

# Intersections of Informatics, Information Science, and Digital Libraries

Professor Lorang  
University Libraries

[liz.lorang@unl.edu](mailto:liz.lorang@unl.edu)

November 1, 2018

# Goals for Today

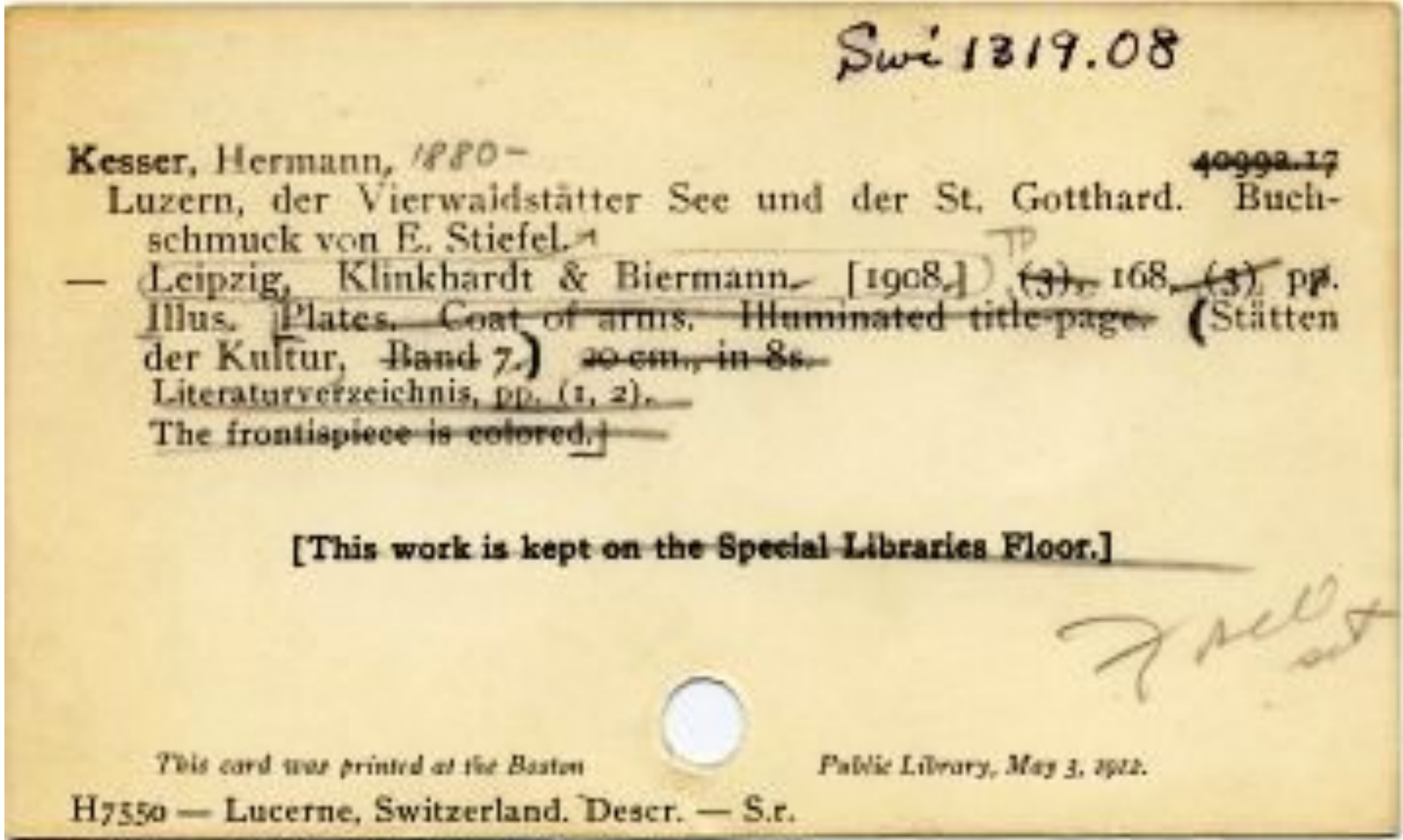
- Introduce students to the intersection of informatics, the work of librarians and archivists, and the development of digital libraries
- Introduce some key concepts in critical information literacy and relevance to informatics
- Explore some current challenges in digital libraries that might be addressed with informatics-driven approaches






Public Domain, <https://commons.wikimedia.org/w/index.php?curid=2651819>






Public Domain, <https://commons.wikimedia.org/w/index.php?curid=32519802>



BOOK

**Luzern, der Vierwaldstätter See und der St. Gotthard,**  
Kesser, Hermann, 1880-1952. / Leipzig, Klinkhardt & Biermann / (Statten der Kultur ; Bd. 7.)  
[1908]

 [Available at Widener Library](#) [Offsite Storage \(Swi 1319.08\)](#) >

TOP

SEND TO


DETAILS


GET IT


SHELF VIEW


TAGS


Send to


 TEXT CALL #


 PERMALINK


 E-MAIL

 CITATION

 EXPORT RIS

 EXPORT BIBTEX

 PRINT

 RESERVES LIST

Details

Title	Luzern, der Vierwaldstätter See und der St. Gotthard,
Attribution	von Hermann Kesser. Buchschmuck von E. Stiefel, Zürich.
Author / Creator	<a href="#">Kesser, Hermann, 1880-1952.</a> >
Published	Leipzig, Klinkhardt & Biermann [1908]
Description	172 p.
Language	German
Series / collection	<a href="#">Statten der Kultur, Bd. 7</a> >
Series (transcribed)	Statten der Kultur ; Bd. 7
Creation Date	[1908]
HOLLIS number	990055658810203941
Permalink	<a href="http://id.lib.harvard.edu/alma/990055658810203941/catalog">http://id.lib.harvard.edu/alma/990055658810203941/catalog</a>
Source	HVD - Core

<https://hollis.harvard.edu>



# Full-Text Digital Access

- Digitized and "born digital"
- Full-text search/retrieval
- Strengths and weaknesses of various tools and strategies



**When is "good enough" and/  
or incomplete information  
o.k.?**

**When do you need "the best"  
and/or comprehensive  
information?**



# Some Challenges to Finding "the Best" and/or Comprehensive Information

- Uneven and incomplete digitization and digital access
- Normative ways of finding and accessing information



# Uneven and incomplete digitization and digital access

- Digitization is expensive and time-consuming
- Some digitized and born digital material is openly available; other material is behind paywalls
- Many decisions along the way about what to digitize and how to make materials (whether digitized or born digital) available. These decisions are not neutral and there are many political, social, and economic factors that come into play.
- All access to information is inherently intertwined with political, social, and economic factors. This point is key both for users of information and for those creating and enabling access to information.



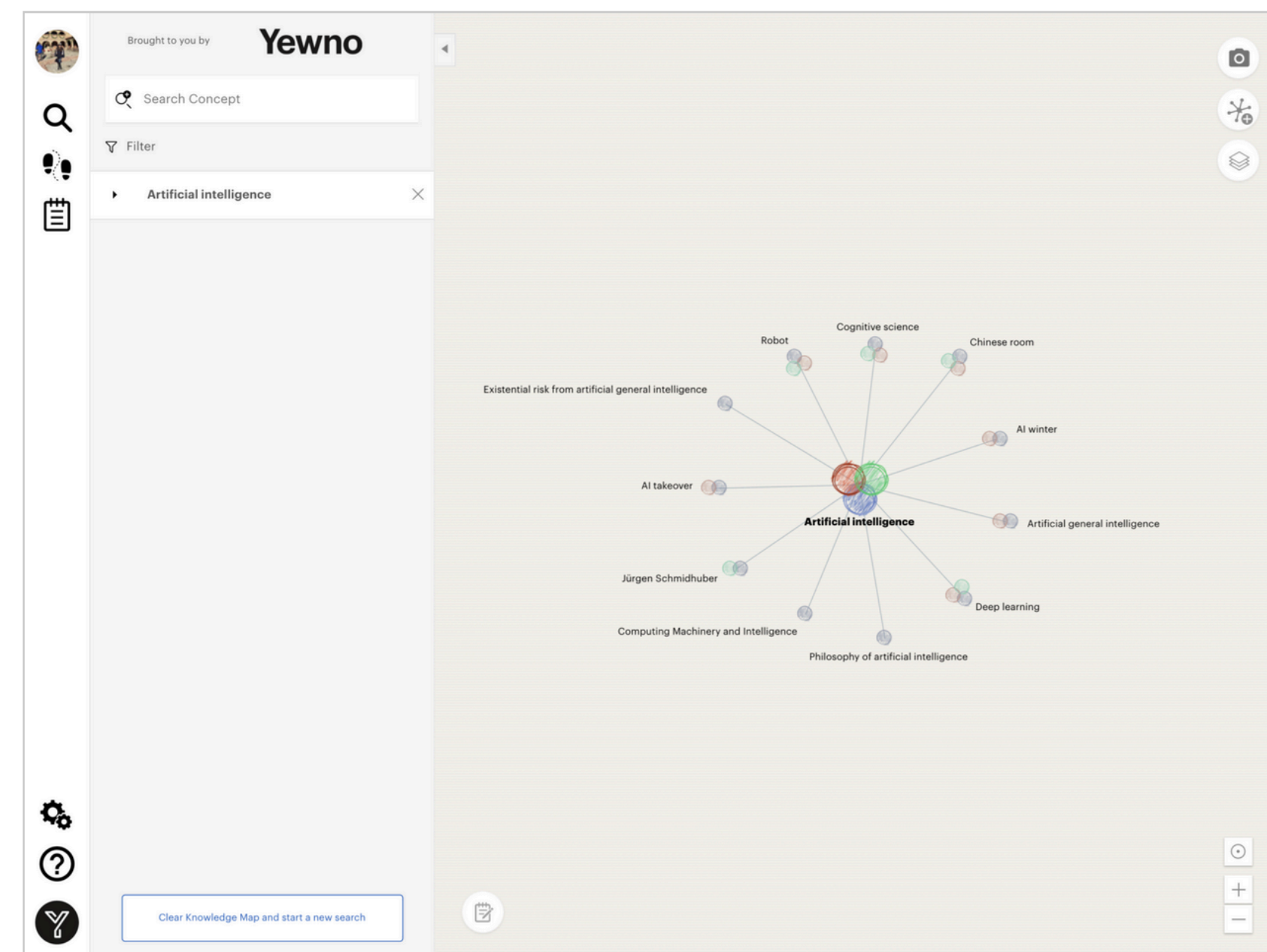
# Normative ways of finding and accessing information

- Legacies of the card catalog and access points
- Primacy of keyword search and text

## Yewno Discover

Yewno's Discover platform harnesses hundreds of millions of semantic connections and conceptual links from millions of scholarly articles, books, and databases across all academic fields. This empowers students and researchers to navigate intuitively across concepts, relationships, and disciplines, learning from resources that might have otherwise been overlooked. This not only enhances understanding and creates more impactful work, but also saves time while ensuring comprehensive and credible coverage.

[Learn More >](#)





## Digital Collections

### Featured Content



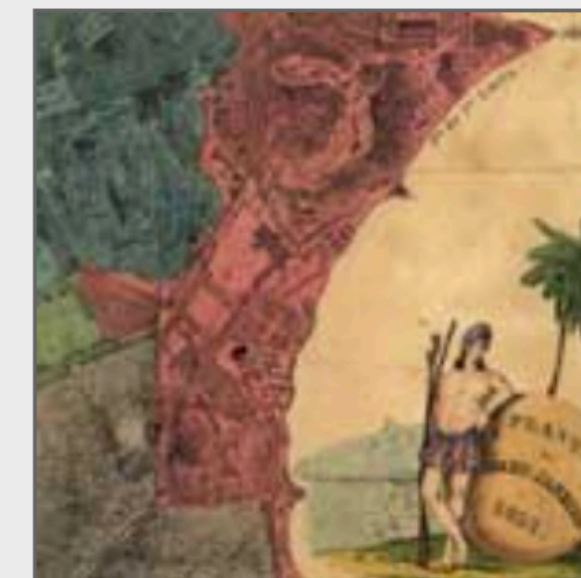
Historic American  
Buildings Survey/Historic  
American Engineering...



Chronicling America:  
Historic American  
Newspapers



Farm Security  
Administration/Office of  
War Information ...



Cities and Towns



Civil War Maps

Results: 1-40 of 335 | Refined by: Part of: Digital Collections

### Refine your results

#### Subject

American History

114

## Digital Collections

View

Gallery

Go

Sort By

Select

Go



## NATIONAL JUKEBOX



- [National Jukebox](#)
- [Browse all Recordings](#)
- [Advanced Search](#)
- [Artists](#)
- [Genres](#)
- [Jukebox Day By Day](#)
- [Playlists of Recordings](#)
- [Victrola Book of the Opera](#)
- [Contact Us](#)
- [About the National Jukebox](#)

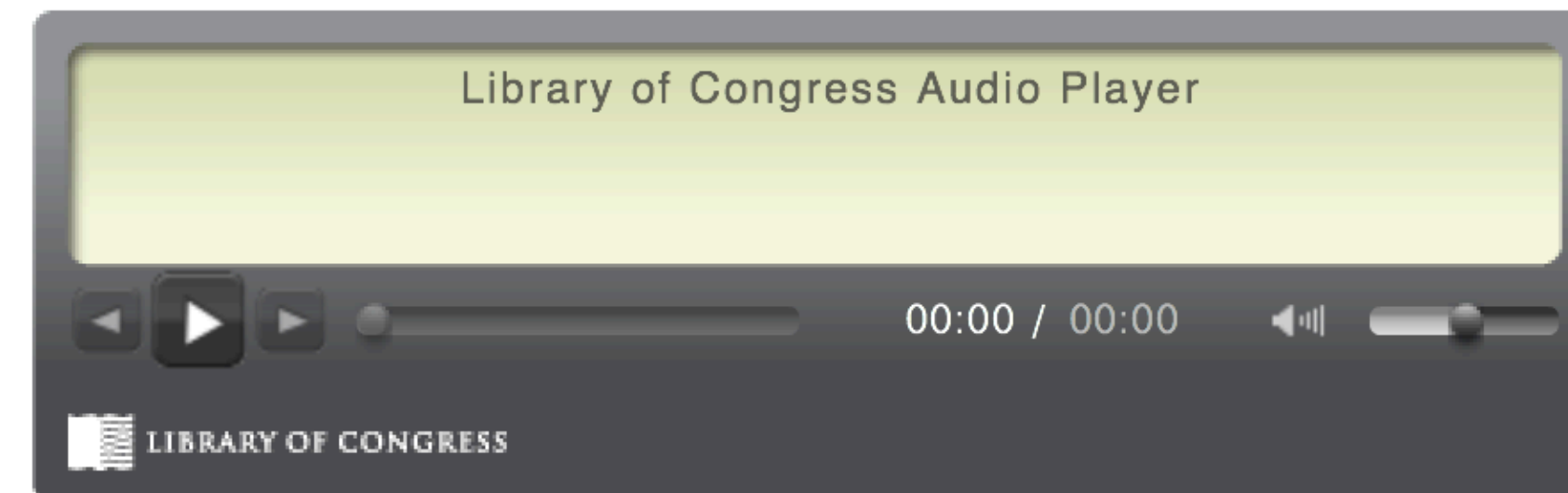
## National Jukebox

Historical Recordings from the Library of Congress

Search

Example: Artist, Song, Lyricist, Genre, Composer, Year, Place

Go

[Advanced Search »](#)[Print](#) [Subscribe](#) [Give Feedback](#)[View Playlist](#)WARNING: Historical recordings may contain offensive language. [Read the disclaimer »](#)

## Now Playing...

- Elk's reunion march
- Le parlate d'amor
- El teléfono a larga distancia
- At the jazz band ball
- Everybody's jazzin' it
- Pack up your troubles in your old kit bag and smile, smile, smile!

[View This Playlist »](#)

## About the National Jukebox

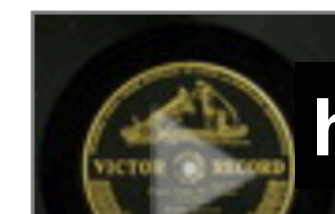
The Library of Congress presents the National Jukebox, which makes historical sound recordings available to the public free of charge. The Jukebox includes recordings from the extraordinary collections of the Library of Congress Packard Campus for Audio Visual Conservation and other contributing libraries and archives.

[Find out more about the National Jukebox project »](#)

## First Time Here?

Find out about playlisting, what "matrix" means, and other ways to [get the most from the National Jukebox](#).

## Featured Recordings and Playlists



Don't take me home

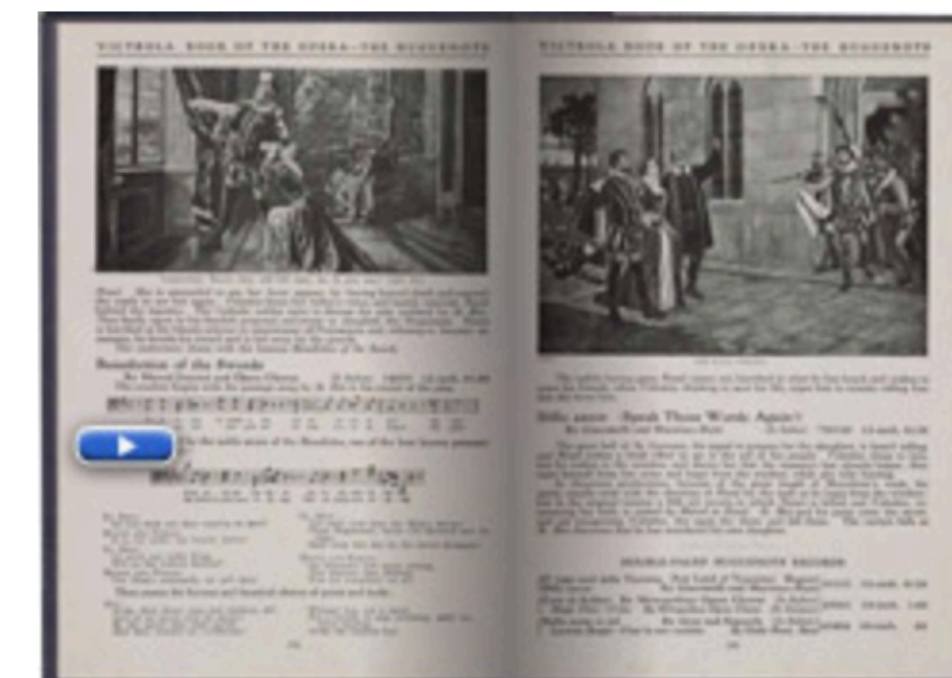
<http://www.loc.gov/jukebox/>

## Featured Playlist:

Rivers



## 1919 Victrola Book of the Opera



Chronicle of the opera with illustrations and



# What's to be done about these challenges?

- Expand imagination
- Pursue new approaches
- Bring together researchers and developers, human-knowledge and machine potential



## Project Team

### University of Nebraska-Lincoln

[Elizabeth Lorang](#), co-director

[Leen-Kiat Soh](#), co-director

[Yi Liu](#), research assistant

Chulwoo (Mike) Pack, research assistant

Delaram Rahimighazikalayeh, research assistant

### University of Virginia

[John O'Brien](#)

Andrew Barrow

[Worthy Martin](#)

### Past Contributors

Maanas Varma Datla (2014-2015)

Spencer Kulwicki (2014-2015)

Joseph Lunde (2013-2014)

Grace Thomas (2013-2015)

# Image Analysis for Archival Discovery (Aida)

Libraries, archives, museums, and other groups are creating millions and millions of digital images as we digitize the cultural record. For the most part, though, these digital images are under-utilized, and we leverage little of their information potential. At the same time, locating relevant materials in digital collections is often already a difficult endeavor and will become increasingly so as more content is digitized. The Aida team is exploring what more we can do with the millions of images that represent the digitized cultural record—particularly digital images of textual materials—and we are interested in the types of discovery that serious attention to digital images might yield.

## Code & Data

Code developed for our project is available via our GitHub [organization page](#).

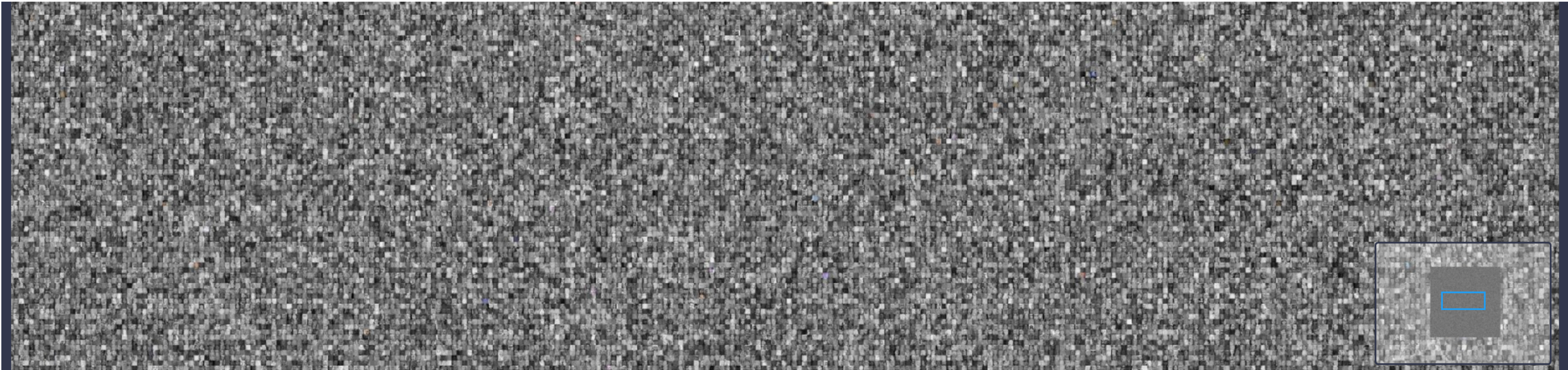
Data generated for our project is made available through appropriate data repositories, designed for long-term storage, preservation, and access. Our first data set is available via the Open Science Framework and the University of Nebraska-Lincoln Libraries Data Repository: <https://osf.io/jd74f/>


## Documents & Publications

- “Using Chronicling America’s Images to Explore Digitized Historic Newspapers & Imagine Alternative Futures,” National Digital Newspaper Program Annual Meeting, Washington, D.C., September 26, 2018
- “Patterns, Collaboration, Practice: Algorithms as Editing for Historic Periodicals,” American Antiquarian Society Spring Symposium 2018, Editorship as Collaboration: Patterns of Practice in Multi-Ethnic Periodicals, Worcester, MA, April 28, 2018

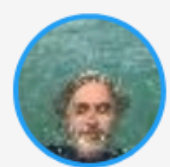
<http://projectaida.org/>





 Download

## Faces of the Tribune collection



by  
wragge

Uploaded 14 days ago  
Category [Uncategorized](#)  
Resolution 23,400 x 23,450  
Extension JPG  
File size 96.10 MB

234  
Views

0  
Likes

0  
Comments

### Comments



Write a comment ...

NO COMMENTS YET

### Description

The results of using OpenCV to detect faces in the Tribune collection of 60,000 negatives, documenting political action in Australia from the 1960s to the 1980s. See: <https://github.com/GLAM-Workbench/facial-detection> Images are courtesy of The Tribune / SEARCH Foundation, and are part of the collections of the State Library of New South Wales.

<https://easyzoom.com/image/127772>



picdescbot

@picdescbot

I'm a bot that describes random pictures from wikimedia commons using Microsoft Cognitive Services. Made by @elad3, Moderated by @elad3 and @0xabad1dea

picdescbot @picdescbot · Oct 30  
a group of people on a beach

[commons.wikimedia.org/w/index.php?cu...](#)



1 8 19

picdescbot @picdescbot · 24 Jun 2016  
a dinosaur on top of a surfboard



picdescbot @picdescbot · Oct 31  
a train traveling down train tracks near a station

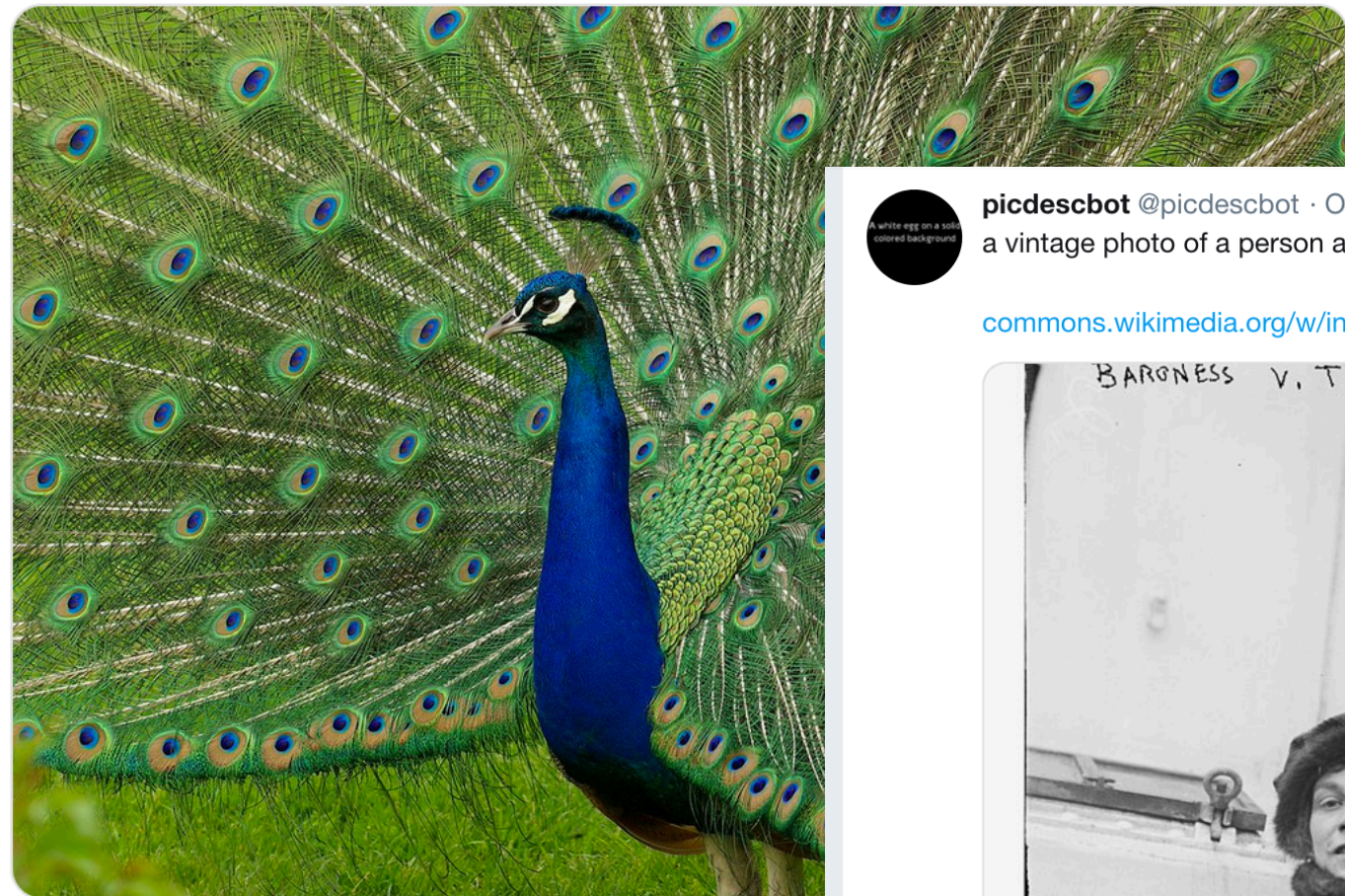
[commons.wikimedia.org/w/index.php?cu...](#)



1

picdescbot @picdescbot · 13h  
a bird flying over a field

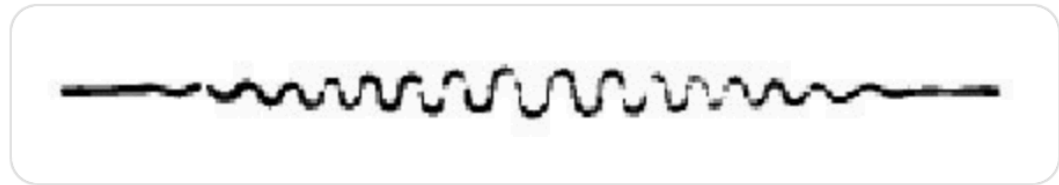
[commons.wikimedia.org/w/index.php?cu...](#)



1 7 10

picdescbot @picdescbot · 4h  
a close up of a persons face

[commons.wikimedia.org/w/index.php?cu...](#)



2 8 21

picdescbot @picdescbot · Oct 30  
a vintage photo of a person and a person posing for a picture

[commons.wikimedia.org/w/index.php?cu...](#)



9



Visuals

1819 items



484: [Taos man holding his granddaughter, facing right with adobe oven on right, and Taos Pueblo in background] [graphic].

Click to view.

A Library of Color

What is this?

<https://library-of-colors.glitch.me/>



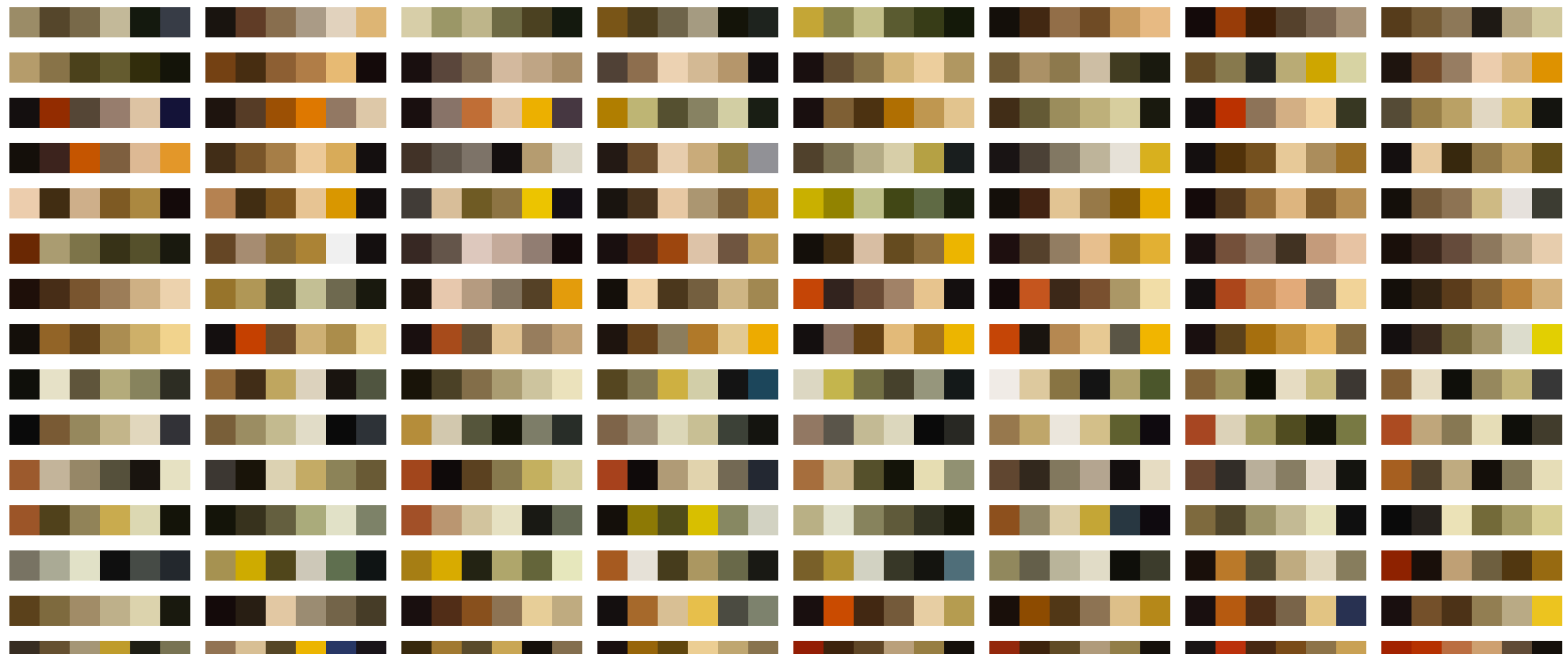
# Library of Congress Colors

The Library of Congress has thousands of digitized images online. What are the color clusters in each image, within particular collections?

Choose a collection:

World War I Posters

[Browse this collection](#)



<https://loc-colors.glitch.me/>

# LC for Robots

We hope this list of APIs, bulk downloads, and tutorials will help you begin exploring the many ways the Library of Congress provides machine-readable access to its digital collections.



## Get Data

- Bulk data for Congress.gov [bills](#), [bill status](#), and [bill summaries](#)
- [MARC records](#) - bibliographic information for most of the Library's collections. 25 million records available for exploration in UTF-8, MARC8, and XML formats.
- [Sample MARC data set](#) and [ReadMe file](#)
- [Chronicling America Bulk OCR Data](#) – text only
- [Chronicling America Bulk Data](#) – image, metadata, and OCR text batches

## APIs

- [Loc.gov JSON API](#) [↗](#) - provides data about Library of Congress digital collections. The API is in progress and subject to change.
- [Chronicling America APIs](#) - over 12 million (and growing) digitized historic newspaper pages from almost every U.S. state and territory.
- [American Archive of Public Broadcasting APIs](#) [↗](#) - tens of thousands of historic public radio and television programs are available for streaming and more content is added periodically. In addition, the website provides data records for approximately 2.5 million items inventoried by public broadcast stations for this project. Further, scholars may request access to JSON and text transcripts for the AAPB's Online Reading Room through the AAPB Transcripts Research Access service. Credentials are required to access the transcripts API and can be obtained by contacting

<https://labs.loc.gov/lc-for-robots/>

# Technical Challenges & Human Challenges