

Bilaga 1

Developing a Wikimedia and ICOM/IFLA project

Collecting and sharing data among cultural professionals is an important role for network organizations. Increasingly, using the linked open data ecosystem is an important part not only of curating individual institutional collections, but bridging across those collections.

In response to the growing need for effective linked open data that can facilitate the work of IFLA, ICOM, and other heritage networks, we are suggesting a mapping of GLAMs (Galleries, Libraries, Archives, and Museums) in the linked open data project Wikidata.

Like the #1lib1ref campaign, that asks every librarian to add one more reference to Wikipedia, while explaining how Wikipedia fits into library strategies, the project serves two functions: first, to collect data about the institutions, including location and other key public information,¹ and second, to increase literacy among cultural professionals about Wikidata, in the context of linked open data and its applications more generally. By integrating structured data about cultural institutions into this central repository of Wikidata, the data, which is released under a CC0 license, can be mixed and matched into other collections of data.² In turn, these applications can help develop a more advanced understanding of the world's GLAM institutions.

Structure of the campaign

The Wikimedia Foundation and Wikimedia Sverige will prepare instructions and campaign pages for participating in the project on Wikidata. Working with IFLA and ICOM (and their broader network of organizations), we will pick a multi-week window in mid-2019 where we can encourage libraries and museums to share their knowledge about their institutions. They will be able to share the data directly on Wikidata or through an online form.

The structure of the campaign will focus on decentralized advocates, both in the Wikimedia community and in the professional networks, encouraging participation through social media and discussion in public forums (modeled by the existing #1lib1ref campaign). Moreover the

¹ The initial data models are at:

https://www.wikidata.org/wiki/Wikidata:WikiProject_Heritage_institutions/Data_structure. There are dozens of potential data items that could be created for each, but we recommend prioritizing: instance of (P31), part of (P361), coordinate location (P625), located at street address (P969), located in the administrative territorial entity (P131), country (P17) and public data that helps with contacting and learning about the org, such as official website (P858) or Twitter username (P2002), and identifiers for the institution, such as GND Integrate Authority File (P277) or ISIL ID (P791). If available, collection or exhibition size (P1436) can be added as well. This is open to reconfiguration based on need, and broadly could be used as a first step. One consideration is having a second step for the campaign that either focuses on significant items in the collection or adding other institutions in your region.

² For the reasoning behind this license, see

https://upload.wikimedia.org/wikipedia/commons/1/18/Arguments_on_CC0-licensing_for_data.pdf

visualizations provided by the crowdsourcing tool for the campaign, will help identify gaps that can be used to advocate for more participation.

IFLA and ICOM will activate their networks, via social media and traditional communications, about the importance of linked open data. This involves ~1 month of communications build-up before the campaign, as well as coordination during and after it.

Using a hashtag and the campaign pages, we will push educational materials about the use of linked open data to map institutional information, and other applications for linked open data.

Some data is already collected by a number of institutions or professional networks: we wouldn't want to duplicate too much of that existing work. Providing a pathway for upload of this data to Wikidata will be an important part of the campaign, supported by Wikimedia Sverige. Our focus would only be datasets that the institutions or professional networks are willing to release under a free license and share with the world.

The inclusion of the data can be done by building a specially designed small computer program that automatically adds the information on Wikidata. This saves a lot of work for the institutions and ensures that there is good baseline coverage with high quality data. Each piece of information on Wikidata can be sourced and by doing so highlight the work of these institutions and professional organizations.

As part of the crowdsourcing and outreach activities, we will encourage participants in the campaign to verify the data in Wikidata, which in turn can help the professional communities enrich their own data.

Year 1 focus: Does the crowdsourcing work? Can we measure it? Can we make the case for this contribution within existing high data-literate networks? Can we make the case for the release of full datasets which can be included?

Year 2 focus: Expand to higher priority networks that would help with broader mission, and begin strategically expanding the kinds of data encouraged for the campaign. The focus would be on filling gaps in high risk or underrepresented regions, around institutions and collections.

Value for Heritage organizations

- Gives ICOM and IFLA a supplementary set of information about institutions that can be linked to authority controls and other more authoritative datasets, allowing synchronization across different efforts.
- Wikidata allows rapid availability of this supplemental data, through its inclusion in OpenStreetMap and other projects, as part of crowdsourcing efforts during humanitarian and disaster response.
- Inclusion in Wikidata increases visibility of institutions, including:
 - providing data for Google search results, OpenStreetMap and a large number of other digital services both commercial and non-profit. By curating

institutional data on Wikidata, we create better representation of cultural heritage organizations in those increasing number of contexts and ensure that the collected data gets a significant usage, therefore increasing the value of the efforts.

- increased attention from Wikimedia Communities on these institutions, increasing the likelihood and opportunity for coverage of the institution in projects like Wikipedia and Wikimedia Commons.
- Wikimedia communities are already collecting data about cultural institutions by crowdsourcing and aggregating existing open data sets about institutions — this collaboration would augment and improve those efforts to better reflect the needs of the professional community.
- Literacy in Wikidata will help bridge the gap between cultural professionals already working on linked open data, and the increasing value of using central identifier systems like Wikidata.

Risks

We have identified a few risks of the project:

- Institutions are sometimes unaware of the consequences of releasing their data under an open license or uncomfortable with it being “outside of their control” — teaching this kind of comfort with open contexts, like Wikidata, will be important.
- There will be datasets that are not possible to include because of unclear copyright, data protection regulations, personal rights issues etc. A vetting process and clear instructions for what data can be included will be important, and only datasets where this has been cleared will be included and data about individuals will be excluded.
- The data that can be collected on Wikidata, needs to be publicly visible data, preferably with a public source for that information. Communicating the difference between already public data, and efforts by ICOM, IFLA and others to collect more sensitive data about collections and historical sites will be important.

Contribution by different participants:

- **Wikimedia Sverige (WMSE)** is the local Wikimedia affiliate in Sweden. WMSE, during 2016-2017, worked closely with UNESCO and a number of other partners to collect information about cultural heritage in danger. This include uploading official datasets about the protected cultural heritage sites, uploading new and historical images of the sites and promoting the knowledge and understanding of the issue through presentations, a photo exhibition and online activities. This experience will be valuable when working to upload new data and include it on the other Wikimedia projects (such as Wikipedia), create targeted information material and to reach out about this project.
- **Wikimedia Foundation (WMF)** will support development of the communications campaign and work with a volunteer developer to optimize a crowdsourcing layer for contributing to the campaign. WMF has experience running these kinds of campaigns with the #1lib1ref campaign, that encourages librarians to add references to Wikipedia. WMF will also work with WMSE to encourage other affiliates and

Wikimedia communities to help participate in the campaign, encouraging engagement with a larger network of institutions.

- **IFLA and ICOM** will connect interested parties/working groups within their networks to the campaign (discussions about their involvement are ongoing). Moreover, as part of the campaign, both organizations will support communications around the relevance and impact of linked open data, for describing cultural institutions to the public. Having endorsement and support of the campaign by IFLA and ICOM, is an important part of ensuring that the professional community is connected with the project, and that the message of the importance of linked open data reaches those communities.

Support from other affiliates and community members:

Once we firm up the project dates, WMSE and WMF will work with regional affiliates in the Wikimedia Community to support communications and activation of heritage institution networks within their communities. We have communities active in multiple regions throughout the world that are interested in working with cultural heritage institutions.

About Wikidata

Wikidata is the Wikimedia community's linked, open, multilingual, structured database that is openly editable on a wiki platform, like other Wikimedia projects. In practice, it serves as a central data hub for structured information that can be used in the individual Wikipedias, but the machine-readable data can be freely accessed and reused by anyone. For the cultural heritage community, the increasing value of Wikidata is the ability to rapidly synchronize and connect different parts of the existing authority control ecosystem across previous and current linked open data efforts, while also gaining the secondary value of the data made available by Wikimedia volunteers.

External links

- Wikidata + GLAM: <https://blog.wikimedia.org/2016/08/23/wikidata-glam/>
- Data modeling and case studies for Heritage Institutions: [WikiProject Heritage Institutions](#)
- Existing support for batch uploading from Wikimedia Sverige: [Connected Open Heritage](#)
- Wikidata queries
 - [Libraries currently mapped in Wikidata](#)
 - [Museums currently mapped in Wikidata](#)
 - [Archives currently mapped in Wikidata](#)