10-25-2018

Wikipedia and Libraries: Improving Metadata Through Collaboration

Eric Willey
Illinois State University, emwille@ilstu.edu

Angela Yon
Illinois State University, ayon@ilstu.edu

Follow this and additional works at: https://ir.library.illinoisstate.edu/fpml

Part of the Library and Information Science Commons

Recommended Citation
https://ir.library.illinoisstate.edu/fpml/96

This Presentation is brought to you for free and open access by the Milner Library at ISU ReD: Research and eData. It has been accepted for inclusion in Faculty and Staff Publications – Milner Library by an authorized administrator of ISU ReD: Research and eData. For more information, please contact ISUReD@ilstu.edu.
Wikipedia and Libraries; Revolution Through Collaboration

TERI EMBREY, PRITZKER MILITARY MUSEUM & LIBRARY
ERIC WILLEY & ANGELA YON, ILLINOIS STATE UNIVERSITY
Wikipedia and Libraries: Improving Metadata Through Collaboration

ANGELA YON AND ERIC WILLEY, ILLINOIS STATE UNIVERSITY
FUNDING PROVIDED BY AN ILLINOIS STATE UNIVERSITY RESEARCH GRANT
Goals

• Build a catalog that will enable users to identify authors that share the ethnicity African American

• To examine the feasibility using Wikipedia sources to identify author characteristics

• Case study to learn what resources (skills, time, tools, workflow) are necessary
Issue

• Catalog records are intended to describe the resource (book, DVD, article, sound recording, etc.)

• **NOT** to describe the characteristics of the creator

• Researchers sometimes select materials based on the life experiences of the creators rather than author/title/subject
The Library of Congress Demographic Group Terms (LCDGT)

Library of Congress Demographic Group Terms

Newspaper editors
  [occupation/field of activity]
  UF  Editors, Newspaper
  BT  Journalists

Newspersons
  USE  Journalists

Newswriters
  USE  Journalists

Ngala speakers (Lingala speakers)
  USE  Lingala speakers

Northern Chinese speakers
  USE  Chinese speakers

Northern Irish
  [national/regional]
  UF  Eireannaigh (Northern Ireland)
  BT  Britons

Cartoonists
  [occupation/field of activity]
  UF  Caricaturists
  Cartoon artists
  BT  Artists

Castellano speakers
  USE  Spanish speakers

Castillian speakers
  USE  Spanish speakers

Catalan speakers
  USE  Catalan speakers

Catalans
  [national/regional]
  UF  Catalanians
  BT  Spaniards

Catalog librarians
  USE  Catalogers

Catalogers
  [occupation/field of activity]
  UF  Cataloging librarians
  Catalogue librarians
  Cataloguers
  Library catalogers
  Library cataloguers

#OCLCARC18
Wikipedia has lots of lists...

People
- Lists of people
  - Lists of atheists
  - Lists of Christians
    - Lists of former Christians
  - Lists of bisexual people
  - Lists of celebrities
  - Lists of centenarians
  - Lists of deaths by year
  - Lists of ethnic groups
    - Lists of black people
    - Lists of African Americans
    - Lists of Celts
    - Lists of Britons
  - Lists of heroes
  - Lists of libertarians
  - Lists of Nobel Prizes
  - Lists of people by cause of death
    - Lists of people executed in the United States
    - Lists of people executed in Texas
  - Lists of people by nationality
    - Lists of people from African Union states
    - Lists of Americans
      - Lists of African Americans
      - Lists of people from Kansas

Natural and physical sciences
- Biology
  - Lists of animals
    - Lists of amphibians by region
    - Lists of birds by region
    - Lists of cats
    - Lists of dogs
    - Lists of elephants
    - Lists of extinct animals
      - Lists of extinct animals of the British Isles
    - Lists of horses
    - Lists of insects of Great Britain
    - Lists of mammals by population
    - Lists of mammals by region
    - Lists of reptiles by region
  - Lists of cities
    - Lists of cities by country
    - Lists of communes of France
    - Lists of ghost towns in Canada
    - Lists of places in Wales
    - Lists of populated places in the United States
      - Lists of places in Kansas
      - Lists of Los Angeles topics
      - Lists of New York City landmarks
      - Lists of New York City topics
      - Lists of Omaha topics
      - Lists of San Francisco topics
      - Lists of U.S. cities with large ethnic identity populations
      - Lists of U.S. cities with non-white majority populations
  - Lists of municipalities
    - Lists of neighborhoods by city
    - Lists of places
      - Lists of places by eponym

History and events
- Lists of disasters
- Lists of incidents of unrest and violence in the United States by city
- Lists of occupations
- Lists of tenants in the World Trade Center (1968–2001)
- Lists of years by topic
Library of Congress Demographic Group Terms (LCDGT)

African Americans [ethnic/cultural]
UF Afro-Americans
Americans, African
Americans, Afro-
RT Blacks

Advocates (Lawyers)
USE Lawyers
Aeronautical historians
USE Aviation historians
Afghan Americans
[ethnic/cultural]
UF Americans, Afghan
Afghans (Afghanistan)
[national/regional]
Residents of Afghanistan.
BT Asians
Afghans (Asian people)
USE Pashtuns (Asian people)
African Americans
[ethnic/cultural]
UF Afro-Americans
Americans, African
Americans, Afro-
RT Blacks
African Methodist Episcopal Church member
[religion]
UF A.M.E. Church members
AME Church members
BT Methodists
LCDGT in Name Authority Records (NARs): Why we didn’t do that

• Demographic terms are about an item’s creator, not creators themselves

• Difficulty in assigning terms that apply to some of an author’s works to their NAR

• No field in NARs for demographic terms (except gender)
Considerations

- Ethical and moral issues on author characteristics
- Ethnicity - very high standard of proof needed for this demographic category
Process

• Compared Wikipedia list of African American writers to our library holdings

• Researched entries for evidence of self-identification as African American and created citations

• Created name-title records

• Linked LCCN's/Wikipedia pages
Bland, Eleanor Taylor. †t Done wrong

010 no2017162601
040 INS †b eng †e rda †c INS-SF †d INS-SF
046 †k 1996 †2 edtf
1001 Bland, Eleanor Taylor. †t Done wrong
386 †n eth †a African Americans †2 lcdn
386 †n nat †a Americans †2 lcdn
670 Brown, W. Dale. Conversations with American writers, 2008: †b page 7 (Eleanor Taylor Bland, "Busy Lady" interview; in the black community we're finally starting)
“Link your data to other data to provide context”

Tim Berners-Lee’s “5 ★ Open Data,”

• LCCN in the authority data box on Wikipedia page

• 024 Wikidata identifier in the Name authority record
Findings - List of African-American Writers

- 387 names on list
- 25% writers with no link to name authority records
- 4% names do not have name authority records
- 70% writers in our catalog
- 91% in catalog self-identified as an African American
Next steps

• Extract bibliographic records of titles from catalog

• Add the demographic term African Americans to the bib

• Load the enhanced records in our catalog VuFind and discovery platform, EBSCO Discovery Service (EDS)

• Track the OCLC#s of enhanced records

• Repeat the process for these records in OCLC
<table>
<thead>
<tr>
<th>Title</th>
<th>Author</th>
<th>Publisher</th>
<th>Year</th>
<th>Pages</th>
<th>Size</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trombone Shorty</td>
<td>Troy Andrews</td>
<td>Abrams Books for Young Readers</td>
<td>2015</td>
<td>1</td>
<td>27 x 27 cm</td>
<td></td>
</tr>
</tbody>
</table>

- African Americans: 2
- Men: 2
- Trombonists: 2
- Jazz musicians: 2
- Americans: 2
Summary

- 25% inclusion of links to Name Authority Records
- 70% Wikidata identifier added to Name Authority Records
- 70% writers holding in the catalog, 91% self-identification as an African American
- $6 per self-identification research and citation
- 2500 bibliographic records in catalog and OCLC
- Enhancement to OCLC records offer opportunities for use by other libraries and faceted navigation
What we will do...

• On the Wikipedia page, we only will add titles if the author is the sole creator (not editor, or any work in anthologies, etc.).

• Only name-title records for works entirely created by the author, not pieces in anthologies.

• With the caveat that we might find the time to do solely authored works found in anthologies or other types of compilations.
Challenges

• Using Wikipedia Lists or Categories compiled from Wikidata?
Challenges

• Defining demographics

• Wikipedia is not necessarily complete
Brawley, Benjamin Griffith. **Benjamin Griffith Brawley** (April 22, 1882 - February 1, 1939) was a prominent **African-American** author and educator. Several of his books were considered standard college texts, including *The Negro in Literature and Art in the United States* (1918) and *New Survey of English Literature* (1925).[1]

---

1 [https://en.wikipedia.org/wiki/Benjamin_Griffith_Brawley#cite_note-Valade-1](https://en.wikipedia.org/wiki/Benjamin_Griffith_Brawley#cite_note-Valade-1)

Phillip Hayes Dean, an African American playwright best known for his biographical portrait of entertainer and activist Paul Robeson. [2]

- Wikipedia page and Wikidata
- Not on the List and not in the Category
- Our catalog includes records for his works

Potential outcomes and future opportunity

- LCDGT Information displays in our catalog and faceted search results
- Grow our documentation and research
- Build on the relationship between Wikipedia and libraries
Impact - search

• With our new VuFind 4.1 interface for our catalog, the Wikipedia page will appear on an author search

• New browsing features for subjects and authors that utilize Authority record data (Authority records are updated weekly)

https://i-share.carli.illinois.edu/vf-isu/Author/Home?author= Hughes%2C+Langston%2C+1902-1967
Impact – full circle linking

• 024 wikidata in the NAR

• LCCN in the authority data box
Resources

• “5★ Open Data.”
• “List of African-American writers.”
• Watch for more information in our IR, ISU ReD.
Thank you!

Angela Yon
Cataloging and Metadata Librarian
Milner Library at Illinois State University
ayon@ilstu.edu

Eric Willey
Interim Head of Cataloging and Acquisitions
Milner Library at Illinois State University
emwille@ilstu.edu
THANK YOU

Angela Yon
AYON@ILSTU.EDU
Illinois State University

Eric Willey
EMWILLE@ILSTU.EDU
Illinois State University