

Wikidata and Knowledge Graphs in Practice: Using semantic SEO to create discoverable, accessible, machine-readable definitions of the people, places, and services in global information community institutions and organizations

Montana State University (Doralyn Rossmann, Jason Clark)
London School of Economics (Helen K. Williams, Neil Stewart)

NISO Plus 2022

What's Ahead

- Research Motivation
- Outside-in & Inside-out Solutions
- Knowledge Graphs and Libraries
- Case Studies
- Research Implications

Research Motivation

“A library is a collection of materials, books or media that are easily accessible for use...”

[Wikipedia entry for "Library"](#)

“Wikidata is being used as a means of documenting and surfacing researchers, publications and research data in a number of ways. It provides an opportunity for sharing faculty scholarship on an open and accessible platform.”

[ARL White Paper on Wikidata: Opportunities and Recommendations](#)

Outside-In Inside-Out

Outside-In, Inside-Out

“***inside-out*** resources which may be unique to an institution (e.g. digitized images, research materials) where the audience is both local and external. Thinking about ***an external non-institutional audience***, and how to reach it, poses some ***new questions for the library.***”

Source: Lorcan Dempsey, [Outside-In and Inside-Out](#) (2010)

An Inside-Out Library

“an opportunity to ***pivot their expertise in organizing information outward***. “Inside-out” library services can include support for special collections, digital scholarship, scholarly communication, and data management. A key characteristic of such services is that an academic ***library takes on broader information management challenges*** at their college or university.”

Source: Mark Dahl, [Inside-Out Library Services](#) (2018)

Inside-Out

Metadata expertise

+

Contributions to external sources about
library resources (Wikidata)

Library contributes to broader
research, teaching, and learning at a
global scale

Knowledge Graphs

Knowledge Graph

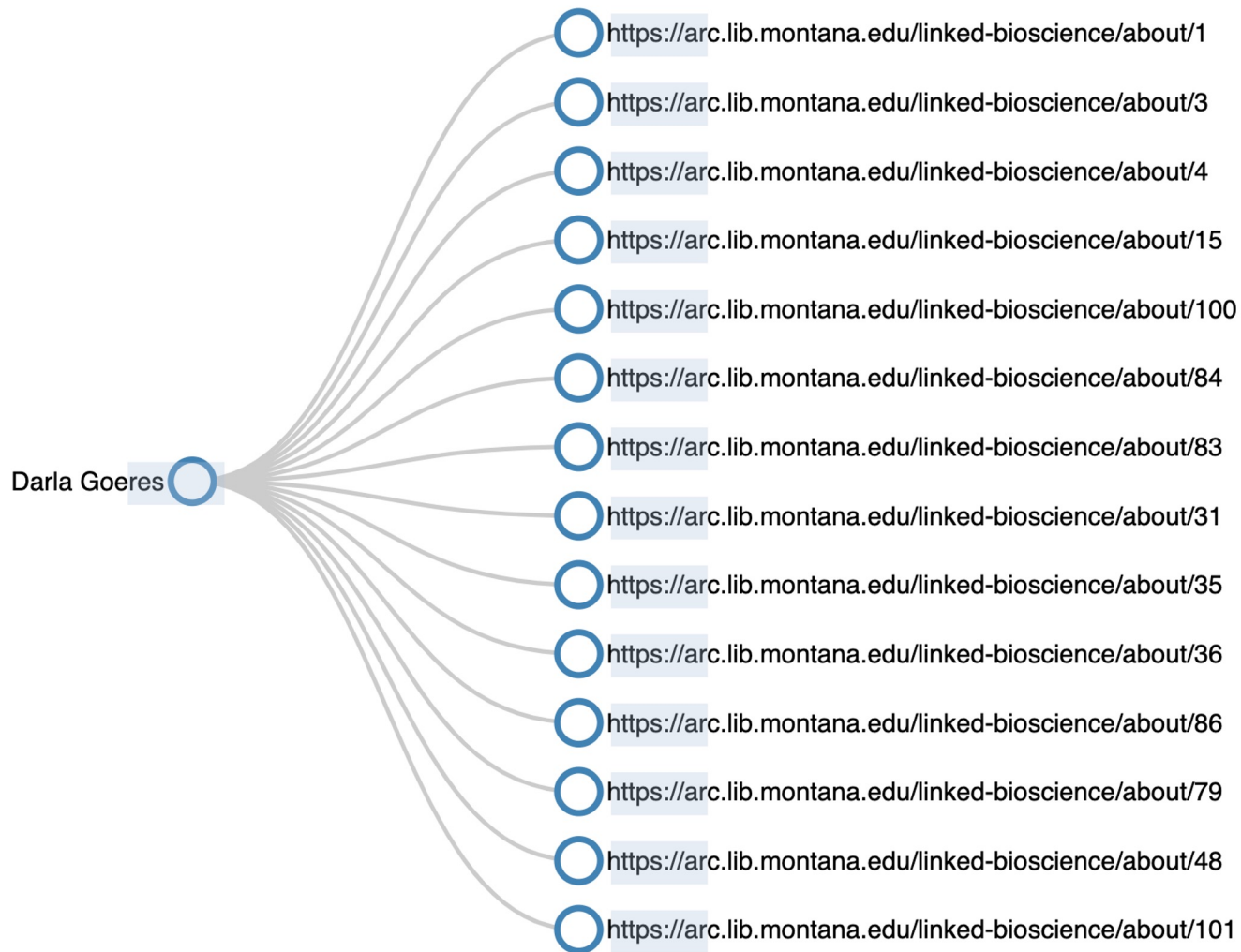
“A knowledge graph, also known as a semantic network, represents a network of real-world entities—i.e. objects, events, situations, or concepts—and illustrates the relationship between them.”

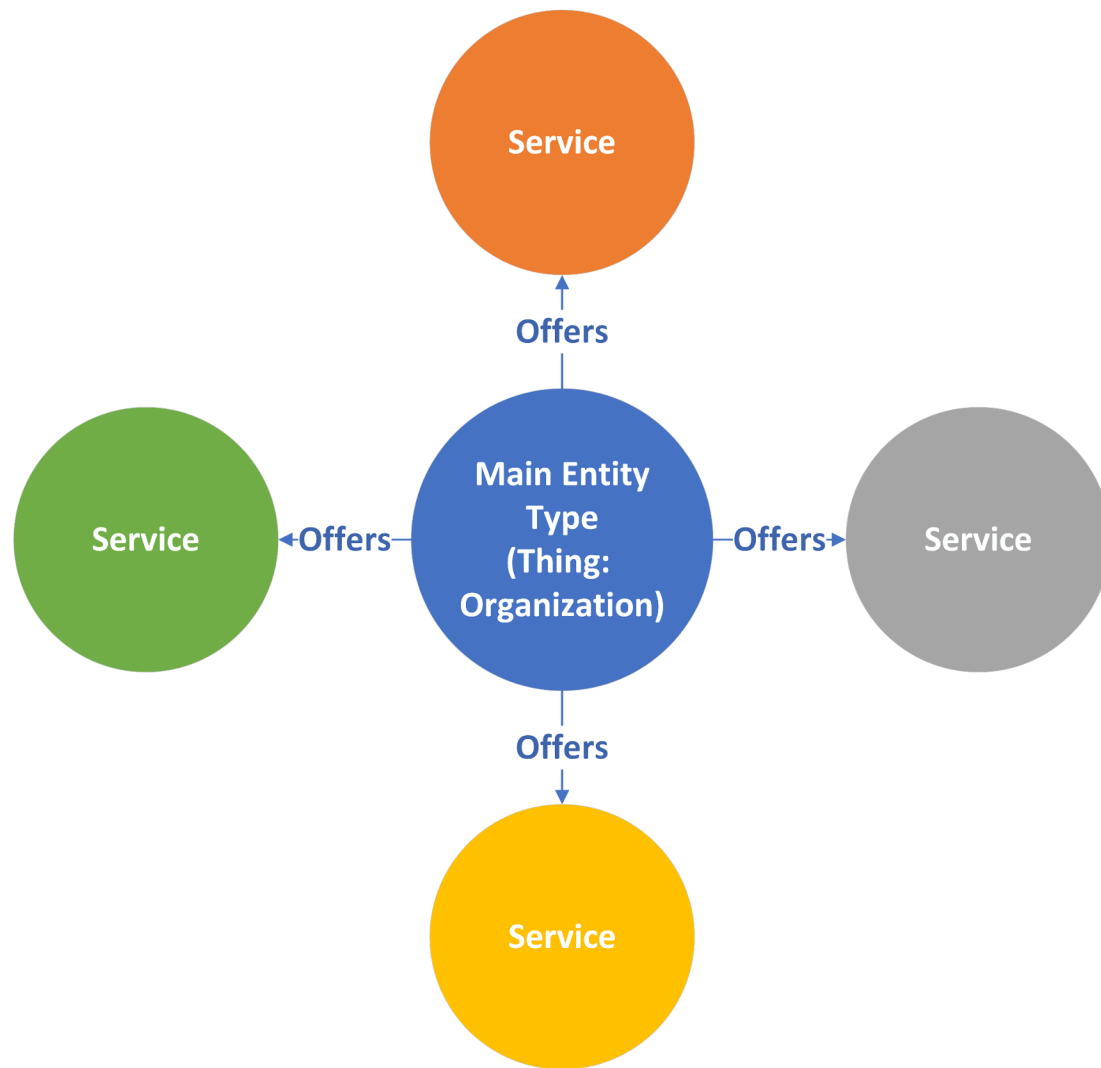
from IBM - Knowledge Graph

Graph Data Model

A connected graph of nodes and relationships with properties and labels.

*from Linked Bioscience project
Code + Demo*







Case Studies

Semantic Search Engine Optimization

Semantic search relies on a network of related entities such as contextually related concepts, ideas, people, places and things to determine what a web page is about and the “intent” which it might fulfill.

Optimization adds coding and indexing to web pages to give search engines more information, proactively.

Set up library
website as a graph of
pages

Website + hasPart + WebPages

Knowledge-graph repository
templates

Machine-readable
templates

Montana State University Library

[Find](#)[Request](#)[Services](#)[Spaces](#)[People](#)[Ask the Library](#)A detailed historical map of Montana, showing various geographical features, rivers, and towns. The map is in sepia tones with some color highlights in red and green. It includes labels for 'Gallatin's', 'Yellowstone', and 'Montana'.
CatSearch for books, articles, and more[Articles and Research Databases](#)[Journals by Title](#)[Advanced Search](#)[Hours](#)[Library Accounts](#)[News and Events](#)[Ask The Library](#)

Decision on pages for
initial library
knowledge graph

[Index \(primary landing page\)](#)

[About](#)

[Find](#)

[People](#)

[Request](#)

[Resources](#)

[Services](#)

[Spaces](#)

Metadata and Markup

Defining library people, places, and things (including our actions)

JSON-LD markup and Schema.org vocabulary

Monitoring in Bing Webmaster Tools and Google Search Console

Social Media Optimization - SMO
(Twitter and Facebook)

```

"name": "People, Staff Directory and Departments – MSU Library | Montana State University",
"description": "People, staff directory and departments at MSU Library Montana State University including staff profiles, staff skills and expertise, and cont
"url": "https://www.lib.montana.edu/people/",
"identifier": "https://www.lib.montana.edu/people/",
"mainEntity": {
  "@type": "Library",
  "@id": "#library-people",
  "name": "Montana State University (MSU) Library",
  "alternateName": "MSU Library",
  "publicAccess": true,
  "hasOfferCatalog": {
    "@type": "OfferCatalog",
    "name": "Library People and Staff Expertise",
    "itemListElement": [
      {"@type": "Offer", "itemOffered": {"@type": "Service", "name": "teaching", "sameAs": "https://www.wikidata.org/wiki/Q352842"}},
      {"@type": "Offer", "itemOffered": {"@type": "Service", "name": "research", "sameAs": "https://www.wikidata.org/wiki/Q42240"}},
      {"@type": "Offer", "itemOffered": {"@type": "Service", "name": "metadata", "sameAs": "https://www.wikidata.org/wiki/Q180160"}},
      {"@type": "Offer", "itemOffered": {"@type": "Service", "name": "user experience", "sameAs": "https://www.wikidata.org/wiki/Q1047808"}},
      {"@type": "Offer", "itemOffered": {"@type": "Service", "name": "scholarly communication", "sameAs": "https://www.wikidata.org/wiki/Q7432048"}},
      {"@type": "Offer", "itemOffered": {"@type": "Service", "name": "data management", "sameAs": "https://www.wikidata.org/wiki/Q1149776"}},
      {"@type": "Offer", "itemOffered": {"@type": "Service", "name": "collections management", "sameAs": "https://www.wikidata.org/wiki/Q17008276"}},
      {"@type": "Offer", "itemOffered": {"@type": "Service", "name": "publishing", "sameAs": "https://www.wikidata.org/wiki/Q3972943"}},
      {"@type": "Offer", "itemOffered": {"@type": "Service", "name": "archiving", "sameAs": "https://www.wikidata.org/wiki/Q166118"}},
      {"@type": "Offer", "itemOffered": {"@type": "Service", "name": "software development", "sameAs": "https://www.wikidata.org/wiki/Q638608"}},
      {"@type": "Offer", "itemOffered": {"@type": "Service", "name": "competitive intelligence", "sameAs": "https://www.wikidata.org/wiki/Q552515"}},
      {"@type": "Offer", "itemOffered": {"@type": "Service", "name": "library circulation", "sameAs": "https://www.wikidata.org/wiki/Q6542553"}},
      {"@type": "Offer", "itemOffered": {"@type": "Service", "name": "information literacy", "sameAs": "https://www.wikidata.org/wiki/Q679789"}},
      {"@type": "Offer", "itemOffered": {"@type": "Service", "name": "instructional technology", "sameAs": "https://www.wikidata.org/wiki/Q1068473"}}
    ]
  },
},

```




MSU Library
@msulibrary



Look at these fine fellas -- they're ready for a Saloon Parade in Twin Bridges, MT, July 4, 1898:

arc.lib.montana.edu/brook-0771/ite... #MontanaState

2:03 PM - 4 Jul 2013



MSU Library
@msulibrary



An MSU Track team of the past @MSUXCTrack @msubobcats! They look like winners to us!
arc.lib.montana.edu/msu-photos/ite... #MontanaState

12:57 PM - 15 Apr 2014



MSU Library



Montana State College Track Team

By MSU Library @msulibrary

Montana State College Men's Track Team with the Chemistry Building and Montana Hall in the background....

[View on web](#)



<!-- Social Media Tags -->

<meta name="twitter:title" content="Montana State College Track Team">

<meta name="twitter:description" content="Montana State College Men's Track Team with the Chemistry Building and Montana Hall in the background....">

<meta name="twitter:image:src" content="http://arc.lib.montana.edu/msu-photos/objects/parc-000247.jpg">

<meta name="twitter:url" content="http://arc.lib.montana.edu/msu-photos/item/247">

<meta name="twitter:card" content="summary_large_image">

<meta name="twitter:site" content="@msulibrary">

<meta name="twitter:creator" content="@msulibrary"> <meta property="og:title" content="Montana State College Track Team, MSU Historic Photo">

<meta property="og:description" content="Montana State College Men's Track Team with the Chemistry Building and Montana Hall in the background....">

<meta property="og:image" content="http://arc.lib.montana.edu/msu-photos/objects/parc-000247.jpg">

<meta property="og:url" content="http://arc.lib.montana.edu/msu-photos/item/247">

<meta property="og:type" content="website" />

<meta property="og:site_name" content="MSU Historic Photo - Montana State University (MSU) Library"/>


<!-- End Social Media Tags -->


Web scraping of SERPs


Bing and Google
Raw data results of search queries





[Scrapers + Crawlers \[code\]](#)
[Outputs \[data\]](#)


Analytics and Benchmarking

 main ▾ [scrapers-crawlers](#) / [outputs](#) / [bing-serp-snapshot-msu-library-2021-12-19.csv](#) Go to file ...

 **jasonclark** Initial commit Latest commit a0fde8e 12 hours ago [History](#)

 1 contributor

21 lines (21 sloc) | 7.02 KB Raw Blame    

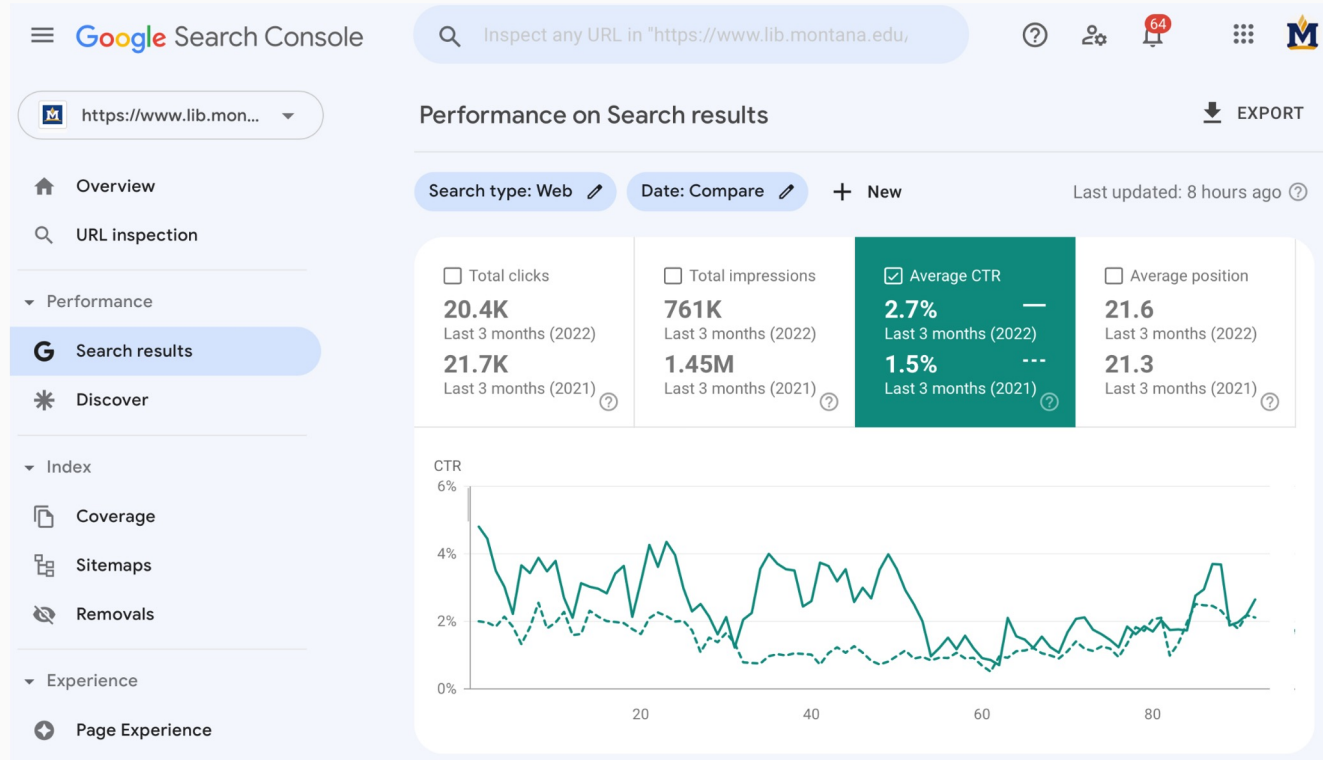
 Search this file...

	title	link	breadcrumbs
1			
2	Home MSU Libraries	https://lib.msu.edu/	https://lib.msu.edu
3	Montana State University Library - MSU Library Montana ...	https://www.lib.montana.edu/	https://www.lib.montana.edu
4	Libraries - Missouri State University	https://libraries.missouristate.edu/	https://libraries.missouristate.edu
5	Find - MSU Library Montana State University	https://www.lib.montana.edu/find/	https://www.lib.montana.edu/find
6	Home - Mississippi State University Libraries	http://library.msstate.edu/	library.msstate.edu
7	University Libraries - Waterfield Library	http://lib.murraystate.edu/	lib.murraystate.edu
8	Harry A. Sprague Library - Montclair State University	http://library.montclair.edu/	library.montclair.edu
9	Morehead State University :: Academic Excellence	https://www.moreheadstate.edu/academics/camden-carroll-library	https://www.moreheadstate.edu/academics/camden-carroll-library
10	MSU Billings Library - MSU Billings MSU Billings	https://www.msutexas.edu/library/	https://www.msutexas.edu/library
11	library.msu.ac.zw - Midlands State University Library	https://library.msu.ac.zw/	https://library.msu.ac.zw
12	Library »MSU Texas	https://msutexas.edu/library/	https://msutexas.edu/library
13	Library Services Minnesota State University, Mankato	https://library.mnsu.edu/	https://library.mnsu.edu

Google Search Console

Impressions
Coverage
Click-through Rate (CTR)

Analytics and Benchmarking

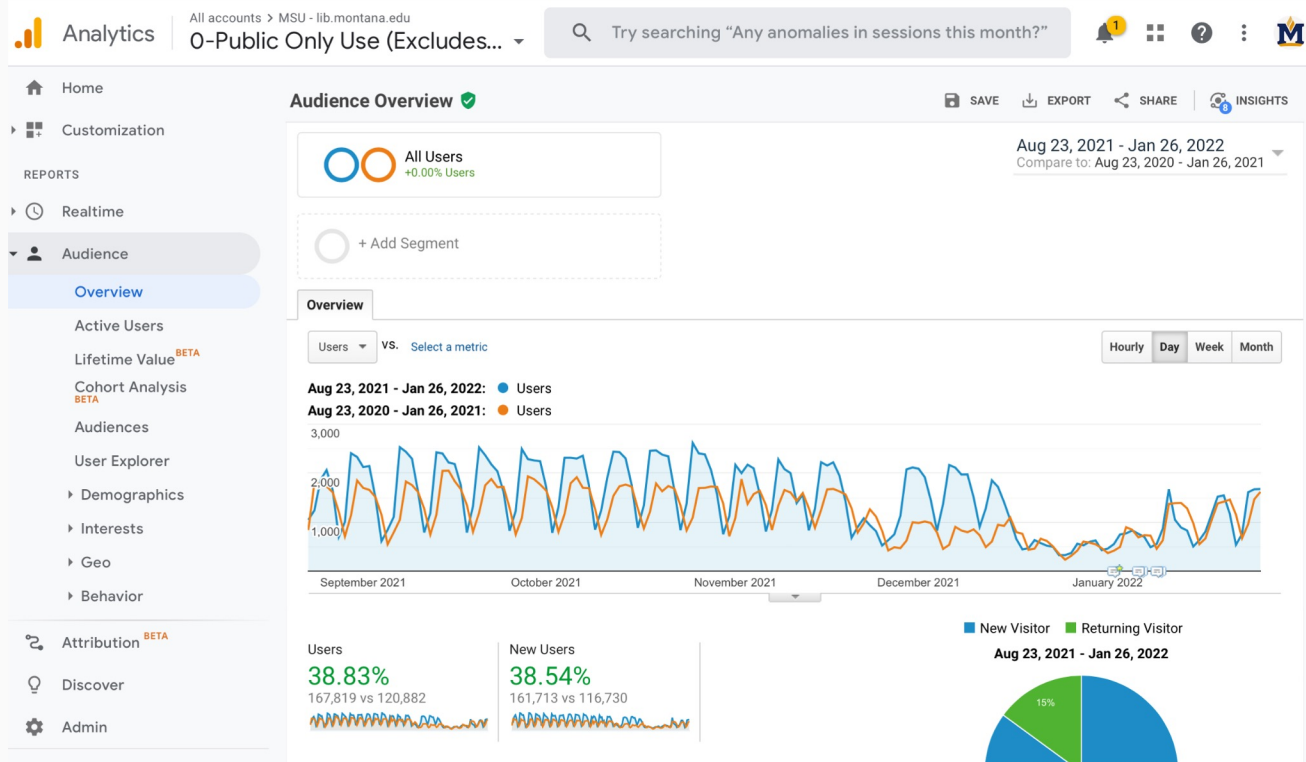


Google Analytics

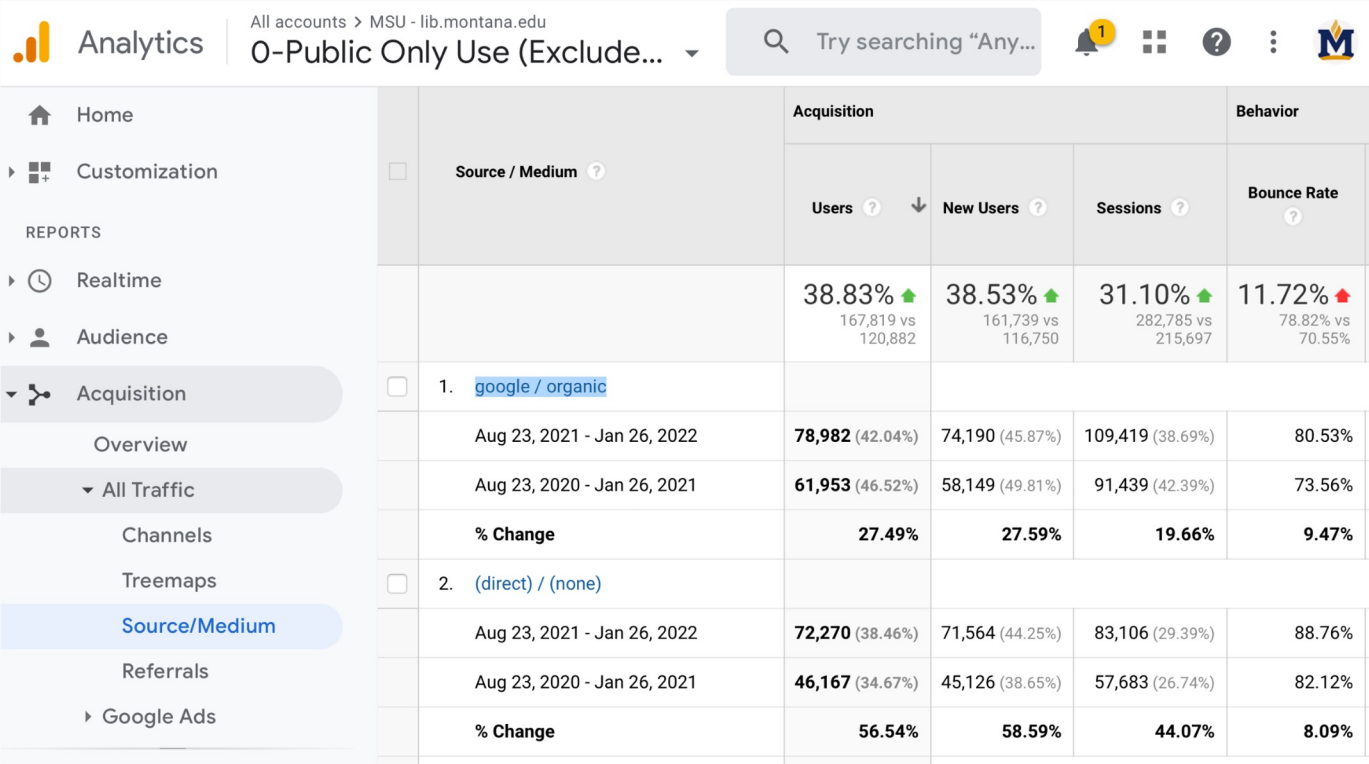
Audience
> Users

Acquisitions
> Source/Medium

Analytics and Benchmarking



Analytics and Benchmarking



Wikidata case study

LSE



Helen Williams
Metadata Manager

[@HelsKRW](https://twitter.com/HelsKRW) H.K.Williams@lse.ac.uk

NISO Plus 2022

Library

Digital shift context

Digital Scholarship & Innovation



Development and exploration of new ways in which our metadata can support research, teaching and learning

Inside out approach

Collaborative creation and management of metadata beyond the Library to a global landscape





Structure
d
database

Multilingua
l

Reusable
data

Open
Knowledgebase

Free



- Creates bridges across domains
- Impacts search engine results
- Content in Linked Open Data Ecosystem
- Identifier hub



Begin at the beginning

Barriers

*Technical *Regulatory *Articulating value



- Creating and editing libraries in Wikidata
- <https://coffeecode.net/creating-and-editing-libraries-in-wikidata.html>

British Library of Political and Economic Science (Q2371017)

has part	LSE Press	
	» 1 reference	
	Houghton Street Press	
	» 1 reference	
	Women's Library	
	» 1 reference	
	Women's Library Archives	
	» 1 reference	
	LSE Library online catalogue	
	» 1 reference	
	LSE Library Archives online catalogue	
	» 1 reference	
	LSE Digital Library	
	» 1 reference	
	LSE Research Online	
	» 1 reference	
	LSE Theses Online	
	» 1 reference	

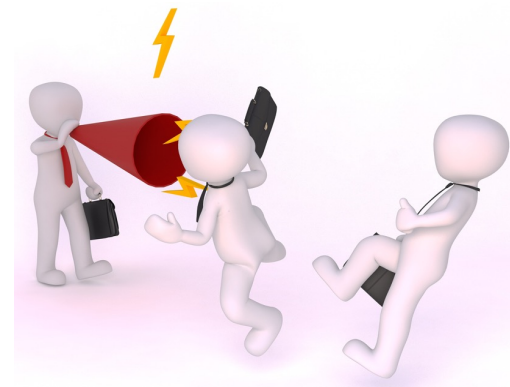


Welcome to LSE Theses Online

Welcome to LSE Theses Online, the online archive of PhD theses for the London School of Economics and Political Science. LSE Theses Online contains a partial collection of completed and examined PhD theses from doctoral candidates who have studied at LSE. Please note that not all print PhD theses have been digitised. For a full catalogue of LSE theses, please search LSE Library Search at <http://www.lse.ac.uk/library/home.aspx>.

Use the "Browse" functions above to browse for theses by year or Department. For a quick search, use the search box below. For an advanced search, click [here](#).

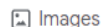
Articulate value



- Contextualise metadata by linking to related external data
- Enable new connections & discoveries



lse theses 1998



About 574,000 results (0.52 seconds)

<http://etheses.lse.ac.uk/view/year/1998.default.html>

Items where Year is 1998 - LSE Theses Online - London ...

Dalmazzo, Alberto (1998) Technological and financial factors in models of wage determination. PhD thesis, London School of Economics and Political Science. Day, ...

<http://etheses.lse.ac.uk/view/sets/LSE-GY>

Geography and Environment - Browse by Sets - LSE Theses ...

PhD thesis, London School of Economics and Political Science. ... Nava-Escudero, Cesar (1998) Local government organisation in London and Mexico City: A ...



OpenRefine

A free, open source,
powerful tool for
working with messy data

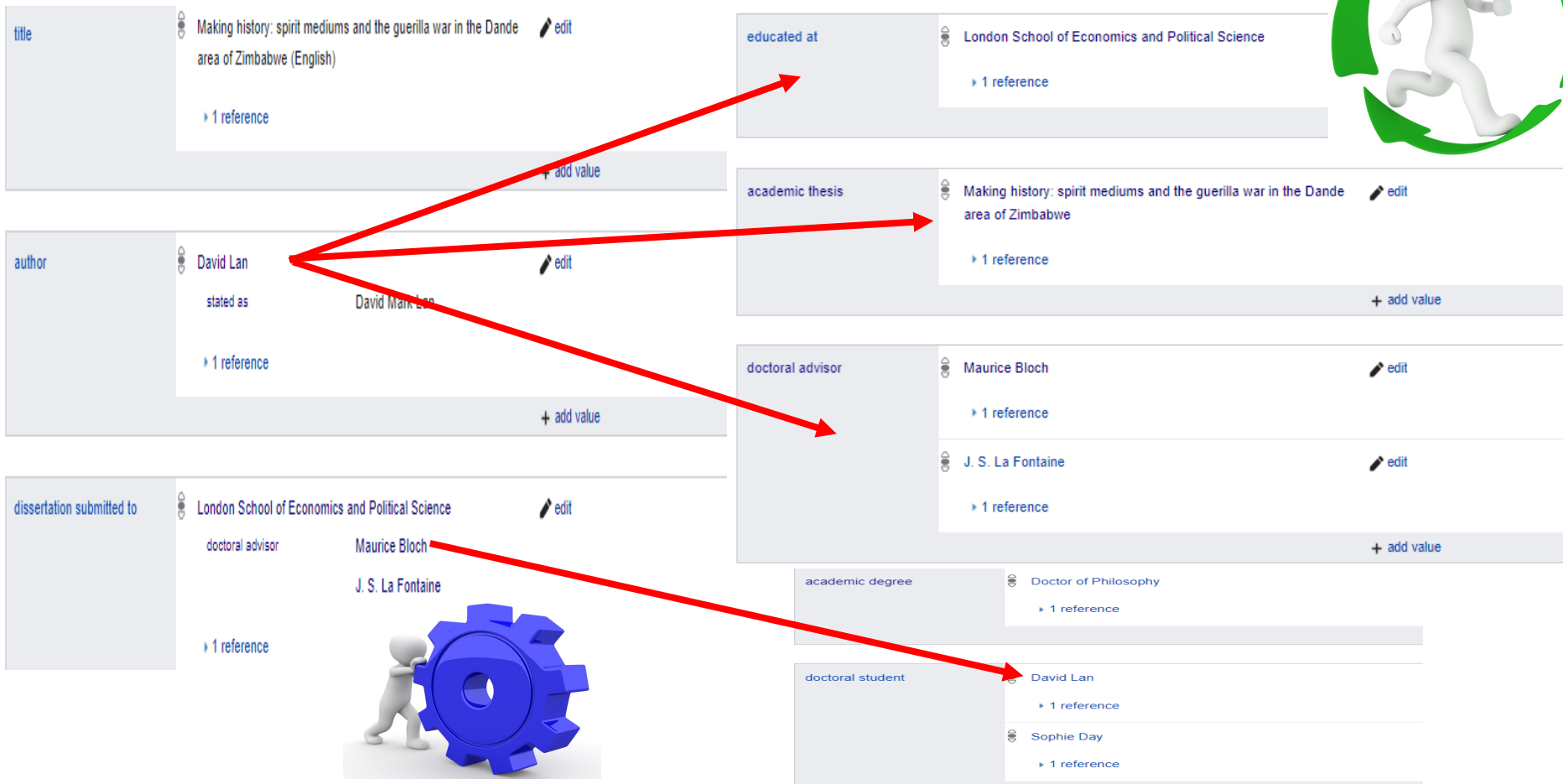


Process

- Data modelling
- Data mapping
- Data reconciliation in OpenRefine
- QuickStatements and OpenRefine for data upload
- Reconcile titles with external identifiers – EThOS, CORE, DART-EUROPE, ProQuest
- Roundtripping within Wikidata

▼ All	▼ URL	▼ Author	▼ Wikidata author id	▼ Author name string	▼ stated as	▼ Description	▼ title	▼ date	▼ pages	▼ doi	▼ Supervisor 1	▼ Supervisor 2
1.	http://etheses.lse.ac.uk/617	Hannah Abdullah Create new item Search for match		Hannah Abdullah		doctoral thesis by Hannah Abdullah	New German painting: painting, nostalgia & cultural identity in post-unification Germany Choose new match	2012	258		Nigel Dodd Choose new match	
2.	http://etheses.lse.ac.uk/379	Taha Afshar Choose new match	Q102403590			doctoral thesis by Taha Afshar	Corporate philanthropy in the UK and US: the impact of cycles, strategy and CEO succession Choose new match	2012	217		Harry Barkema Choose new match	Fei Qin
3.	http://etheses.lse.ac.uk/683	Adel Al Torafi Choose new match	Q4881548			doctoral thesis by Adel Al Torafi	Understanding the role of state identity in foreign policy decision-making: the rise of Saudi-Iranian rapprochement (1997-2009) Choose new match	2012	349		Katerina Dalacoura Create new item Search for match	

Round tripping within Wikidata



<https://w.wiki/jwZ>

Institutional theses ranking in Wikidata

count	institutionLabel
9665	Delft University of Technology
6219	University of Leeds
5652	University of Oxford
3954	London School of Economics and Political Science
3824	Universidad Nacional de La Plata
2421	Faculty of Applied Sciences
1416	Faculty of Electronic Engineering, Mathematics and Computer Science
1133	Faculty of Mechanical, Maritime and Materials Engineering
1105	Faculty of Civil Engineering and Geosciences
810	University of North Carolina at Charlotte
736	University of Washington
396	Faculty of Aerospace Engineering, Delft University of Technology
360	Faculty of Architecture and the Built Environment
304	University of Edinburgh



Wikidata Query Service

Examples

Query Builder

Help

More tools



```
1 SELECT (COUNT(?thesis) As ?count) ?institutionLabel WHERE {  
2   ?thesis wdt:P31 wd:Q187685;  
3   wdt:P4101?institution  
4   SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en" }  
5 } GROUP BY ?institutionLabel ORDER BY DESC(?count)
```

Interim analysis of reach and engagement

Downloads

- Total downloads Feb-May 2021 **14% higher** than same time period in 2020 (For comparison previous 3 years saw increase of 6.9% and decreases of 5 and 12% respectively)
- Analysis of 80 titles added to Wikidata - Downloads in the 6 months after adding titles to Wikidata on average **47% higher** than the 6 months before addition to Wikidata



Interim analysis of reach and engagement

Google Analytics for LSETO

- Primary referral source Google Scholar – approx. 40% of traffic
- Second referral source Twitter – approx. 10% of traffic
- 10 sources referring 1-6% of traffic each
- Long tail of 300 sources referring 0.x or 0.0x% of traffic each
- 6 months preceding Wikidata work - average 3.82% traffic from Wikipedia
- 6 months after Wikidata work - average **9.31%** traffic from Wikipedia
 - With most recent week of analysis **13.61%**



Etheses.lse.ac.uk on Twitter

38 mentions Feb-May 2020, increasing to 74 for same time period in 2021



Visualisation



<https://www.wiki/2767>

 Wikidata Query Service

Examples

Query Builder

Help

More tools

English

1

#defaultView:Graph

2

SELECT ?author1 ?author1Label ?rgb ?author2 ?author2Label

3

WITH {

4

SELECT ?author1 ?author2 ?rgb WHERE {

5

{ ?author1 wdt:P108 wd:Q174570 . } union { ?author1 wdt:P1416 [wdt:P361* wd:Q174570] . } union { ?author1 wdt:P69 wd:Q174570 . }

6

{ ?author2 wdt:P108 wd:Q174570 . } union { ?author2 wdt:P1416 [wdt:P361* wd:Q174570] . } union { ?author2 wdt:P69 wd:Q174570 . }

7

{ ?author1 wdt:P184 ?author2 }

8

UNION

9

{ ?author2 wdt:P185 ?author1 }

10

{ ?author1 wdt:P21 ?gender1 .

11

BIND (IF(?gender1 = wd:Q6581097, "3182BD", "E6550D") AS ?rgb)

12

}

13

} AS %result

14

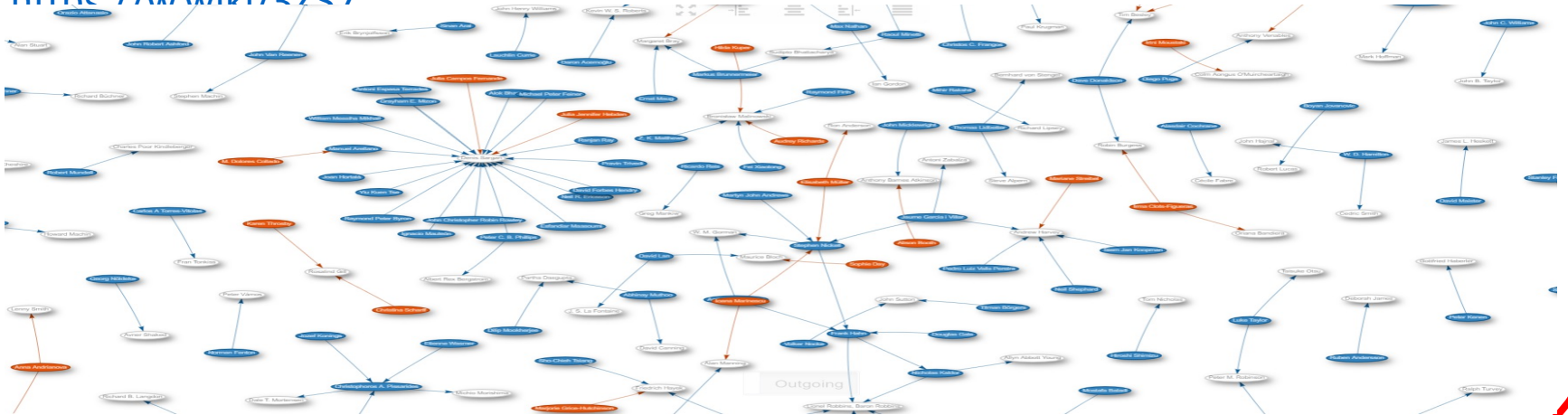
WHERE {

15

INCLUDE %result

16

SERVICE wikibase:label { bd:serviceParam wikibase:language "en,da,de,es,fr,jp,sv,ru,zh".

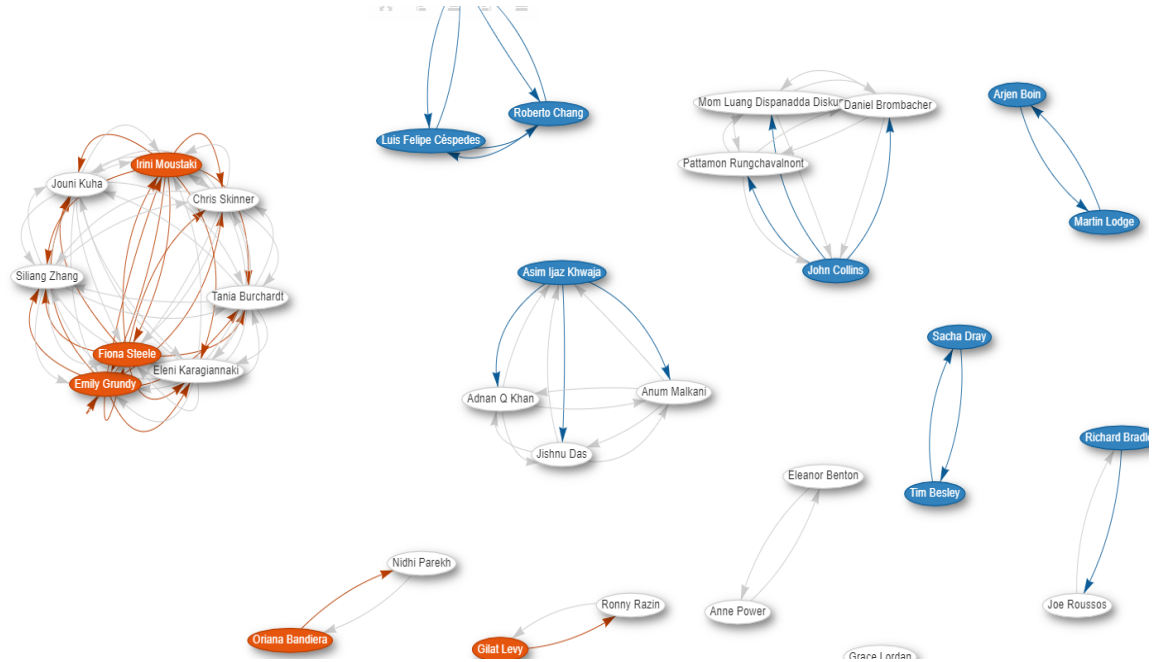




Where next?

LSE Public Policy Review author collaborations

- Extend reach of content & data unique to LSE
- Develop digital scholarship expertise



Where next?



- Open Access focus
- Special collections focus
- Digitised content focus
- Researcher focus



In conclusion



Wikidata

- Supports global & collaborative metadata
- Helped us extend reach & engagement of LSE theses
- Demonstrates role of metadata in expanding access & visibility of our libraries & their resources
- Improves the accessibility & reach of global information communities



Photo credits

Slides 2-15: CC0, [3dman-eu: Pixabay](#)

Slides 2, 5: London School of Economics and Political Science www.lse.ac.uk

Slides 3-4, 6-8, 11-12: CC0, <https://www.wikidata.org>

Slide 5: Google <https://www.google.co.uk/>

Slide 6: OpenRefine <https://openrefine.org/>



Library Services
British Library of Political
and Economic Science

Research Findings + Implications

Initial Findings

LSE

Increased downloads of LSE theses

Increased referrals from Wikipedia

Increased reach and engagement

Expands access and visibility

New opportunities to visualise the data

Initial Findings

MSU

Impressions (GSC)

Coverage (GSC)

Organic Search (GA)

Traffic Sources & Twitter Analytics -
SMO

Future Research

Applying library knowledge graph to search experiences

Additional metadata and content in Wikidata

Additional markup and descriptions to enhance definitions of library

Closing thoughts

Graph data models and sources have an impact for libraries.

- Create new understandings for humans and machines.
 - Allow information professionals to apply expertise.
 - Enhance discovery and usage.

Semantic Web Optimization and Wikidata - Tools and Getting Started

- Bing Webmaster Tools
 - <https://www.bing.com/webmasters/about>
- Google Search Console
 - <https://search.google.com/search-console>
- Google Analytics
 - <https://analytics.google.com/>
- Creating Annotations in Google Analytics
 - <https://www.lovesdata.com/blog/google-analytics-annotations>
- Structured Data Markup Validator
 - <https://validator.schema.org/>
- Twitter Analytics
 - <https://analytics.twitter.com>
- Wikidata, Wikipedia
 - <https://www.wikidata.org>, <https://www.wikipedia.org/>

References and Follow-up Resources

What is a Knowledge Graph

<https://www.ibm.com/cloud/learn/knowledge-graph>

Linked Biosciences

<https://github.com/jasonclark/linked-people-bioscience>

Connecting Researchers with Bioscience

<https://medium.com/read-write-participate/linked-bioscience-90fe4dbb3798>

General Structured Data Guidelines

<https://developers.google.com/search/docs/advanced/structured-data/sd-policies>

Knowledge Graph Structured Data

<https://github.com/msulibrary/knowledge-graph-structured-data>

Utility Tool to Generate Abouts and Mentions

<https://inlinks.net/#surl>

Works Consulted

Dempsey, Lorcan. (2010). Outside-In and Inside-Out
<https://www.lorcandempsey.net/orweblog/outside-in-and-inside-out/>

Dempsey, Lorcan. (2017). Library Collections in the Life of the User: Two Directions
<https://www.oclc.org/research/news/2017/01-26.html>

Dahl, Mark. (2018). Inside-out Library Services.
<https://doi.org/10.1108/S0732-0671201800000039003>

****Thank you****