

Critical Archiving for Linked Open Data

How can archivists deploy linked open data (LOD) to address ongoing issues of biased representation and outright erasure?

CRITLOD answers this question by investigating the potential for linked open data to improve the digital representation of Black female artists.

Finding communities in our data model

```
Wikidata Query Service [Examples] [Help]
1 #Artist and
2 #Art movement
3 #Work works in their
4 #Art location
5 SELECT ?itemLabel ?label
6 WHERE {
7
8 ?item wdt:P21 wd:Q6581072
9     wd:P272 wd:Q698954
10    wd:P56Awd:P279 wd:Q391743
11    wd:P379 ?o .
12 ?o wdt:P21 wd:Q697984 .
13
14 SERVICE wikibase:label { bd:serviceParam
15 }
```

We proposed that the key elements of our artist data model could be used to infer relationships among entities. These relationships might lead to the discovery of surprising connections among otherwise isolated entities. Building on the work of other LOD archival projects, we referred to this broad relationship as “influence.”

Using the property “has works in the collection” from our model, we graphed the relationships between artists cataloged on Wikidata. We found the existence of six clearly distinguished communities in the graph.



Further research

The next stage of research is to develop a data model of agent-to-agent properties that allows researchers to graph semantic relationships among artists. We propose that these properties extend the information most frequently available on Wikidata.

The next stage of research should also look to model other domains with similar barriers in order to evaluate whether or not this modelling process will work in other areas.

Examples of agent-to-agent heuristics	Derived from
Collaborator	Work location; has works in the collection
Peer	Educated at; award received
Associated with	Exhibited with; movement; participant

Environmental Scan

Our environmental scan found significant overlap in linked open data methods (domain modelling, edit-a-thons, and the use of Wikidata) and specific needs for Black history, art, and culture projects confronted with missing information or information scarce sources.

Notable projects: Linked Jazz Black Lunch Table & Enslaved.org.

Common needs included crowdsourcing, preservation, and specialized semantic properties.

Data collection

We took a selection of sixty four artists appearing in two exhibitions of artwork predominantly by African American women artists.

Each artist was classified by which show they appeared in or if they appeared in both shows. We also noted their presence in major authority files: Wikidata, Library of Congress, and the Getty.

23 of the artists were missing from at least one authority file, while 5 were missing from Wikidata and 4 did not appear in any authority file.

This data on prominent artists reflected the scope of missing resources of scarce description.

Shifting: African American Women Artists and the Power of Their Gaze (The David C. Driskell Center, 2017)

We Wanted a Revolution: Black Radical Women, 1965-85 (Brooklyn Museum, 2017)

Exhibited in both shows

Name	Wikidata ID
Ahuja, Mequitta	Q3306173
Amos, Emma	Q5372719

Data modelling

From selection criteria derived from the data collection process, we used Wikidata queries to model the most common properties used to describe African American female visual artists.

These two figures show our final query and visual representation on our model.

```
Wikidata Query Service [Examples] [Help]
1 #Frequency of properties
2 #for African America
3 #for visual artists
4 #no subqueries
5 SELECT ?p (?COUNT?) AS ?count
6 WHERE {
7
8 ?item wdt:P21 wd:Q6581072
9     wd:P272 wd:Q698954
10    wd:P56Awd:P279 wd:Q391743 .
11
12 SERVICE wikibase:label { bd:serviceParam
13 }
14 ORDER BY ?c
15 WHERE NOT ?p=?count
```

"Most frequent properties for African American women visual artists, from Wikidata"

What is linked open data (LOD)?

Linked open data is structured data that supports semantic queries (such as “all **singers born in 1890**”) and platform interoperability.

What is the goal of CRITLOD?

Our goal is to draw generalizable conclusions about the use of linked open data in any domain plagued by missing information, propose directions for future research, and develop methods for centering community relationships.

Where can I find linked open data on the web?

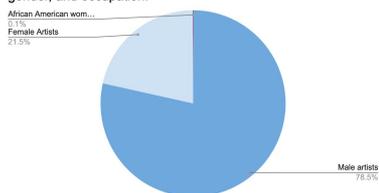
The Wikidata platform, managed by Wikimedia, has become a central hub for storing and querying linked open data.

Why investigate linked open data for archives?

Linked open data, along with Wikidata, has been suggested as a solution for information sharing and improved description. It has been successfully used by music, bibliographic, and historical archives to enrich archival description.

The problem of making up for missing information: Wikidata has only 371 records for female African American visual artists, with varying levels of description and source materials. With so little information available, are there ways that different elements of linked open data can close the descriptive gap?

"Breakdown of artists on Wikidata based on queries of ethnicity, gender, and occupation."



```
Wikidata Query Service [Examples] [Help]
1 #ones of the subset of items
2 #with the properties
3 #African American, female,
4 #and visual art or subset of visual artist
5 #ORDER GROUPED BY ?label
6 #USE
7
8 ?item wdt:P21 wd:Q6581072
9     wd:P272 wd:Q698954
10    wd:P56Awd:P279 wd:Q391743 .
11
12 SERVICE wikibase:label { bd:serviceParam
13 }
14 ORDER BY ?l
15 WHERE NOT ?l=?count
```

- Properties that emphasize relationships between entities leverage interoperability to facilitate discovery across platforms.
- Relationally minded data models enrich the efforts of volunteer and crowd driven efforts to archive or enrich information about their communities.
- Linked open data has the potential to raise the visibility of entities and materials that otherwise have little associated information.