TOOLS FOR DETERMINING EQUITABLE REPRESENTATION OF WOMEN IN LIS PUBLICATIONS

AMALIA MONROE-GULICK, UNIVERSITY OF KANSAS LIBRARIES
MARLA SCHLEUDER, ROBERT J. DOLE INSTITUTE OF POLITICS UNIVERSITY OF KANSAS
OUTLINE

- Precedent Method & Method Analysis
  - Precedent Article
- Methods & Tools
  - EBSCOhost API
  - OpenRefine
  - R
- Obtaining Data
- Cleaning Data
- Predicting Gender
- Results
  - Gender & Authorship
  - Gender & Journals
- Next steps
PRECEDENT METHOD & METHOD ANALYSIS
PRECEDEINT ARTICLE


- Found underrepresentation of women authors in top political science journals (even accounting for the fewer women in the field)
- 10 "top ranked" political science journals over a 15-year period
METHODS & TOOLS
TOOLS & USES

**EBSCOhost Integration Toolkit**
- Get citation data:
  - from specific Journals,
  - for specific Year range,
  - quickly,
  - in a structured format

**OpenRefine**
- Pull in structured data (XML),
- clean it,
- and parse into more useful format (tabular)

**R**
- Combine data from all pull requests
- Run package to predict gender based on Author first name
- Create new columns derived from data in other columns
- Compile statistics/ analyze results
JOURNALS

- College and Research Libraries
- Information Technology and Libraries
- Journal of Academic Librarianship
- Journal of Documentation
- Library & Information Science Research
- Library Hi Tech
- Library Quarterly
- Library Resources and Technical Services
- Library Trends
- portal: Libraries and the Academy
PRECEDENT METHOD & CURRENT METHOD: OBTAINING DATA

- Precedent Study
  - Obtained data using web scraping techniques

- Current Study
  - Obtained data using API
GETTING STARTED WITH THE API

SearchService

The Search Service provides access to EBSCOhost powerful reference system through a variety of proprietary web methods. Using this service, you can search and browse full text databases and popular databases from leading information providers. The comprehensive databases range from general reference collections to specially designed, subject-specific databases for public, academic, medical, corporate and school libraries.

The following operations are supported. For a formal definition, please review the Service Description.

Info
The Info method allows the Data Partner to receive a list of databases available to the given profile as well as the attributes of those databases and any attached authorities and their supported format types.

Browse
The Browse method is used to view indexes such as Authors or Subject Terms on an EBSCOhost database. Note that the index availability can be different per database.

Search
The Search method is used to perform basic searches on EBSCOhost reference system. Abstracts as well as the full text for documents and articles can be retrieved using this method. When available, full text articles can also be downloaded in pdf format.

AuthoritySearch
The AuthoritySearch method is used to perform authority searches on EBSCOhost reference system.

GetClusters
Finds and returns the set of clusters that would be returned by the given search.
OPENREFINE: CLEAN UP
In order to focus our study on scholarly articles, we removed:

- Editorials
- Reviews
- Correction notices
- Letters to the editor
- Speeches
- Table of Contents

We went from over 9,000 articles down to about 6,000
PRECEDENT METHOD & CURRENT METHOD: PREDICTING 'GENDER'

- Precedent Study
  - Predicted gender using Python’s ‘genderize’ package using ‘Social Media accounts’

- Current Study
  - Predicted gender using R’s ‘gender’ package ssa method – data from Social Security Administration
Comparison of Gender Distribution in Librarianship

Sources:
- ARL 2018-2019- US Only (All) (Figure 3: Race/Ethnicity and Sex Distribution of Professional Staff In ARL University Libraries, FY 2018–2019)
R GENDER PACKAGE: SSA METHOD
RESULTS
<table>
<thead>
<tr>
<th>Sample</th>
<th>Women Authored Articles</th>
<th>Women Solo-Authored Articles</th>
<th>Articles with Women as Primary Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Articles (6,191)</td>
<td>56%</td>
<td>54%</td>
<td>58%</td>
</tr>
<tr>
<td>Articles With At least 1 U.S. Affiliated Author (3,386)</td>
<td>61%</td>
<td>58%</td>
<td>62%</td>
</tr>
<tr>
<td>Articles With Only U.S. Affiliated Authors (3,224)</td>
<td>62%</td>
<td>52%</td>
<td>63%</td>
</tr>
</tbody>
</table>

Percent of Women Authored Articles by Sample
GENDER & AUTHORSHIP, BY AUTHORSHIP TYPE

Percent of Women Authored Articles by Sample

<table>
<thead>
<tr>
<th>Sample</th>
<th>All Articles (6,191)</th>
<th>Articles With At least 1 U.S. Affiliated Author (3,386)</th>
<th>Articles With Only U.S. Affiliated Authors (3,224)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>56%</td>
<td>61%</td>
<td>62%</td>
</tr>
<tr>
<td>Solo-Authored Articles</td>
<td>54%</td>
<td>58%</td>
<td>52%</td>
</tr>
<tr>
<td>Multi-Authored Articles with Woman Primary Author</td>
<td>58%</td>
<td>62%</td>
<td>63%</td>
</tr>
</tbody>
</table>

Women in the Profession, Average: 74%
Gender & Journals, Summary

Percent of Women Authored Articles by Journal 2004-2019

- CRL: 64%
- PLA: 64%
- LRTS: 63%
- LT: 63%
- JAL: 59%
- LQ: 59%
- LISR: 58%
- ITL: 48%
- JOD: 46%
- LHT: 41%

2004-2019
GENDER & JOURNALS BY YEAR

Gender Composition of Authors per 5 Year Span and Journal
GENDER & JOURNALS: COLLEGE & RESEARCH LIBRARIES

Sources:


![Graph showing the percentage of women authors from 1987-2019](image-url)
NEXT STEPS

- Subject analysis
- Update after COVID-19 impact
CONTACT

- Amalia Monroe-Gulick, almonroe@ku.edu
- Marla Schleuder, mdschleuder@ku.edu