University,
- Pitts Theology Library and Stuart A. Rose Manuscript, Archives, and Rare Book Library at Emory University,
- The Center for Popular Music at Middle Tennessee State University,
- The John Jacob Niles Center for American Music and UK Libraries at the University of Kentucky,
- The University of Michigan Library and William L. Clements Library at the University of Michigan,
- The Department of Special Collections and University Archives and Helmerich Center for American Research at the Gilcrease Museum at the University of Tulsa, and
- The Southern Baptist Theological Seminary Archives and Special Collections.

# Metadata in Sounding Spirit: Values and Considerations

## Metadata and Research Data Working Group

- Diana Folsom (Director of Digital Collections, Gilcrease Museum)
- Erin Fulton, co-chair (Musicology PhD Candidate, University of Kentucky)
- Andrew Hankinson (Sr. Software Developer, RISM)
- Jesse P. Karlsberg, co-chair (Sr. Digital Scholarship Strategist, Emory Univ.)
- Cory Lampert (Head of Digital Collections, UNLV)
- Em Nordling (English PhD Candidate, Emory University)
- Sara Palmer (Digital Text Specialist, Emory University)



# Metadata in Sounding Spirit: Values and Considerations

1. User-friendly **browsing** and faceted **search**
2. **Interoperability** through data-sharing partnerships
3. Useful for quantitative **research**

# Metadata Application Profile

	Data Type	Repeatable	Repeatability	Controlled	Entry	Style Needed	Description	Source	Values	Notes	DCMI	DIG Field	DIG Status	RISM Field	RISM Status	ATLA Field	ATLA Status	Emory Field	Emory Status	Field Index	Contra Index
URL	URL	Y	N	N	Reader	rule	Volume landing page in Sounding Spirit digital library.	Based on the form of URLs in the Sounding Spirit digital library where PID corresponds with the PID for the volume.	https://library.soundingspirit.org/volume/PID/page/1	For RISM: Note about external resource (RIS 3a) Link to record in Sounding Spirit Digital Library. Link type (RIS 3a) Digitized source. For RISM: Threni Numberspage (RIS 1a) entire correspond to a "Sounding Spirit Digital Library" Bibliographic reference (RIS 1a) that will need to be added to the Secondary literature Authority.	Identifier (name identifier)	URL	required	External resource URL (RIS 3a) Digitized source Numberspage (RIS 1a)	optional	URL	required	Access URL	required if applicable	1	n/a
PID	string	Y	N	N	Predefined	n/a	Permanent unique identifier for the volume.	Based on the file naming convention spotted out in the digitization standards document. Page image of the title page of the volume. Page image of the front cover or bibliographic record for the work in the institutional catalog if the title page is missing.	xxxx-First six letters of volume's title (include capitalization, exclude alpha24) character repository code (four digit sequence number)		n/a	n/a	n/a	n/a	optional	n/a	n/a	Identifier - Local	required	2	1
Short Title	string	Y	N	N	Manual	Basic	Short title of the work.	A shortened title, in this case. For a work with a subtitle, the pre-canon title. For a long title with no subtitle, an identifying subset of the title. The full title, in sentence case, with leading articles capitalized and moved to the end of the short title and the short title repeated from the subtitle. If a subtitle is present, by " " with spaces on either side of the separator. If the title page is missing and the front cover does not feature comparable information, enclose the full title in square brackets.		Alternative title (name alternative)	n/a	n/a	Additional title (730)	optional	Alternative Title, optional	n/a	n/a	3	3		
Title	string	Y	N	N	Manual	Basic	Full title of the work.	Page image of the title page of the volume. Page image of the front cover or bibliographic record for the work in this institutional catalog if the title page is missing. The full contents of the title page, transcribed diplomatically exactly as it appears on the title page, in the format requested by RISM. In practice, this may include all information on the title page, such as the composer's name or other indications. For printed music, include everything on the title page including the imprint. Transcribe the title in its entirety. Do not abbreviate. Indicate line breaks with ( ) (the vertical pipe) with a single space before and after. Use an exclamation point or brackets ( ) or ( ) to indicate an unusual or incorrect spelling. If reading order diverges from line order, as with title-by-side text in volumes, prioritize reading order and treat sections on the same horizontal plane as their own lines, leaving vertical pipe characters to mark these as line breaks in the same way that we mark vertical line breaks. Transcribe lead and small caps in all caps. If the title page is missing but the title appears on the cover, transcribe that page and prefix the entry with "Covered". If working with a single page document (a broadside), include information that would be found on a title page or cover title, author, publishing information, etc.) but do not transcribe the entire contents of the page and prefix the entry with "Caption title, p. [1]". replacing "[1]" as needed with the page number of the source of the transcribed text. Leave the field blank if the title page is missing and the book does not include title information.	Title (name title)	Title	required	Standardized title (245 \$a)	required	Title	required	Title	required	4	3		
Title on Source	string	Y	N	N	Manual	Basic	Full diplomatic title of the work.	Page image of the title page of the volume. Page image of the front cover.			n/a	n/a	n/a	Title on source (245 \$a)	required	n/a	n/a	n/a	n/a	5	3
Responsibility	string	N	N	N	Manual	Basic	All credited authors or editors of a work, presented in the order in which they are credited, in a human-readable format.	Page image of the title page of the volume. Cover, title page verso, table of contents, or running title if the responsibility is unclear from the title page or the title page is missing. Bibliographic record for the work in the institutional catalog if the responsibility is unclear from the volume's pages. Page image of the title page of the volume. Cover, title page verso, table of contents, or running title if the responsibility is unclear from the title page or the title page is missing. Bibliographic record for the work in the institutional catalog if the responsibility is unclear from the volume's pages.	One or more names separated by commas with first and middle names followed by last names. Names are recorded in the same manner (with initials and abbreviations preserved) as they are found in the source. Titles of authors or editors, such as "Rev." or "Dr.", are omitted except for individuals credited by the names of their partner, as in a married woman credited by their husband's names and an appellation such as "Mrs." if the responsibility was derived from the catalog record, enclose it in square brackets.	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	6	5		

# Metadata Application Profile

## Data and data entry

- Data Type
- Requirement
- Repeatabe
- Controlled
- Entry
- Skills Needed

	A	B	C	D	E	F	G
1	Fields	Data Type	Requirement	Repeatabe	Controlled	Entry	Skills Needed
11	Contributor URI	URL	N	Y	LCNAF; VIAF; RISM Online	Manual	Basic
12	Place of Publication	string	N	Y	N	Manual	Basic
13							

# Metadata Application Profile

## About the field

- Description
- Source
- Values
- Notes

	A	H	I	J	
1	Fields	Description	Source	Values	Notes
30	Publisher URI	The name of the publisher or publishers that issued the work, excluding any firms that are distributors but not publishers. If no publisher is credited but a printer is, the name of this printer.	Page image of the title page of the volume. Library of Congress Name Authority File. Virtual International Authority File. RISM Online Authority File. <a href="https://id.loc.gov/authorities/names.html">https://id.loc.gov/authorities/names.html</a> ; <a href="https://viaf.org/">https://viaf.org/</a> ; <a href="https://rism.online/?mode=institutions">https://rism.online/?mode=institutions</a>	URIs of LCNAs when these exist. URIs of VIAFs if these exist but LCNAs do not. New RISM Online Name authority URIs created in collaboration with RISM if neither LCNAs or VIAFs exist.	For names that leave this field RISM. See the field for context
31	Language	The language or languages in which the work is written.	Representative page images of the volume. ISO 639-3: <a href="https://iso639-3.sil.org/code_tables/639/data">https://iso639-3.sil.org/code_tables/639/data</a>	Full English name of the language as specified in ISO 639-3. If English and another language appear in the same work, list the non-English language first for ease of sorting. List of languages: Cherokee, Choctaw, Creek, Czech, Delaware, Dutch, English, German, Hebrew, Latin, Portuguese, Potawatomi, Spanish	
32	Language Code	The language or languages in which the work is written.	Representative page images of the volume. ISO 639-3: <a href="https://iso639-3.sil.org/code_tables/639/data">https://iso639-3.sil.org/code_tables/639/data</a>	Three-letter code for the language as specified in ISO 639-3. If English and another language appear in the same work, list the non-English language first for ease of sorting. List of languages: chr, cho, mus, ces, del, nld, eng, ger, heb, lat, por, pol, spa	
33				For paginated works, a number indicating the overall number of pages in the work, with any sequences of separate pagination in their original Roman and Arabic format indicated by the insertion of a comma, followed by the word "pages"; for sequences of pages split into Latin and Arabic numbered sub-sequences, list each subsequence but include starting as well as ending numbers, separated by "-" (en dashes) for the sub-sequences after the first; for unnumbered sequences of pages with content in paginated	



# Metadata Application Profile

## Crosswalk

Equivalent fields and requirements for

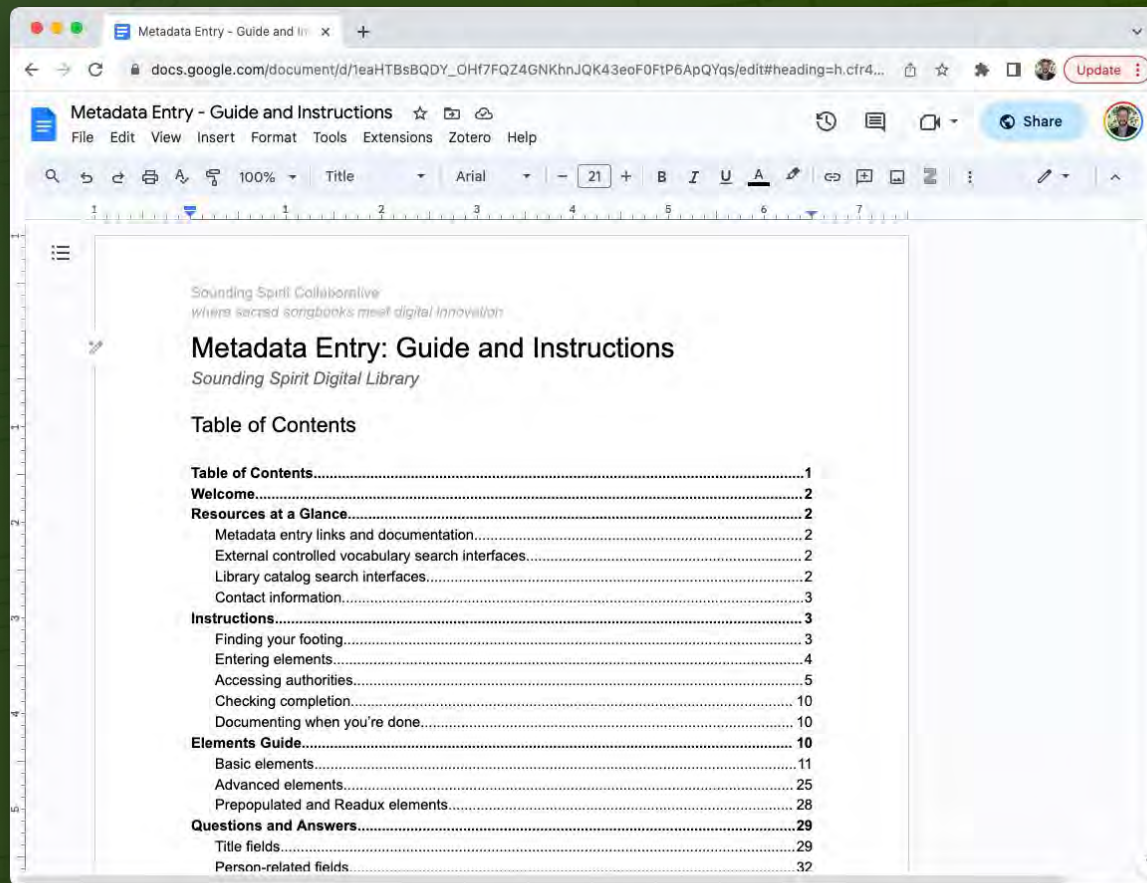
- DCMI
- DLG/DPLA
- RISM
- ATLA
- Emory

	A	L	M	N	O	P	Q	R	S	T
1	Fields	DCMI	DLG Field	DLG Status	RISM Field	RISM Status	ATLA Field	ATLA Status	Emory Field	Emory Status
2	URL	Identifier (terms: identifier)	URL	required	External resource URL (856 \$u)	optional	URL	required	Access URL	required if applicable
3	PID	n/a	n/a	n/a	Bibliographic reference Number/page (691 \$n)	optional	n/a	n/a	Identifier - Local	required
4	Short Title	Alternative Title (terms: alternative)	n/a	n/a	Additional title (730)	optional	Alternative Title	optional	n/a	n/a
5	Title	Title (terms: title)	Title	required	Standardized title (240 \$a)	required	Title	required	Title	required



# Metadata Application Profile

## Guide and Instructions



The screenshot shows a Google Docs interface with a document titled "Metadata Entry - Guide and Instructions". The document content includes a header with the text "Sounding Spirit Collaborative" and "where sacred songbooks meet digital innovation", followed by the title "Metadata Entry: Guide and Instructions" and the subtitle "Sounding Spirit Digital Library". Below this is a "Table of Contents" section with a list of topics and their corresponding page numbers.

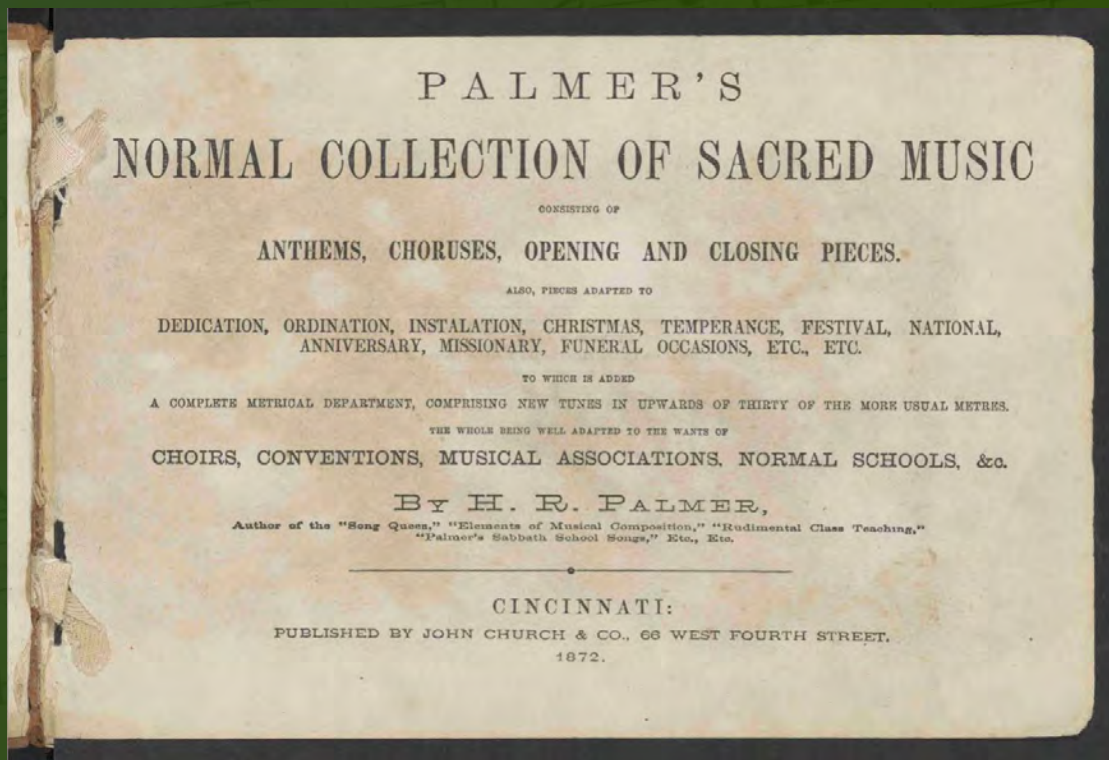
<b>Table of Contents</b>	<b>1</b>
<b>Welcome</b>	<b>2</b>
<b>Resources at a Glance</b>	<b>2</b>
Metadata entry links and documentation	2
External controlled vocabulary search interfaces	2
Library catalog search interfaces	2
Contact information	3
<b>Instructions</b>	<b>3</b>
Finding your footing	3
Entering elements	4
Accessing authorities	5
Checking completion	10
Documenting when you're done	10
<b>Elements Guide</b>	<b>10</b>
Basic elements	11
Advanced elements	25
Prepopulated and Readux elements	28
<b>Questions and Answers</b>	<b>29</b>
Title fields	29
Person-related fields	32

# Browse and Search

Pre-colon title for  
digestible display

Short Title:  
Palmer's Normal Collection  
of Sacred Music

vs...

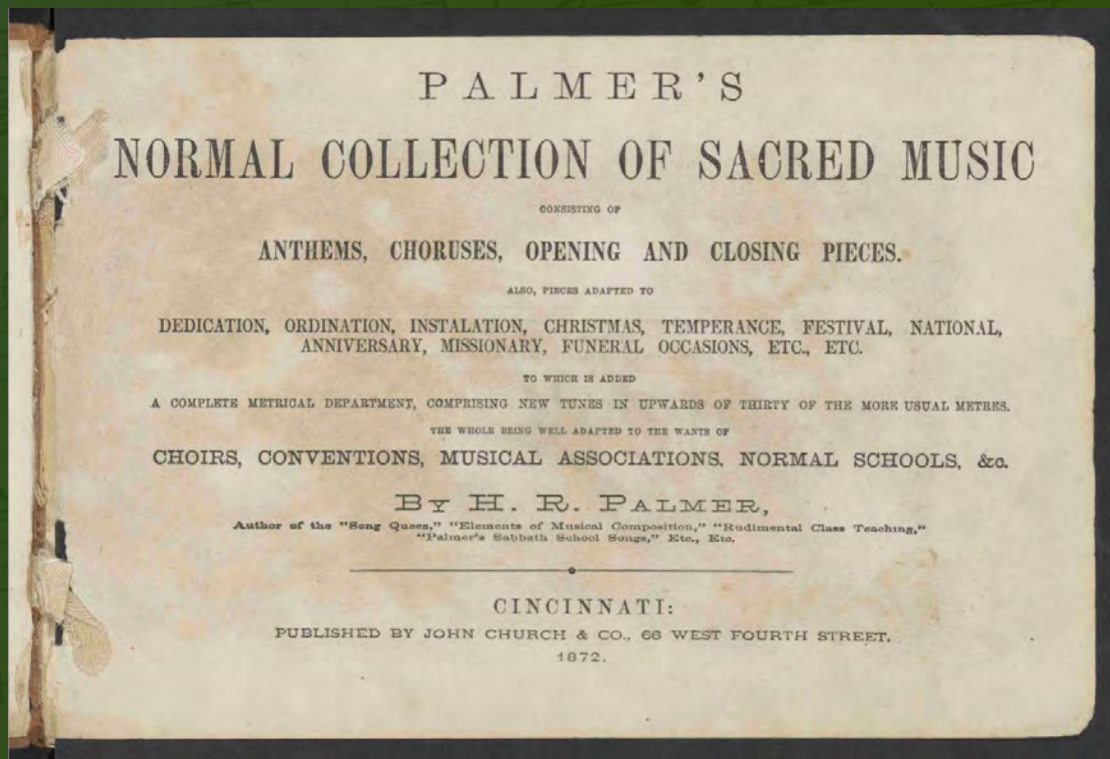


H. R. Palmer, ed., *Palmer's Normal Collection of Sacred Music* (1872). 1872-Palmer-EMU.

# Browse and Search

## Title:

Palmer's normal collection of sacred music ; consisting of anthems, choruses, opening and closing pieces. Also, pieces adapted to dedication, ordination, instalation [sic], Christmas, temperance, festival, national, anniversary, missionary, funeral occasions, etc., etc. To which is added a complete metrical department, comprising new tunes in upwards of thirty of the more usual metres. The whole being well adapted to the wants of choirs, conventions, musical associations, normal schools, &c.



H. R. Palmer, ed., *Palmer's Normal Collection of Sacred Music* (1872). 1872-Palmer-EMU.



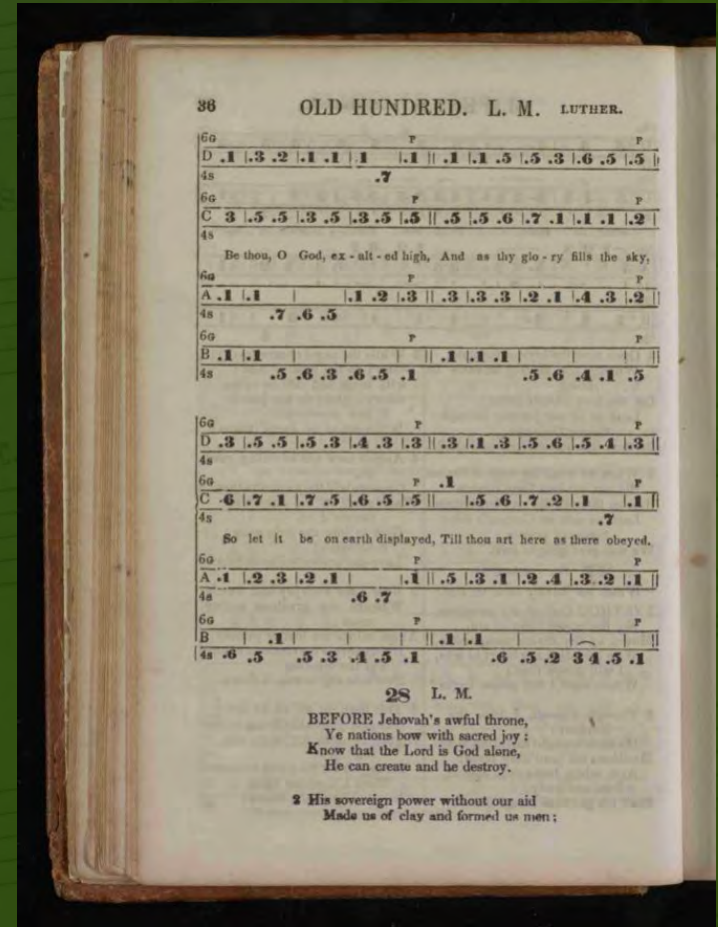
# Browse and Search

External and locally controlled  
vocabularies for faceting

Notation: Numeral

Possible values: Nil, Round, Seven-shape  
Aikin, Seven-shape Walker, Seven-shape  
Funk, Seven-shape Johnson, Seven-shape  
Swan, Seven-shape Hood, Four-shape,  
Numeral, Numeral staff

(Also: Characteristics, Content, Still in Use)



# Interoperability

## **External controlled vocabularies facilitate interoperable data**

Creators, Contributors, Publishers, and Institutions:

Library of Congress Name Authority File (LCNAF)

Places: Getty Thesaurus of Geographic Names (TGN)

Subjects: Faceted Application of Subject Terminology (FAST) Subject Headings

Languages: ISO 639-3

Dates: Extended Date/Time Format (EDTF)

# Interoperability

## Elements meeting requirements of data sharing partners

All: Subject, Holding Archive, Rights Statement URI

DLG: Medium, Digital Format, Type

RISM: Scoring Summary, Total Scoring, Title on Source, Physical Description, Local Identifier



# Interoperability

## Elements meeting partner requirements

Scoring Summary: V, Coro

Total Scoring:

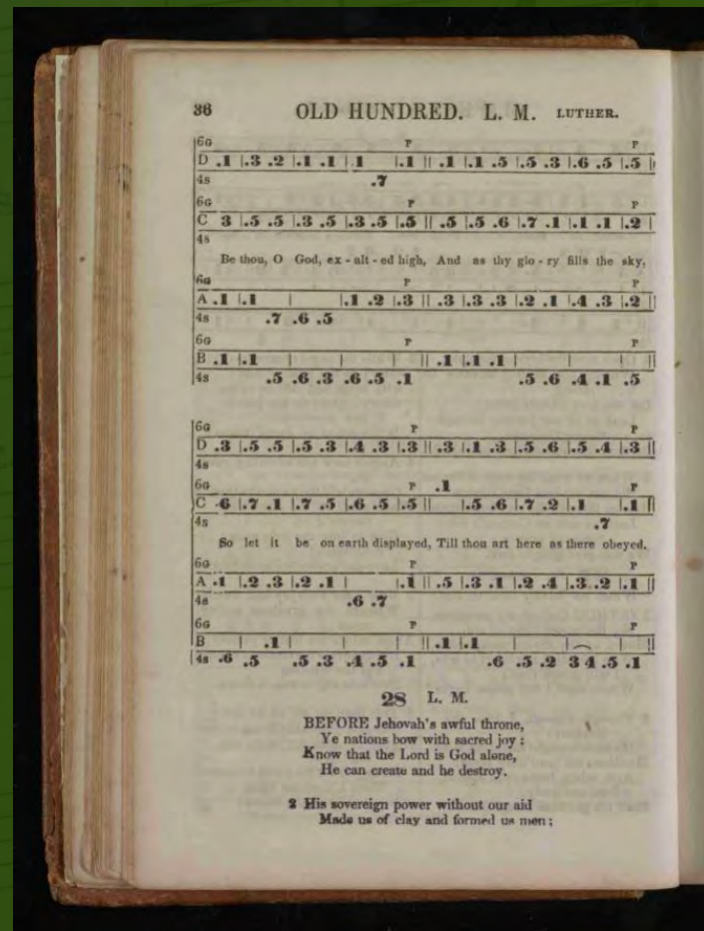
V

Coro Air

Coro Double Air

Coro Counter

Coro Bass

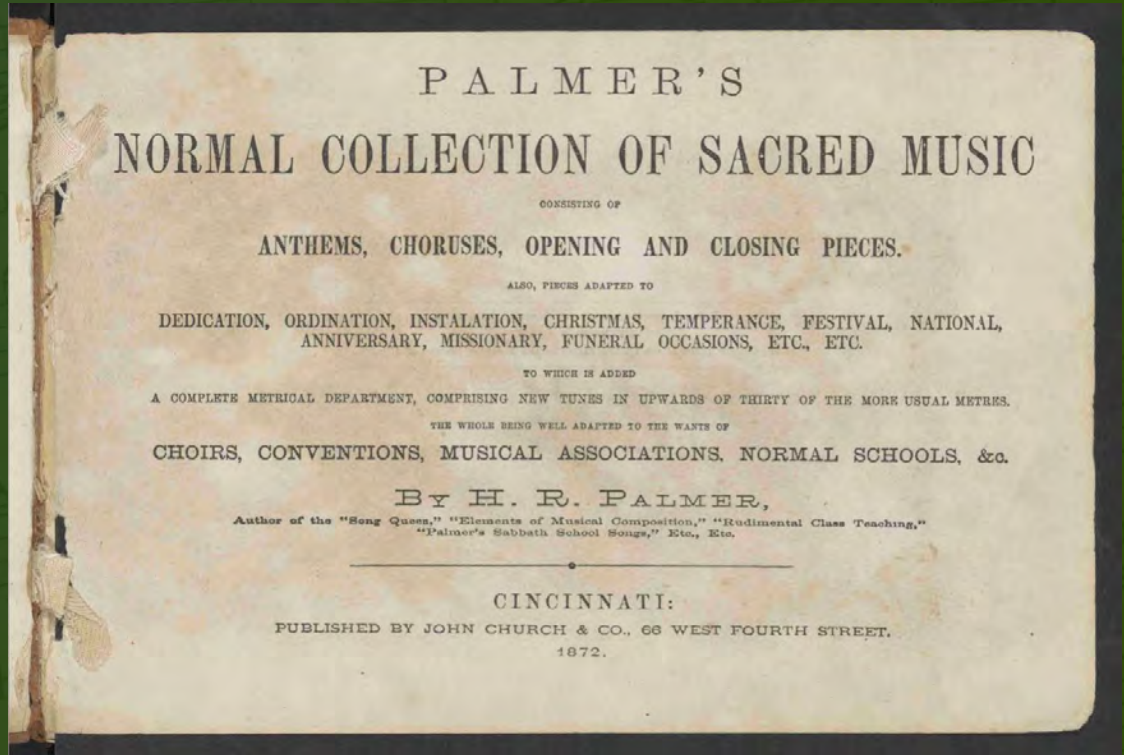


Silas W. Leonard and A. D. Fillmore, eds.,  
*The Christian Psalmist* (1851). 1851-Christ-UKY.

# Interoperability

Title on Source:

PALMER'S | NORMAL COLLECTION OF SACRED MUSIC | CONSISTING OF | ANTHEMS, CHORUSES, OPENING AND CLOSING PIECES. | ALSO, PIECES ADAPTED TO | DEDICATION, ORDINATION, INSTALATION [sic], CHRISTMAS, TEMPERANCE, FESTIVAL, NATIONAL, | ANNIVERSARY, MISSIONARY, FUNERAL OCCASIONS, ETC., ETC. | TO WHICH IS ADDED | A COMPLETE METRICAL DEPARTMENT, COMPRISING NEW TUNES IN UPWARDS OF THIRTY OF THE MORE USUAL METRES. | THE WHOLE BEING WELL ADAPTED TO THE WANTS OF | CHOIRS, CONVENTIONS, MUSICAL ASSOCIATIONS, NORMAL SCHOOLS, &c. | BY H. R. PALMER, | Author of the "Song Queen," "Elements of Musical Composition," "Rudimental Class Teaching," | "Palmer's Sabbath School Songs," Etc., Etc. | CINCINNATI : | PUBLISHED BY JOHN CHURCH & CO., 66 WEST FOURTH STREET. | 1872.



H. R. Palmer, ed., *Palmer's Normal Collection of Sacred Music* (1872). 1872-Palmer-EMU.



# Interoperability

**Data harvesting of custom metadata in Readux** (<https://readux.io>)

Built on Django/Python framework.

Ingest of custom metadata in spreadsheet form.

Display in sidebar and customizable incorporation into faceted search.

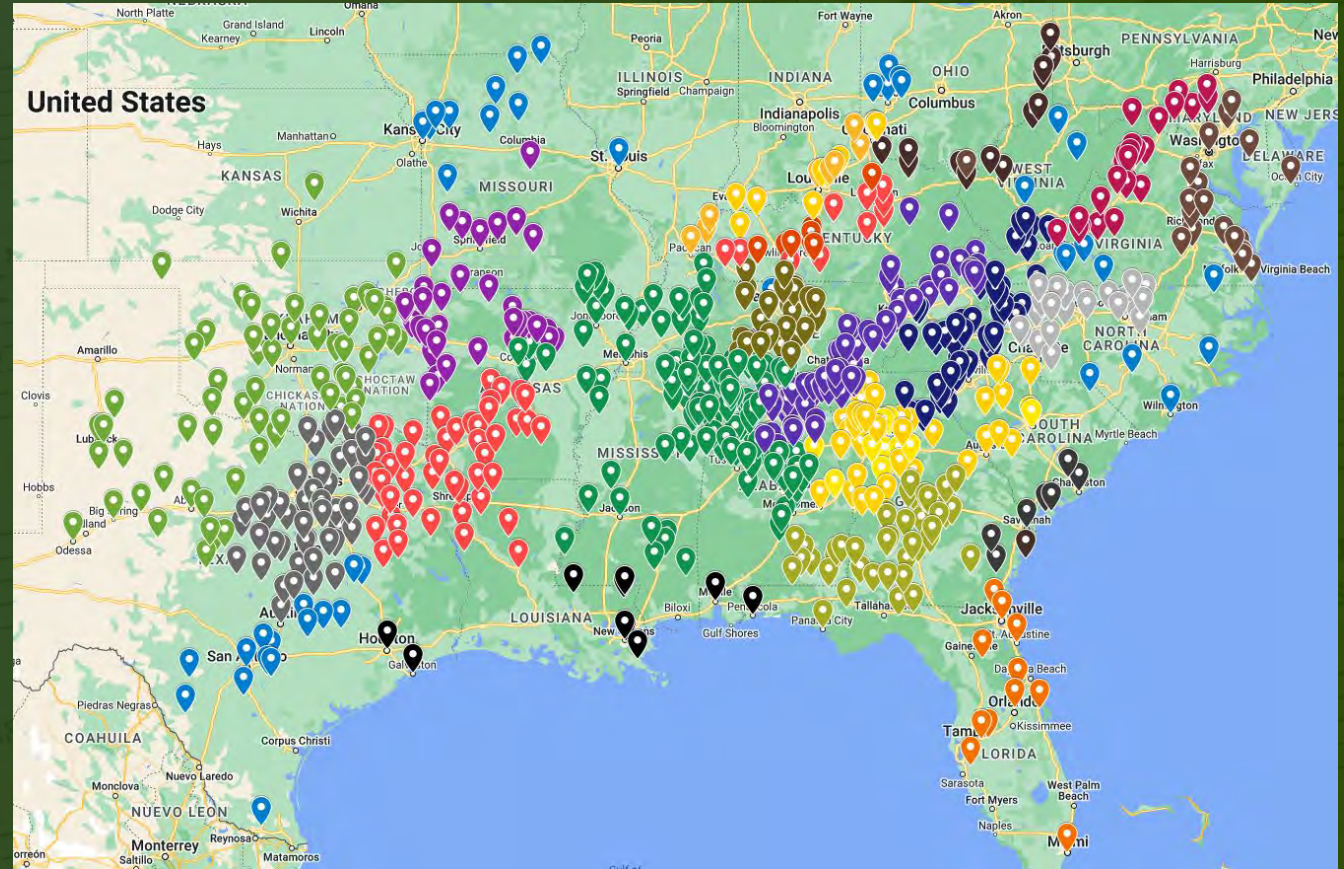
Available for data harvesting in IIIF manifests.

# Research

## Place of ...

- Contributor
- Affiliation
- Field Research
- Publication
- Manufacture
- Distribution
- Use

(Also: Dimensions,  
Characteristics)



# Futures

## Resources

- Metadata Application Profile: <https://bit.ly/sounding-spirit-metadata-map>
- Project site: <https://soundingspirit.org>
- Contact Jesse: [jesse.p.karlsberg@emory.edu](mailto:jesse.p.karlsberg@emory.edu)

**Next steps:** publishing library and metadata, indexing songs, sharing our work

**Thanks to:** Erin Fulton, Em Nordling, Lucy Wallitsch, Patrick Pender, Brenda Umutoniwase, Solomon Kim, Anne Evers, Bethany Branscome, Meredith Doster, and the members of the MRD WG. This project has been made possible in part by the National Endowment for the Humanities.



# Questions and Answers



Photo by [Matt Walsh](#) on [Unsplash](#)

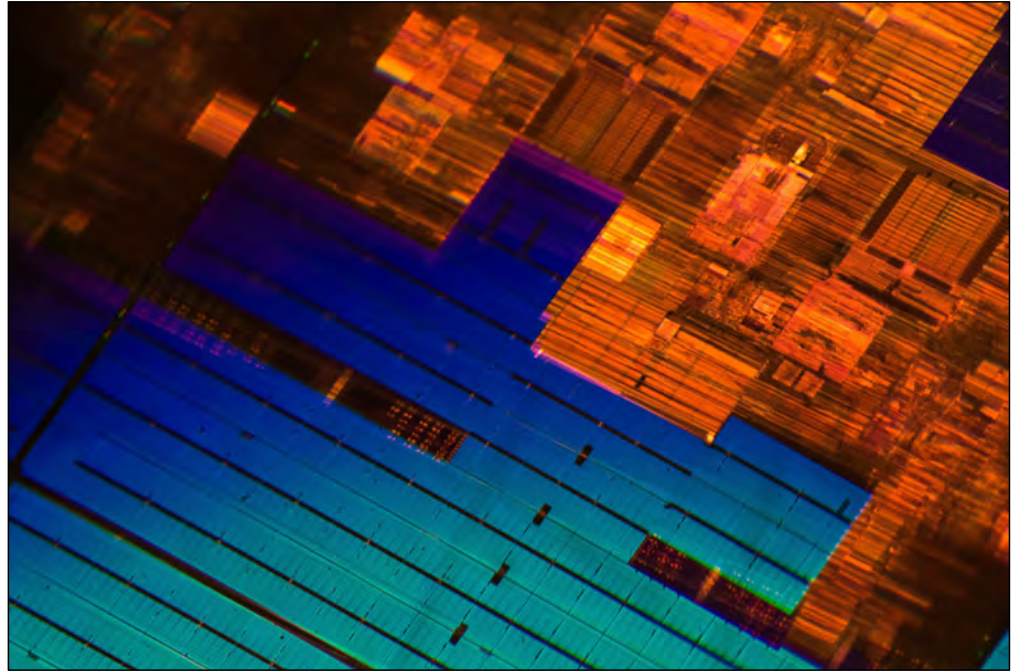


Photo by [Jason Leung](#) on [Unsplash](#)

# Announcements

- ALA Annual Program
  - Staying in MARC: Expectations for BIBFRAME to MARC conversion and the hybrid metadata future, *Sun. June 30 1-2 PM Pacific*
    - **Presenters:** Jeff Mixter (OCLC) and Elizabeth Miraglia (University of California, San Diego)
    - Note: Preliminary Schedule date and time
- Let's continue the conversation on ALA Connect!
  - Core Metadata Interest Group:  
<https://connect.ala.org/core/communities/community-home?CommunityKey=a38252d1-29e5-42e7-8b5d-a43a19f99aea>

# Thank You!



Photo by [Aaron Burden](#) on [Unsplash](#)