

Leveraging Wikidata to Enhance Authority Records in the EHRI Portal

Nancy Cooley

Abstract

This paper will discuss how the European Holocaust Research Infrastructure (EHRI) leveraged Wikidata to expand and enhance its authority records for Holocaust-era camps and ghettos. As an international project integrating descriptions for Holocaust-related archival material held in institutions across the world, EHRI faces challenges in developing name authorities that can be efficiently and effectively applied across multiple languages and various metadata standards. The integration of the EHRI name authorities for camps and ghettos into Wikidata demonstrates how the linked open data capacities of Wikidata can be used to enhance local metadata and expand the reach of that data.

Introduction

In the spring of 2017, the European Holocaust Research Infrastructure (EHRI) began a project to expand and enhance its authority records for Holocaust-era camps and ghettos. As a research infrastructure integrating archival records from across the world, the EHRI portal poses many challenges to developing a functioning name authority file (NAF) that provides points of access across languages and metadata standards. In recent years some arbiters of authority control have turned to Wikidata as a solution to the problem of siloed ontologies that frequently share authoritative names for the same entities. The linked open data capacities of Wikidata allow it to function as a central access point, bridging these NAFs and increasing the connections between sources of knowledge. This paper will discuss how EHRI was able to leverage Wikidata as a data integration platform, enhancing its authority records for camps and ghettos while also enriching Wikidata with validated information on the Holocaust.

Background

Started in 2010 with funding from the European Commission, EHRI functions as a digital infrastructure integrating descriptions for Holocaust-related archival material held in institutions across the world and making these resources more accessible through an online platform, the EHRI portal. EHRI also serves as a human network, connecting Holocaust researchers, archivists, and digital humanists through conferences, workshops, and online courses (About EHRI, 2010). The mandate for the EHRI project emerged in response to logistical challenges facing Holocaust research, specifically the fragmentation and geographical dispersal of archival sources documenting the Holocaust (Anderson and Blanke, 2013). As of January 2017, the EHRI portal includes descriptive records for archival materials from more than 1,800 institutions in 51 countries, providing access to more than 200,000 archival descriptions (Factsheet EHRI, 2017). The institutions represented in EHRI include the libraries, archives, and museums of various cities, counties, countries, ethnic groups, universities, and organizations.

With the diversity of organizations contributing records comes a range of descriptive standards in varying languages (and alphabets) uniquely applied as access points in records according to the practices of each institution. EHRI project leaders responded to this challenge in the first phase of the initiative (EHRI-1) by developing a multilingual thesaurus comprised of a hierarchical list of subject terms and five name-authority lists.¹ Relevant to this paper are the two name-authority lists for Holocaust-era camps and Holocaust-era ghettos. Originally based on camps identified by the International Tracing Service and ghettos listed in *The Yad Vashem Encyclopedia of the Ghettos During the Holocaust* (Miron and Shulhani, 2009), these authorities had no additional descriptive information in their records and demonstrated few hierarchical relationships.

With the second phase of EHRI (EHRI-2), project leaders sought to expand and enhance the camps and ghettos name authorities with descriptive information for each entry, including alternative names, geographic locations, and relationships between entries (Rodriguez et al., 2016). Considering there were approximately 1,109 ghettos and 2,055 camps listed in the authorities, manual addition of descriptive data for each authority record was not realistic. It was deemed essential for EHRI-2 to deploy an integration platform that incorporates numerous sources, allows for multilingual alternative name labels, structures various descriptive elements, supports collaborative work, and generates linked open data. Wikidata emerged as a promising tool that fulfilled the project leaders' requirements for an integration platform and provided the additional appeal of being globally known and highly accessible.

Wikidata in brief

Wikidata was started in 2012 as a central storage repository for the Wikimedia project, structuring data from Wikipedia, Wikimedia Commons, Wikivoyage, Wikisource, and other Wikimedia endeavors. It has also expanded to include data from other encyclopedias, catalogs, and NAFs. Under the Creative Commons CC0 License, Wikidata makes its data freely accessible for viewing, extraction, and modification in human- and machine-readable formats (Wikidata:Introduction, 2018).

Wikidata entries employ a semantic triple structure through statements about data, making human-readable strings of text machine-readable. In Wikidata *items* (also referred to as *entries* in this paper) represent objects or concepts and are described by *statements* which use *properties* to make factual claims about items. Items and properties receive unique identifiers in the form of a sequence of numbers prefaced by a “Q” for items and a “P” for properties.

Illustrating this structure is the Wikidata item for the poet Melania Fogelbaum (Q11777018), which has a statement using the property “place of birth” (P19) to assert that she was born in Łódź (Q580). There is no minimum nor limit to the number of statements that can be used to describe Wikidata items. Even if items have few statements, they usually have a label, description, and aliases (also referred to as alternative name labels in this paper).

The flexibility of the Wikidata ontology, which makes it so accommodating to various Wikimedia projects as well as metadata sources beyond Wikimedia, is most clearly revealed in the thousands of properties that can be used to describe items. Properties cover a range of classes from the broad to the narrow. For example, the most frequently used property “instance of” (P31) indicates the general category to which an item belongs, while the property “stated in” (P248) is used in the reference section of statements to indicate a specific information source. The variety of properties in Wikidata ensures that virtually all descriptive metadata categories found in other ontologies can be represented in a Wikidata item’s page.

Wikidata as an integration platform

The simplicity and adaptability of Wikidata’s interface as well as the low barrier to entry have allowed for contributions from users with varying technical skills from around the world, increasing the multilingualism and perspectives present in the database. As noted by Allison-Cassin and Scott (2018), contributions to Wikidata can be easily made without the user fully understanding the complexities of the linked open data generated from these modifications. Given the ease of use and global interaction present on its platform, Wikidata poses a golden opportunity for metadata managers from across cultural institutions to link their data within a publicly accessible platform.

Almost since its creation, Wikidata was identified as a resource for libraries to expand the reach of their metadata and to free it from library data silos, particularly in the case of metadata concerning authority control (Klein and Kyrios, 2013). Wikidata has seen the contribution of authority control metadata, particularly the unique identifiers that link to external vocabulary databases, from numerous organizations and consortia around the world. These include the Virtual International Authority File (VIAF), the Library of Congress Name Authority File (LCNAF), the Art & Architecture Thesaurus (AAT), the Musical Instrument Museums Online (MIMO), and Europeana.

Bartholmei et al. (2016) described Wikidata as having the potential to link data across languages and ontologies, which could provide greater support to researchers, promote discovery across disciplines, and democratize access to information. As illustrated by Klein and Kyrios (2013), Neubert (2017), and van Veen (2017), Wikidata can act as a “linking hub” or rather the central axis of a wheel where every spoke is the identifier unique to an outside source. Instead of creating individual bridges between each source, all the sources are connected to a central hub.

The lofty goals for Wikidata posed by librarians and metadata managers correspond with many of the ambitions and requirements needed for the success of the EHRI project. The diversity of descriptive metadata present in the EHRI portal necessitates a platform that can unify multiple languages, alphabets, sources, descriptions, external identifiers, relationships between entries, and allow participants at different technical levels to contribute. Wikidata aligns with these requirements and also fulfills the secondary mandate of EHRI to increase the accessibility of its data.

Updating the EHRI ghettos authority

Updates to the camps and ghettos authorities began with a pilot project examining the effectiveness of using Wikidata to enhance the smaller of the two authorities, the list of Holocaust-era ghettos. The original authority list of ghettos numbered 1,109 and was derived from the online edition of *The Yad Vashem Encyclopedia of the Ghettos During the Holocaust* (Rodriguez et al., 2016). Each authority record included the English and Hebrew name for the ghetto as well as a unique identifier within the EHRI portal. The priority of the pilot project was to expand these entries to include alternative names (e.g. translations in other languages), the place name associated with the ghetto, and geographic coordinates. The secondary aim was to contribute validated information on ghettos to Wikidata. At the start of the project in March 2017, Wikidata included entries for only 80 ghettos.

The first step in the pilot project was to associate each ghetto with the administrative locality to which it belonged and to determine its geographic coordinates. The names of most Holocaust-era ghettos included the name of the locality in which they were located. However, localities throughout Europe were known by several names during the twentieth century due to shifts of borders and changes to the government in power. While the most common name for a ghetto might not be the currently used name for the place where it existed, we chose to use this preferred name as the English-language label for the ghetto with the hope that current and historic names would emerge in the alternative name labels.

To determine these places and their coordinates, we extracted from Wikidata a list of places potentially associated with each of the ghettos. We then compared the coordinates listed in Wikidata for each place with coordinates data from *The Yad Vashem Encyclopedia of the Ghettos During the Holocaust* and location descriptions from the *USHMM Encyclopedia of*

Camps and Ghettos.

During this process we made a few discoveries about our data set. First, a number of known ghettos were not included in the original authority list. While comparing the sources from USHMM and Yad Vashem, an additional 282 ghettos were identified and added to the authority list. Most of these ghettos were extracted from volumes 2 and 3 of the *USHMM Encyclopedia of Camps and Ghettos* and are not included in *The Yad Vashem Encyclopedia of the Ghettos During the Holocaust*. The 1,366 ghettos that presently comprise the EHRI ghettos authority are not an exhaustive representation of all the ghettos that existed during the Holocaust. Undoubtedly, there were more ghettos and there are many resources describing these ghettos that we could have assessed for inclusion in the EHRI ghettos authority. However, for the purposes of our project, we thought it prudent to focus on only two well-known and respected resources as they would provide validated information as well as a natural end point to our research.

Another discovery we made about our data set was the number of ghettos sharing the same name, which led us to craft a standard method for differentiating name labels. While identical preferred name labels in Wikidata are technically permitted, they are not ideal. Given our hope that the entries will be populated with additional information in the future, it is essential for the labels to fully represent the exact ghetto the entry is regarding so there is no confusion to researchers and contributors. Following the standard practice of the LCNAF, ghettos sharing the same name were differentiated by one of their locality's administrative divisions (country, province, county, etc.), which was determined by the level at which the two places stopped having administrative divisions in common.

Wikidata as a source for multilingual labels

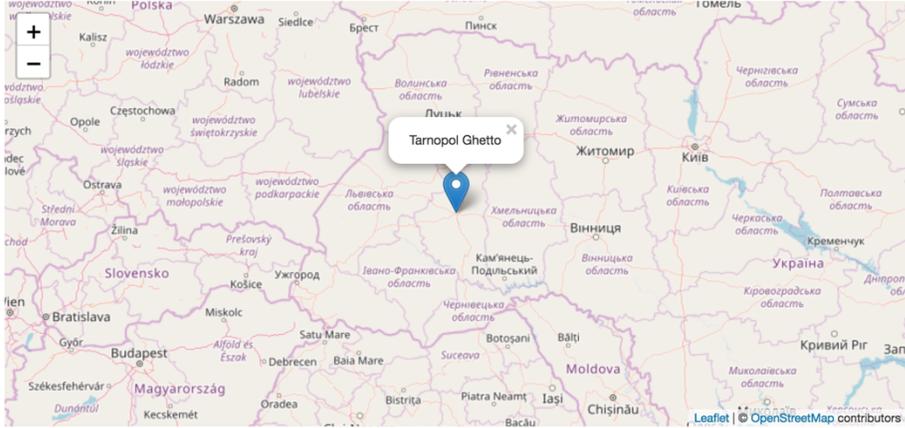
The remarkably low number of ghettos listed in Wikidata at the start of our project did not bode

well for ingesting a lot of new information into the EHRI ghettos authority. However, we were able to leverage Wikidata's prominence as a gazetteer that represents geographical places across the world and includes 1,812,457 administrative territorial entities as of April 2018 (Wikidata:Statistics, 2018). Many of these Wikidata entries link to Wikipedia pages in various languages. Making use of this feature, we extracted name variants for each place based on the Wikipedia labels listed in each locality's Wikidata entry. We prioritized 15 languages, including Bulgarian, Czech, Dutch, English, French, German, Hebrew, Hungarian, Italian, Polish, Romanian, Russian, Serbian, Ukrainian, and Yiddish. We were able to include name variants in Serbo-Croatian and Croatian for some of the ghettos as well. Using these exports and applying separate rules for each language, we created alternative name labels for the ghettos in up to 17 languages.

For example, we were able to generate alternative name labels in 16 languages for the Tarnopol Ghetto, which in the first iteration of the authority was limited to labels in only English and Hebrew. These labels capture the various spellings for the city of Tarnopol, including the current Ukrainian spelling Тернопіль and the transliteration *Ternopil'* (figure 1). Having a lingual diversity of cross-references is important to the EHRI portal as it ensures that potential access points can be more easily identified and applied to archival records that exist in languages other than English and Hebrew.

Tarnopol Ghetto

See Also http://www.yadvashem.org/yv/he/research/ghettos_encyclopedia/ghetto_details.asp?cid=461
<http://www.wikidata.org/entity/Q20981329>



Keywords

This item has multiple descriptions. Alternate descriptions may source from different repositories or represent a (partial) translation.

- Bulgarian - Тернопол (Гето)
- Dutch - Ghetto van Ternopil
- Serbo-Croatian - Ternopol (Geto)
- Romanian - Ghetoul Ternopil
- Italian - Ghetto di Ternopil'
- Polish - Tarnopol (Getto)
- Serbian - Тернополь (Гето)
- Ukrainian - Тернопіль (Гетто)
- Hungarian - Ternopil (Gettó)
- Russian - Тернополь (Гетто)
- German - Ghetto Ternopil
- Czech - Ghetto Ternopil
- French - Ghetto de Ternopil
- Croatian - Ternopilj (Geto)
- English - Tarnopol Ghetto
- Hebrew - גטו טרנופול

Figure 1. The authority record for Tarnopol Ghetto in the EHRI portal.

Importing into Wikidata

After collating the associated place name, the geographic coordinates, and the multilingual labels for each ghetto, the team had an import for Wikidata that included the following descriptions for each ghetto:

- the English name of the ghetto
- a statement qualifying the entry as an instance of a “ghetto in Nazi-occupied Europe” (Q2583015)
- variant names of the ghetto in languages other than English
- the Wikidata unique identifier for the place where the ghetto was located

- the geographic coordinates for the ghetto
- the EHRI-assigned unique identifier for the ghetto
- unique identifiers for the ghetto from online resources (if they existed), including *The Yad Vashem Encyclopedia of the Ghettos During the Holocaust* and the online USHMM Holocaust Encyclopedia²

One of the challenges we faced while compiling this import was finding an appropriate Wikidata property to apply to the EHRI-assigned unique identifier, a feature which would be essential for future EHRI participants to extract and update data. On a Wikidata page, most unique identifiers are found under the “Identifiers” subheading and the authority to which they link is classified as a property. At the time of our pilot project, the structure of the data in the EHRI portal was not conducive to the requirements for a Wikidata external identifier property, particularly for a property for authority control.³ During discussions with the Wikidata community, the property “catalog code” emerged as a reasonable alternative that allowed the ghetto entries in Wikidata to include the EHRI unique identifier (Wikidata:Property proposal/Identifier, 2017).

Extracting from Wikidata

After updating Wikidata with 1,286 new entries for ghettos and enhancing the 80 existing ghetto entries, we extracted data relevant to the EHRI ghettos authority and integrated the data into the EHRI portal. Each entry in the EHRI portal has been enhanced to include many more multilingual alternative name labels, geographic coordinates, a map of the location, and, if possible, a link to the ghetto’s entry in *The Yad Vashem Encyclopedia of the Ghettos During the Holocaust* or the USHMM Holocaust Encyclopedia. Because Wikidata is an open project with data that can be altered by any user, the data extracted from Wikidata for future synchronizations

will need to be validated.

Updating the EHRI camps authority

Preparing the enhancement of the EHRI camps authority provided unique challenges unlike those presented by the EHRI ghettos authority. The ways in which the two data sets were similar extended only to both starting with existing data sets and both requiring similar prioritized descriptions (geographic coordinates, multilingual variant labels, and unique identifiers). The camps data set was more complicated partially because of the errors existing in the original data set, but also due to the general complexity surrounding the history of concentration camps during the Holocaust. Within the scope of an authority, where hierarchical relationships, location, functions, and names are essential for classification, this particular subject matter does not always fit tidily within an ontological structure.

Some of the challenges we faced were varying hierarchical relationships, where subcamps at various points in their existence were administrated by different main camps, sometimes after closing for a period of time but not always. Geographic coordinates were not always easy to determine as camps were moved to new locations or the source materials may be vague on the locality. Additionally, one camp could be classified for many different purposes, such as a forced labor facility, an extermination center, a transit hub, etc. Deciding which name variants to include and leave out was also a challenge, as many were formed with naming protocols unique to their section of the Nazi camp system and, for our purposes, some of these names were too vague to be of real use in the identification of access points in archival records. For example, the code names used by the Nazi administration for various camps were frequently one or two letters (e.g. the subcamp of Buchenwald created in Oberndorf was also known as

“Ms” and “Mu”). Given the difficulty of matching these code names to archival records, we chose not to include many of them in the authority records.

Camp labels

It was essential to establish a few protocols for crafting labels. Following the standard practice of the LCNAF, every camp was qualified as a “concentration camp” (Q152081) regardless of its primary function and even if it did not exist in a permanent location. Camps that shared names were differentiated within their labels by a variety of distinguishing factors, including the camps’ dates of operation, the gender of the inmates, the camps’ management, and the camps’ precise area within a locality. Usually these disambiguating descriptions were included in parentheses, but this standard is not consistent, as some camps were disambiguated in the source materials in varying formats.⁴

Consider the following list of camps and their various disambiguating details:

Differentiated by gender:

Dessauer Ufer men's concentration camp

Dessauer Ufer women's concentration camp

Differentiated by the camp’s forced-labor management:

Lippstadt (Lippstädter Eisen-und Metallwerke) concentration camp

Lippstadt (Westfälische Metallindustrie) concentration camp

Differentiated by area within a locality:

Budapest (Margit Boulevard) concentration camp

Budapest (Rökk-Szilárd Street) concentration camp

Differentiated by date of operation:

Heidenheim concentration camp (1941-1942)

Heidenheim concentration camp (1944-1945)

Within the Wikidata community, the guidelines regarding the disambiguation of labels can be interpreted in various ways. The Wikidata help page for labels states, “When a [Wikipedia] page title includes disambiguation, either through commas or parentheses, the disambiguation should not be included in the Wikidata label. Disambiguation information should instead be part of the description” (Help:Label, 2018). After reading the discussions on this page, however, we determined that the disambiguating elements in the camp labels were part of the name established in the EHRI camps authority as well as other resources, such as the *USHMM Encyclopedia of Camps and Ghettos* and they would not be in violation of Wikidata norms.

Hierarchical relationships

Expressing relationships between camp entries in Wikidata was relatively easy despite the aforementioned complexity inherent to the history of Holocaust-era concentration camps. Keeping it simple was the key, and we decided to limit relationship options to the Wikidata properties “parent organization” (P749) and “subsidiary” (P355). The camps that were administrated by several main camps over their life spans include each of these parent camps in their Wikidata entries, and each of the subcamps are listed in the Wikidata entry for main camps as subsidiaries.

For example, Helmbrechts concentration camp (Konzentrationslager Helmbrechts) was opened in July 1944 as a subcamp of Ravensbrück concentration camp and, in September 1944, it was transferred to the administration of the Flossenbürg concentration camp (Schmidt and Pallavicini, 2009). The Wikidata entry for the Helmbrechts subcamp lists both Ravensbrück and Flossenbürg as parent organizations and each of the Wikidata entries for the parent camps include the Helmbrechts subcamp in the list of subsidiary statements.⁵

Wikidata also offers the option to add a qualifying description under each parent organization statement that indicates the duration for which the subsidiary was under the control of the parent organization. Ideally, the relatively sparse Wikidata entries we contributed for camps will be updated in the future to include this level of detail.

Contributing to Wikidata

Challenges also existed with regards to the varying frameworks used within Wikipedia to describe camps. These variants emerged while we were running a range of queries in the DBpedia and Wikidata SPARQL query interfaces to determine how these camps were already listed in Wikipedia. At the beginning of our search approximately 400 Wikidata entries described specific camps; however, descriptions of individual camps were identified throughout Wikipedia and found to be structured in a variety of manners. Several large camp systems, such as Auschwitz, Buchenwald, and Dachau, had Wikipedia pages dedicated to listing their subcamps.⁶ Some of these subcamps linked to their own Wikipedia entries, but most had either a red link (indicating that the link-to page does not exist and needs to be created) or they linked to the geographical locality where the camp existed. Information regarding many of the mobile camps appeared in the subheadings of Wikipedia articles but did not link to dedicated entries.⁷ We are hoping that by contributing individual entries for each camp into Wikidata a more consistent framework for describing Holocaust-era camps can be injected into the Wikimedia landscape.

Synchronizing with EHRI

We used a similar process of extraction and synchronization for the camps data as we did the ghettos data. Each camp entry in the EHRI portal includes the camp's primary name label, alternative name labels, hierarchical relationships with other camp entries, the URL for the

Wikidata entry, and a map showing the location of the camp (see figure 2).

EHRI Camps / Ravensbrück concentration camp / Flossenbürg concentration camp

Helmbrechts concentration camp

KZ-Außenlager Helmbrechts

See Also <http://www.wikidata.org/entity/Q3566563>

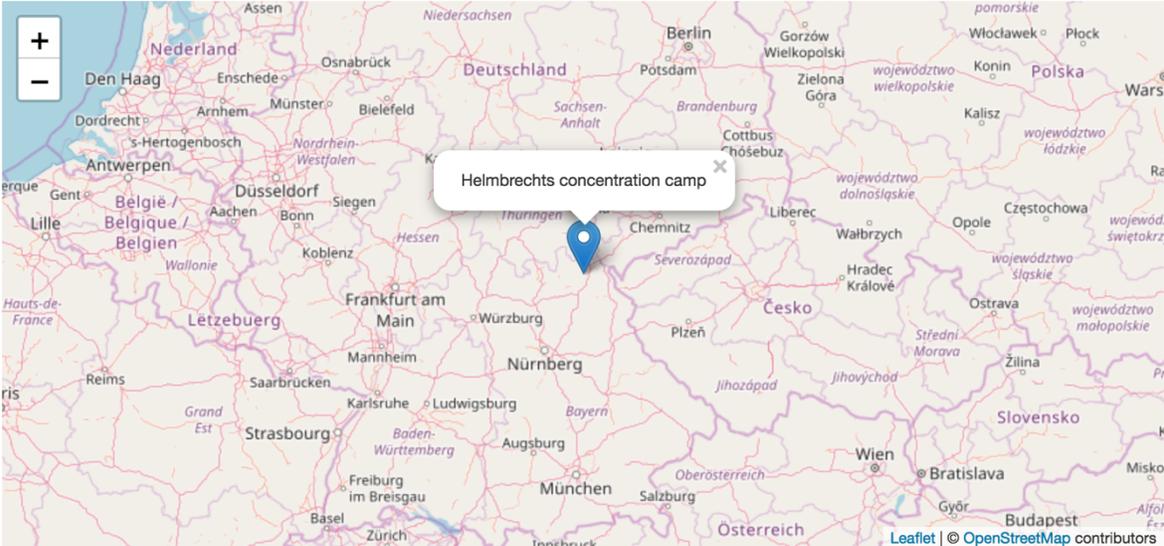


Figure 2. The authority record for Helmbrechts concentration camp in the EHRI portal.

Results

As of September 2018, Wikidata includes entries for over 3,000 and 1,300 Holocaust-era camps and ghettos respectively. Each entry has an identifier associated with EHRI and many entries also include geographic coordinates, alternative names, identifiers for other authorities, and relationships between entities. For most of the data points references have been included to indicate the source of the statement. Figures 3, 4, and 5 show mapped visualizations of the camps and ghettos listed in Wikidata before and after EHRI's contributions (only camps and ghettos with coordinates included in their Wikidata entry are present on the maps).

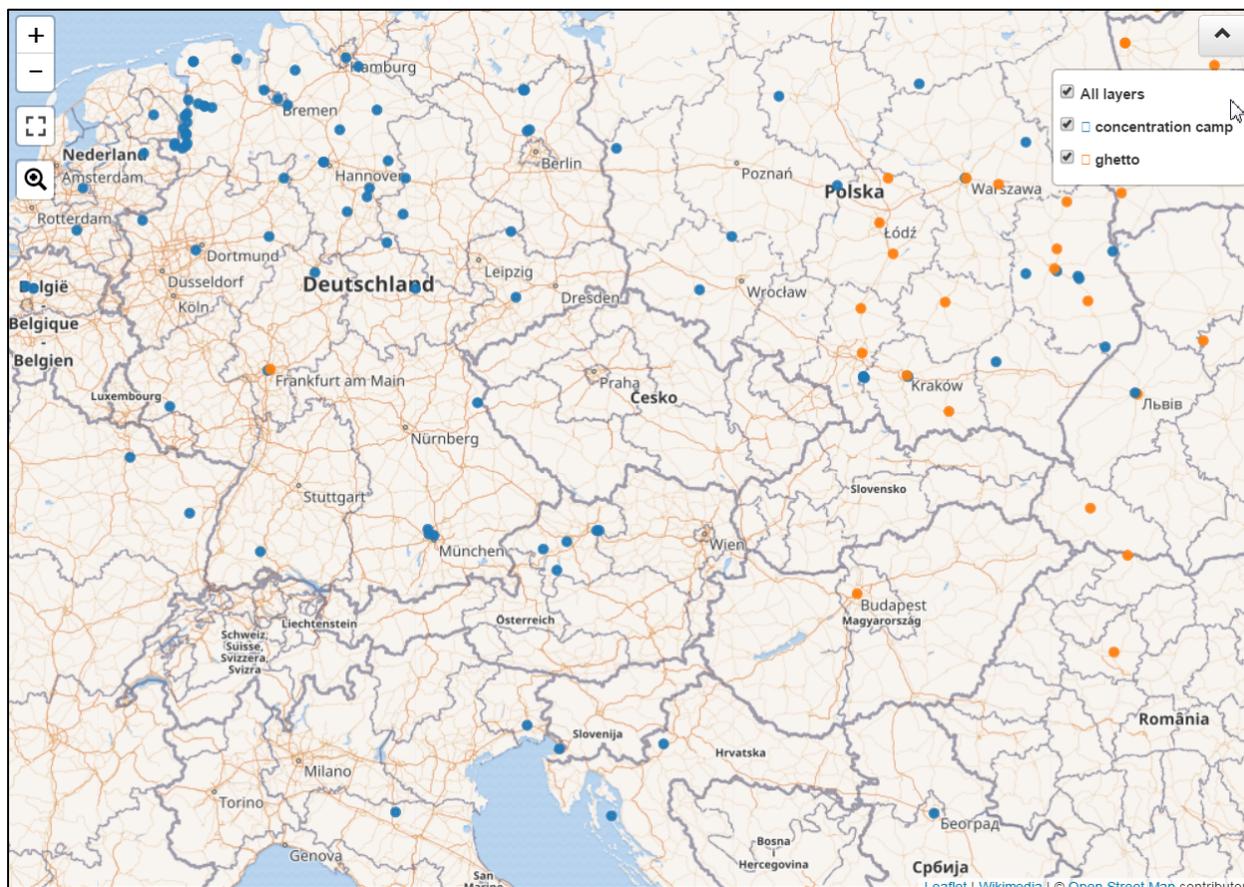


Figure 3. A mapped visualization of the camps and ghettos included in Wikidata before EHRI's contributions. Ghettos are represented by orange dots and camps are represented by blue dots.

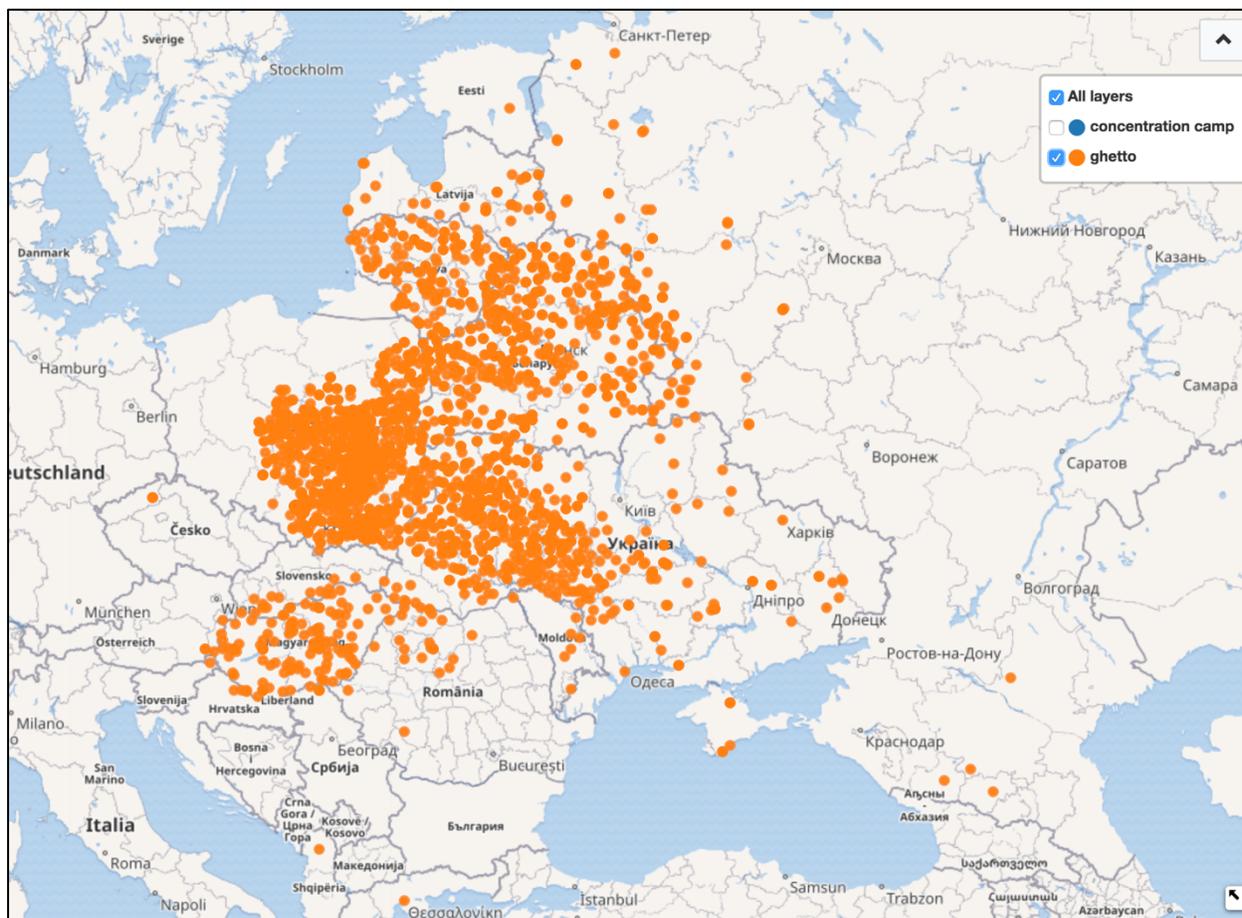


Figure 4. A mapped visualization of the ghettos included in Wikidata after EHRI's contributions.

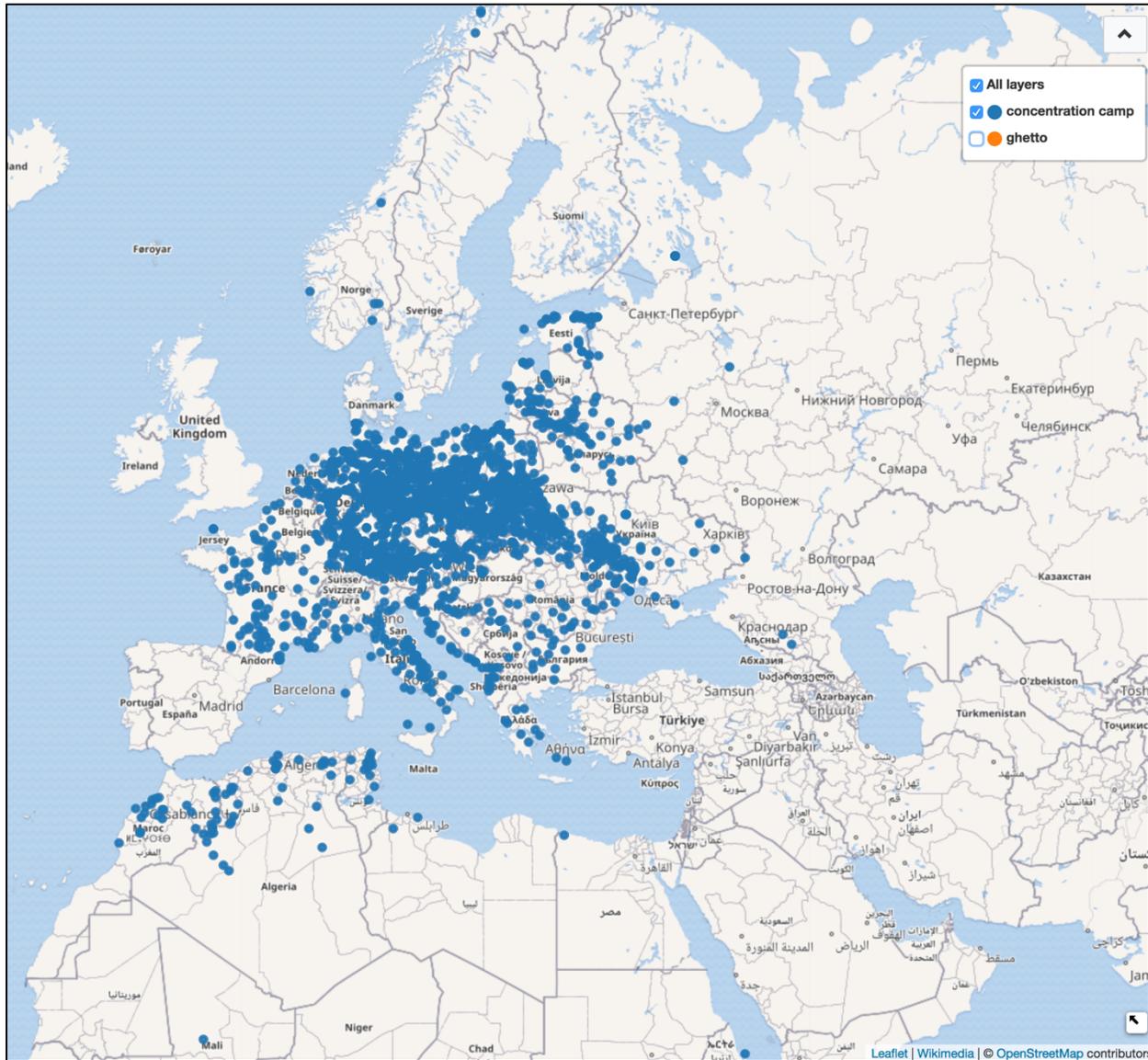


Figure 5. A mapped visualization of the camps included in Wikidata after EHRI's contributions.

Since the EHRI portal was updated with data extracted from Wikidata, there is more descriptive information in each authority record, including alternative name labels, geographic coordinates, and hierarchical relationships. In addition, the EHRI portal has benefitted from the linked open data capacities of Wikidata by incorporating identifiers from other resources into the records for some of its authority records. Within the EHRI authority records, these identifiers

have been incorporated into URLs that link users to the entry for each camp or ghetto in Wikidata as well as to entries in resources from USHMM and Yad Vashem.

Looking towards the future

Moving forward, we hope that the Wikidata entries for camps and ghettos are expanded to include more descriptive data contributed by EHRI participants as well as Wikidata users. The human network of EHRI provides many opportunities to work with content experts to contribute more information to EHRI and Wikidata entries. Much can be improved in the Wikidata entries for camps and ghettos, as many are just skeleton records without any links to Wikipedia entries. Most of the camp entries also lack multilingual alternative name labels. In an ideal world, these records would be enhanced when other institutions with Holocaust-focused authorities and thesauri integrate their metadata into Wikidata. Future EHRI projects may also include matching and adding VIAF and LCNAF identifiers to the camp entries listed in Wikidata and incorporating relational data between these entries and other authority lists from the EHRI portal, like those for persons and corporate bodies.

While the crowd-sourced nature of Wikidata will require validation of new data before it is integrated into the EHRI authorities, we believe that the benefits of crowd contributions will outweigh the potentially negative aspects. Significantly, the information that has been contributed to Wikidata by EHRI was derived from reliable sources from Yad Vashem, USHMM, the International Tracing Service, and Bundesarchiv, indicating to researchers at all levels that it can be depended upon for accuracy. By being part of the linked open data environment of Wikidata, this reliable data on Holocaust-era camps and ghettos is freely usable by anyone, including the Wikimedia community, students, researchers, and other name authority managers.

The accessibility of Wikidata permits substantial possibilities for how the camps and ghettos authorities can be used and applied by researchers and institutions, potentially opening new avenues for inquiry and analysis.

Conclusion

By leveraging the capabilities of Wikidata, EHRI was able to enhance its descriptive metadata and increase potential points of access across its portal. From a broader point of view, EHRI has helped to increase the visibility of information related to Holocaust-era camps and ghettos and demonstrated how other institutions can perform similar projects.

Wikidata holds great potential for integrating ontologies and bridging sources of knowledge. Its ease of use and global presence make Wikidata a promising tool for libraries and cultural institutions, which often do not have the time, resources, or qualified personnel to invest in large scale data mapping projects on their own or with partner organizations. By contributing their data to Wikidata, leaders in thesaurus management and authority control can improve their local metadata while also expanding the reach of that data.

This is an Accepted Manuscript of an article published by Taylor & Francis in the *Journal of Library Metadata* on 30 April 2019, available online:

<http://www.tandfonline.com/10.1080/19386389.2019.1589700>

Notes

¹ The five authority lists are for ghettos, camps, people, corporate bodies, and administrative districts.

² The USHMM Holocaust Encyclopedia is available from <https://encyclopedia.ushmm.org/>. It should not be confused with the *USHMM Encyclopedia of Camps and Ghettos*.

³ The lack of a consistent formatter URL that could be used by all the authority records in the EHRI portal was the primary issue. Because of the modularization of the camps and ghettos authorities in EHRI, the unique identifiers exist in sets independent of one another resulting in URLs uniquely formatted to each authority set. Proposing individual properties for both the ghettos and the camps authorities did not seem advisable.

⁴ Most of the labels compiled for the updated authority came from the *USHMM Encyclopedia of Camps and Ghettos* (volumes 1-3), which has its own format for constructing labels detailed in the volume 1 “Reader's Guide to Using the Encyclopedia” on page xxxviii. Other sources for primary and alternative name labels, geographic coordinates, and hierarchical relationships came from the International Tracing Service, Wikidata, Wikipedia, and Bundesarchiv. The latter generously provided a copy of their camps and detention centers database for easier querying. This database is searchable online at <https://www.bundesarchiv.de/zwangsarbeit/haftstaetten/index.php>.

⁵ See the Wikidata entries for Helmbrechts, <https://www.wikidata.org/wiki/Q3566563>, Ravensbrück, <https://www.wikidata.org/wiki/Q159483>, and Flossenbürg <https://www.wikidata.org/wiki/Q274354>.

⁶ See the following Wikipedia pages:

https://en.wikipedia.org/wiki/List_of_subcamps_of_Dachau,

https://en.wikipedia.org/wiki/List_of_subcamps_of_Buchenwald, and

https://en.wikipedia.org/wiki/List_of_subcamps_of_Auschwitz.

⁷ See for example: <https://de.wikipedia.org/wiki/SS-Baubrigade>.

References

- About EHRI. (2010, September 1). Retrieved from <https://ehri-project.eu/about-ehri>
- Allison-Cassin, S., & Scott, D. (2018). Wikidata: a platform for your library's linked open data. *The Code4Lib Journal*, (40). Retrieved from <https://journal.code4lib.org/articles/13424>
- Anderson, S., & Blanke, T. (2013, December). Intermediating the Human and Digital: Researchers and the European Holocaust Research Infrastructure, 19. Retrieved from https://www.ehri-project.eu/sites/default/files/downloads/EHRI%20working%20paper_final.pdf
- Bartholmei, S., Franks, R., Heilman, J., Joseph, M., McDonald, V., Raunik, A., Ridge, M., and Robertson, M. Opportunities for Academic and Research Libraries and Wikipedia. Columbus, Ohio, 2016. Retrieved from <https://2016.ifla.org/wp-content/uploads/sites/2/2016/08/112-IFLAWikipediaAcademicandResearchLibrariesDiscussioDRAFT.pdf>
- Factsheet EHRI. (2017, January). Retrieved from https://ehri-project.eu/sites/default/files/downloads/ehri_downloads/Factsheet%20EHRI%20January%202017.pdf
- Help:Label. (2018, August 16). Retrieved from <https://www.wikidata.org/wiki/Help:Label>
- Klein, M., & Kyrios, A. (2013). VIAFbot and the Integration of Library Data on Wikipedia. *The Code4Lib Journal*, (22). Retrieved from <http://journal.code4lib.org/articles/8964>
- Megargee, Geoffrey P. (Ed). *The United States Holocaust Memorial Museum Encyclopedia of Camps and Ghettos, 1933-1945* (Vol. 1-3). Bloomington: Indiana University Press, 2009, 2012, 2018.
- Miron, G., Shulhani, S., & Yad a-shem, rashut ha-zikaron la-Sho'ah v'ela-gevurah. *The Yad Vashem encyclopedia of the ghettos during the Holocaust*. Jerusalem: Yad Vashem, 2009.
- Neubert, Joachim. (2017). Wikidata as a linking hub for knowledge organization systems? Presented at the NKOS Workshop at TPDFL, Thessaloniki, Greece. Retrieved from <http://ceur-ws.org/Vol->

1937/paper2.pdf

Rodriguez, K.J., Alexiev, V., Brazzo, L., Riondet, Ch., Gherman, Y., Speck, R. (2016). D.11.2: Road Map Domain Vocabularies. EHRI 2 Deliverables.

Schmidt, Alexander, and Stephen Pallavicini. "Helmbrechts." *The United States Holocaust Memorial Museum Encyclopedia of Camps and Ghettos, 1933-1945*, edited by Geoffrey P. Megargee, vol. 1: Early Camps, Youth Camps, and Concentration Camps and Subcamps under the SS-Business Administration Main Office (WVHA) (pp. 608-610). Indiana University Press, 2009.

van Veen, T. (2017, October). Wikidata as a universal (library) thesaurus. Presented at WikidataCon in Berlin, Germany. Retrieved from https://commons.wikimedia.org/wiki/File:Wikidata_as_universal_library_thesaurus_-_tvv.pdf.

Wikidata:Introduction. (2018, June 18). Retrieved from <https://www.wikidata.org/wiki/Wikidata:Introduction>

Wikidata:Property proposal/Identifier. (2017, May 22). Retrieved from https://www.wikidata.org/wiki/Wikidata:Property_proposal/Identifier

Wikidata:Statistics. (2018, April 30). Retrieved from <https://www.wikidata.org/wiki/Wikidata:Statistics>