Tracking Looted Art with Graphs: A Case Study
Laurel Zuckerman

1 Independent researcher

1 Introduction

Art looting networks operate on many levels, many of them hidden, over long periods of time. Documenting and visualising the connections - and linking them across disparate sources of information - are tasks that graphs are particularly well suited for. This paper describes efforts to track Nazi-art looting networks using the native graph functions of Wikidata enhanced by federated queries.

1.1 Background

Between 1933 and 1945, Nazi Germany operated the greatest looting campaign in history. State-orchestrated theft of Jewish property started in Germany and spread to Austria, Poland, France, the Netherlands, Italy, Belgium, Hungary, and every other country controlled by Hitler's Third Reich. Expropriations were often disguised under a legal veneer (Kreutzmüller), a habit of dissimulation continued to the present day by subsequent generations fearful of restitution issues.

This study builds on advances in three key areas: Holocaust research into the relation between mass plunder and mass murder1; digitized collections databases2; and linked data, whose possibilities for tracking looted art were suggested by Fink, Szekely and Knoblock already in 1999 (Fink) —more than a decade before most of the tools we rely on today were invented.3 Mindful of Smithies’ critical analysis of sustainability issues in innovative new digital projects, (Smithies), this study attempts to achieve its goals by exploiting more fully services that already exist, rather than creating new ones. It first explores leveraging existing data services and then describes implementation. As Wikidata is a crowdsourced wiki built on the shared contributions of diverse individuals, care was taken to respect existing rules and constraints, using standard Wikidata identifiers and properties to build out the dataset, connecting entities both internally and externally via authorities. Results are presented in terms of progress

1 "From the end of World War II until five years ago the Holocaust was primarily viewed as the greatest murder in history. And indeed it was. But since the spring of 1996 it has become ever more apparent that the Holocaust was also the greatest robbery in history."- Greg Bradsher, National Archives (Bradsher)

2 Museums and archives have made tremendous progress. (Commission) Holocaust related projects include Yad Vashem’s victims names database, the Commission for Looted Art in Europe aggregator site, lootedart.com, the German Historical Museum’s Munich Collecting Point archives and the German Lost Art Foundation’s Lostart.de and Proveana).

made in analysing art looting networks using Sparql queries, lessons learned, and issues raised.

2 Leveraging what exists

How to leverage existing data services and tools to track Nazi art looting networks — without building custom services, new ontologies or data silos?

The backbone of our approach is Wikidata, an open community of over 11,000 contributors and a knowledge graph of 1.34 billion statements, supported by the Wikimedia Foundation.(Pinschter)

Tracking looted art means crossing boundaries - geographic, linguistic, temporal, scholarly. It is as much a study of crime networks(Petropoulos)(Calvó-Armengoi and Zenou) as it is of Jewish history, the art market, genocide, genealogy, corporate ownership, provenance, or the dissemination of false information. Aspects of all these domains are present in both looted art and in Wikidata.

2.1 Gap analysis

How well we can visualize the connections using existing resources depends on how complete the graph is with respect to our subject. To assess the state of completeness of the knowledge graph we performed gap analysis using Sparql queries to identify red flag names, persecuted art collectors, artworks, and other entities to be enhanced or created, constructing dashboards to track progress.

In 1946, for example, the OSS Art Looting Investigation Unit (ALIU) investigated Nazi art looting networks and published an annotated list of a thousand ”Red Flag Names”, with their key connections.(ALIU) This is important information for tracking looted art.

How many of these Red Flag Names have entities in Wikidata? To find out we run a simple Wikidata Sparql Query:

```
SELECT ?item ?itemLabel
WHERE
{
  ?item wdt:P1840 wd:Q30335959. #Investigated by(P1840) ALIU(Q30335959)
  SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE]" . }
}
```

The answer (on November 20, 2021) was 178.

How well are these ”ALIU Red Flag Names” referenced in Wikidata, with links to authority files and external identifiers? To find out we run another query and learn that 119 Red Flag Names are linked to Viaf IDs, 104 to Worldcat, 86 to GND. They are linked to over 200 different identifiers in Wikidata, including however, only 18 in Proveana, which we would expect to have many more. Later, we will also learn about their links to businesses, partners, things they owned, and family networks. All useful information for tracking looted art.

---

4 short link to query “https://w.wiki/4Nai”
5 short link to query “https://w.wiki/4Nam”
Table 1: Dashboard. Red Flag Names and Artworks in Wikidata.

<table>
<thead>
<tr>
<th>Description</th>
<th>in Wikidata</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALIU Red Flag Names</td>
<td>172 (out of 1000)</td>
</tr>
<tr>
<td>Nazi art collectors dealers, historians</td>
<td>134</td>
</tr>
<tr>
<td>Art collections</td>
<td>1000+ collections</td>
</tr>
<tr>
<td>Paintings</td>
<td>405,700+</td>
</tr>
<tr>
<td>Paintings created before 1945</td>
<td>30,000</td>
</tr>
<tr>
<td>Art collectors who died in Nazi camps</td>
<td>39 in Wikidata</td>
</tr>
<tr>
<td>Art collectors who fled as refugees</td>
<td>?</td>
</tr>
</tbody>
</table>

With this we begin to build a dashboard. The number of artworks in Wikidata, particularly those created before 1945, is also important. Again, information about the completeness of the knowledge graph can be obtained by Wikidata queries.

2.2 Dashboard

It is, one can see, more difficult to obtain clear information about the number of art collectors who were persecuted by the Nazis because there is not a specific property for this in Wikidata. One is obliged to look in several ways and piece together the information. Proveana, the German Lost Art Foundation’s site specializing in provenance, lists 364 persecuted people. (Proveana) France’s ERR database of looted art lists hundreds more just for France. Restitution claims mention the names of the persecuted collectors. Wikipedia and Dbpedia have categories like „Jewish art collectors“, „German Jews who died in the Holocaust“ and „Jewish emigrants from Nazi Germany to the United States“ which can be helpful for queries, but only for people with Wikipedia pages. The effort of trying to put together a dashboard highlights issues with the data that need to be addressed. It also alerts us to potential incoherencies with related knowledge graphs. The low number of Red Flag Names with Proveana IDs, for example, suggests that either Wikidata has not been updated yet, or that Proveana is missing Red Flag Names, or that they are not being coded the same way.

Information about looters, victims and artworks, while incomplete in Wikidata, was nevertheless sufficient to begin to conceive strategies for exploiting the graph for looted art.

3 Implementation

3.1 Identifying key properties

We want to know which artworks connect to the Holocaust. The table below summarizes key properties (edges) that help us to make these connections. (Figure 1).6

Every one of the properties in Figure 1, added to Wikidata by different people at different times for different reasons, solidifies and expands the knowledge graph for Nazi looted art.

---

6 “See table of key properties for tracking looted art”
3.2 Using Sparql Queries to Track looted art

The next task was to adapt Sparql queries from Wikipedia’s large library for the task of tracking looted art networks, objects and events. Useful queries for data analysis and visualization were developed. They are listed in the attached spreadsheet with short links to executable Wikidata queries. Chains of nested queries, if one has the skill to develop them, are of special interest. As are federated queries with Wikibase projects like the Jewish Digital Cultural Recovery Project: JDCRP. Results, summarized below, were promising.

4 Results

4.1 Wikidata queries help track looted art networks:

Sparql queries showed promise for the following tasks:

**Interactive Data Visualizations**
- revealing Red Flag Looting Networks and adjacent people
- making dynastic art dealing networks visible
- revealing the fates of Jewish art collectors and their families with information about their collections, emigration, Nazi transports, camps, death, and events such as Aryanisation, forced sales, and claims for restitution
- detecting conflict of interest in catalogs raisonné written or published by Red Flag, Red Flag adjacent or Nazi people
- uncovering hitherto undetected intermediate networks (between Red Flags and current possessors)

---

7 “Sparql Queries for Tracking Looted Art”
8 “Jewish Digital Cultural Recovery Project”
• linking artworks to publications and authors attesting to their ownership history

• tracking itineraries of artworks and owners starting from a specific event, like the “Sonderbund Exhibition of 1912”.

• tracking artworks through long chains of connections (for example, 1) collector in possession of looted art; 2) dealers who supplied the collector, 3) their networks, 4) other artworks that passed through same networks (These are nested Sparql queries.)

• transforming specific information sealed inside news articles and scholarly texts to linked data that 1) transcends language barriers 2) connects to entities outside its narrow focus and 3) is both human and machine readable and 4) can be visualized used Wikidata properties depicts, depicted by, topic, described by source, for news articles, legal documents, scholarly articles, and books).

• potential for visualizing complex art crime networks mixing art dealing families, collectors, museums, donors, and art historians that sell via connected intermediaries over long periods of times

• placing art changes of ownership on the timeline of Holocaust events

• potential for tracking citation networks of false ownership histories (Example, false provenance in Artwork1, attested in source1, authored by Author1, leads to several queries: What other artworks have provenances attested in Source1 or written by Author1? Who is citing Source1 or Author1 as if it were a reliable source? How are Author1 and the person citing Author1 are connected?)

Resource for institutions and persons linking to Wikidata via API or OpenRefine

• providing crucial contextual information (Nazi, Red Flag, Jewish refugees or Holocaust victim) to museums that link to Wikidata for information about artwork owners

• providing crucial contextual information for researchers using OpenRefine for entity linking and enrichment

• connecting to custom Wikibase projects like the Jewish Digital Cultural Recovery Project (JDCRP) via Federated queries, enriching both.

• monitoring related digital looted art projects for completeness and accuracy

• iterative data enrichment pipelines, using artwork ids in Wikidata, to link out to provenance information, and to update Wikidata with “owned by” and “owner of” information

• possibility for bridging data silos, separating information about Nazi-looted art recovered by the Allies from information about art in German museums today (via LostArtId and Museum artwork ID),
4.2 Observations

Sequences of nested queries, in which the results of a first query provide the input for a second query, which in turn provides the input for a third query, offer a powerful tool for exploring and exposing the complex collaborative networks characteristic of Nazi looted art. Visualization tools developed both in and around Wikidata, like Metaphacts\(^9\), Entitree\(^10\) and the Panama Paper’s Linkurious\(^11\), offer additional functionalities for navigating complex connections. (Figure 2) Our approach is to surf on top of these achievements, adding targeted Wikidata Queries to the mix. Building dashboards, though frustrating, is highly recommended, not only for spotting gaps and tracking progress, but also for setting priorities to improve data quality in an iterative process. Disambiguation is crucial for all entities connected to looted art. Clarification of labels, descriptions and addition of external identifiers provide immediate benefits for the entire community. Wikidata’s role as a central hub for authorities makes it especially useful. External identifiers crucial for tracking looted art have recently been added by the Wikidata community to the benefit of all. If, reciprocally, other institutions were to systematically add Wikidata Qids to their own entities, progress would be turbocharged. Of special promise for tracking looted art is the Wikidata Wikiproject Provenance (“an effort to get every item about an artwork described by ownership and/or location since its inception on Wikidata”). All these developments suggest that the open platform and crowdsourced process offer significant advantages, even in the highly specific domain of looted art.

4.3 Issues to address

- Though growing, the knowledge graph is still incomplete, particularly concerning Jewish art collectors and their collections. (Zuckerman)

---

\(^9\) “Metapactory demo system based on the open Wikidata knowledge graph!”

\(^10\) “EntiT ree Visualize connected Wikidata items on a dynamic, navigable tree diagram. Discover properties of People, Organizations and Events with a direct link to Wikipedia Articles.”

\(^11\) “PANAMA PAPERS: HOW LINKURIOUS ENABLES ICIJ TO INVESTIGATE THE MASSIVE MOSSACK FONSECA LEAKS”
• Dealers are often intermediaries, not owners. In relying on “owned by” one is limited in representing important connections.

• Art provenances have many versions, each with its own author, publisher, date, text, and series of statements. How to represent multiple versions of provenance texts in Wikidata or in Ci-DOC CRM, or to accurately code as triples the evolutions of the discrete statements that compose them has not yet been resolved.

• One of the distinguishing characteristics of looted art is deliberately falsified provenance.\(^{12}\)(Thomkins) There is, at present, no status like “retracted” or “proven false” for a provenance text in Wikidata. Nor is there a property for “provenance text”.

• Not all looted artworks are ‘notable’ enough to fit within the scope of Sum of All Paintings, and not all looted artworks are paintings.

• Some museums publish no provenance information or isolate provenance information in separate databases or PDF files.

• Time outs and limits on processing power mean that Sparql queries that are theoretically possible might not be practical without assistance from Sparql mavens.

• Wikidata may in the future face capacity issues, and options floated at the 2021 Wikidatacon involved removing external identifiers or scholarly articles in case of emergency.(Pham)

5 Conclusion

As the realization has grown in Holocaust scholarship of the connection between dispossession and murder, so have the digital tools available to study it. This paper concludes that, for studying Nazi looted art, leveraging existing graph structures has deep advantages in terms of the richness, permanence, connectivity, evolution and reuse of the data, and that Wikidata is well suited to this task.

References


---

\(^{12}\) Max Silberberg (Q1913457), Julius Priester (Q57241365), Alfred Flechtheim (Q74128), Ismar Littman (Q43197037), Friedrich Gutmann (Q5503811), Arthur Feldmann (Q63074029), and Fritz Grunbaum (Q87857) are just a few of the many Jewish art collectors whose histories of ownership were inaccurately represented in the provenances of artworks published by museums and auction houses.


