

Thanks for filling out [Crowdsourcing/UGC in cultural heritage](#)

Here's what we got from you:

[Edit response](#)

Crowdsourcing/UGC in cultural heritage

This survey is being carried out within the Common Culture project in order to assess the state-of-the-art in Cultural Heritage crowdsourcing activities across Europe.

The majority of the questions in this survey are either tick box or multiple choice. There are options throughout to leave any comments you may have.

Should you have any questions or experience any difficulties in completing the survey please contact Rob Davies: robert.davies@cut.ac.cy

DATA PROTECTION

This data is being gathered as part of a survey to assess the situation of Cultural Heritage Crowdsourcing activities in Europe. It will be used during 2020 for a one-time analysis. No personal information will be retained.

In addition to your feedback we also collect some personal information such as your name and email address to be able to contact you for follow up if requested.

We respect your trust and protect your privacy therefore we will never sell or share this data with any third parties.

We will securely store this data until the end of December 2020 when the research period is over.

By filling out this form you agree that we will process your data in line with our privacy policy.

If you have any questions or change you mind please contact our Data Protection Officer, Maria Ioannou: gdpr.dpo@cut.ac.cy

Email address *

david.haskiya@wikimedia.se

Project details

Your name

David Haskiya

Organisation

Wikimedia Sverige

Country

Sweden

Crowdsourcing project

What is the name of your crowdsourcing project?

GLAMWiki-collaboration (ongoing activity, not a project)

What is the topic scope of your crowdsourcing project?

How GLAMs can work with the Wikimedia movement and communities to reach and engage more people. For most GLAMs working with the Wikimedia community on the Wikimedia platforms is about increasing their audience reach and providing raw materials for volunteers to contextualise in Wikipedia articles, but GLAMs increasingly work with Wikimedians in order to enrich their own collections data, see eg.

<https://diff.wikimedia.org/2019/12/13/data-roundtripping-a-new-frontier-for-glam-wiki-collaborations/>

The work is typically centred around the GLAMs collections and will, again typically, involve publishing metadata and mediafiles for parts of the GLAM's collection, or the entire collection, on Wikimedia platforms such as Wikidata, Wikimedia Commons, and Wikipedia. Platforms where the public can then work to update, translate, and enrich the information, metadata or media files. The nature of the collaboration varies with the GLAM, their collection and their particular goals with the collaboration. A core aspect of this work is to advocate for open licensing of GLAM-content, a prerequisite for using the Wikimedia platforms.

Note that Sweden is by no means the only Wikimedia chapter to work with GLAMWiki-collaboration - it's a global format. To get a sense of the scope take a look at the monthly This Month in GLAM-newsletter,

<https://outreach.wikimedia.org/wiki/GLAM/Newsletter> which is written by Wikimedia

volunteers. In the latest issue alone, and limiting ourselves to Europe, there are reports from projects run in Albania, France, Netherlands, North Macedonia, Serbia, Sweden, Switzerland, and the United Kingdom.

What is the geographic scope of your crowdsourcing project?

Please tick all that apply.

International
|

Europe

National
|

Regional/local
|

If 'National' please select country

Sweden

If regional/local please provide details

Most of our ongoing work is with Swedish GLAMs. We also work with UNESCO Archives and participate in EU-funded projects.

Is there anything else you'd like to tell us about your answer?

The answer above is for Wikimedia Sverige specifically. GLAMWiki-collaboration in general is part of the global Wikimedia movement.

What were the funding sources for your crowdsourcing project?

Please tick all that apply



EU



National

|



Regional/local

|



Foundation/NGO

Private subsidy



Income-funded

Is there anything else you'd like to tell us about your answer?

Again, the answer is for our GLAMwiki collaboration in general, not for a specific project (of which there are too many to answer this survey individually).

What is/was the duration of your crowdsourcing project?

Please give start date (if known)

MM

/

DD

/

YYYY

Please give end date (if known)

MM

/

DD

/

YYYY

Is there anything else you'd like to tell us about your answer?

Not filling this in as we're not providing information about a specific project. We've been doing GLAMwiki-collaboration since at least 2010.

What are/were the main aims of your crowdsourcing project?

Please tick all that apply

Broadcast search – solving a problem



Community engagement/participation

Create common goods

Distributed human intelligence



Knowledge discovery and management

Peer-vetted creative production – assessing public support

Preservation/documentation

- Other: The goal of the Wikimedia movement is to provide free knowledge to everyone. The goal of participating GLAMs is typically to educate and inform the public on the basis of their collections.

Is there anything else you'd like to tell us about your answer?

What activities have you carried out in the project?

Please tick all that apply

- Collect digitised objects or data (aggregation/repository building)

Collectively build digital artifacts (e.g. 3D models)

Convert analogue to digital (digitisation)

- Edit or enhance, improve quality

Evaluate digital items/research data

Prepare data/metadata for aggregation elsewhere

Process/analyse/interpret information or datasets

Transcription

Other:

.....
Is there anything else you'd like to tell us about your answer?

What type of organisation facilitates/coordinates the project?

Please tick all that apply

Archive

Community organisation

Gallery



Government body

Higher education
institution

Individual



Library



Museum



NGO

Private sponsor

Research organisation

Standalone initiative

Other:

Is there anything else you'd like to tell us about your answer?

What was the mode of collection?

Please tick all that apply

Analogue



Mobil
e

Social media



Web-based

Other:

Is there anything else you'd like to tell us about your answer?

The Wikimedia platforms are all developed to work at least decently on mobile devices. For some there are official mobile applications.

What tasks were crowdsourced?

Please tick all that apply

Audio
transcription

Autotagging/recognition

Captcha

Categorisation

Classification

Counting 'likes'

Digital art

Digitising/uploading

Error checking/cleaning

Gaming

Georeferencing

Indexing

Linking

Photo-masking

Ranking algorithms

Tagging/annotation

Other: Translation of descriptive metadata, Text transcription/OCR correction, Photographic documentation (of cultural and natural heritage environments and objects), Writing of encyclopedic articles

Is there anything else you'd like to tell us about your answer?

What types of content were collected?

Please tick all that apply

AR/VR

Archival or text-based documents

Cultural artifacts

Online data

Online metadata

Photographs

Social media posts

Sound

Video

Web archives

Other:

Is there anything else you'd like to tell us about your answer?

What supporting events were organised (online or physical)?

Please tick all that apply



Continuous



Hackathon

Research/fieldwork

Roadshow

Transcribathon



Other: Editathons (=time limited intensive and collaborative editing of Wikipedia articles, Wikimedia Commons file metadata, etc.), Photography campaigns (Wiki Loves Monuments, Wiki Loves Earth)

Is there anything else you'd like to tell us about your answer?

We've mostly partnered with others in arranging hackathons, eg with the Swedish National Heritage Board.

What was the context for the project?

Please tick all that apply



Citizen Science

Storytelling



Wikidata



Other: Collaborative editing of encyclopedias, Sharing of educational information related to heritage, Awareness raising in regards to cultural and natural heritage, Writing Wikipedia articles as part of school lessons

Is there anything else you'd like to tell us about your answer?

Were any of the following technical framework or tools employed?

Please tick all that apply

CrowdHeritage

EnrichEuropeana

Metadata
games

OAI-PMH

Pybossa

Zooniverse

Offline
tools

None



Other: Mediawiki incl. its Wikibase extension, Multiple accessory/productivity tools based on the Mediawiki APIs

Is there anything else you'd like to tell us about your answer?

How is/was the crowdsourced content stored?

Please tick all that apply

Internal repository

Cloud-based



Other: Contributed data is stored on Wikimedia servers. These are primarily located in the US.

Is there anything else you'd like to tell us about your answer?

Some GLAMs we work with choose to include user-created metadata and links added to their content on the Wikimedia platforms in their own databases. This can be eg. improved, translated or corrected metadata. It's becoming increasingly common to align Wikidata-object URIs with GLAM's own vocabularies, see eg <https://wikimedia.se/2020/06/10/kulturnav-a-hub-for-museum-vocabularies-connected-to-wiki-data/>

Were any of the following metadata and identifier standards employed?

Please tick all that apply

DOI

Dublin Core

EAD

EDM

FOAF

Getty vocabularies

IIF

METS

Schema.org



Home-grown

Persistent
Identifier

None

Other:

Is there anything else you'd like to tell us about your answer?

Varies from GLAM to GLAM. File formats are often standard (XML, CSV) but data models vary.

There are a number of tools for wrangling and batch uploading data to Wikimedia Commons and Wikidata, as long as the in-file format is, or can be converted to, XML or CSV we can work with it.

How was the metadata created or generated?

Please tick all that apply



Manually created



Captured from external source



Auto-generated

Other:

Is there anything else you'd like to tell us about your answer?

The original data comes from the GLAM's Collections Management System or equivalent.

We assume that is what is mean by "captured from external source" above (?).

Some metadata is autogenerated as the batch upload is being prepared. These are typically administrative metadata, Wikimedia Commons categories, language tags, and copyright statements.

For batch uploads manually created data is added post-upload to the Wikimedia platforms.

What was the quantity of crowdsourced items collected?

Please enter numerical values to the nearest 100 below

Digital Objects

25000

Metadata records

25000

User interventions (tags etc.)

Is there anything else you'd like to tell us about your answer?

The digital objects and metadata records already exist - what we do is that we publish them where they can reach a large amount of people. So original digital objects and metadata records are typically not created in our GLAMWiki-collaborations, but are published on our open platforms and thus become available for collaborative editing and contextualisation. About 260 000 media files have been published on Wikimedia Commons as a direct result of our content partnerships with Swedish GLAMs, see https://commons.wikimedia.org/wiki/Commons:Wikimedia_Sverige#Opening_and_sharing_free_material

Wiki Loves Monuments Sweden, which has support from the Swedish National Heritage Board, the Museums of Work Secretariat, and the National Maritime Museums, has resulted in 25 200 new/original photographs with associated metadata - importantly always with a reference-link to the monument's ID in the the official Swedish Cultural Environments Registry. This allows the National Heritage Board to include links to the photographs in the national aggregator SOCH/K-samsök. For all WLM-stats, see https://commons.wikimedia.org/wiki/Commons:Wiki_Loves_Monuments

Were any of the following exchange protocols used?

Please tick all that apply

OAI-PMH

ResourceSync

SPARQL

SWORD

XML/RDF

None



Other: We work with whatever data dumps or technical interface a GLAM can offer. If we can work it, we accept it. We set no formal technical requirements, but do have hard requirements concerning copyright.

Is there anything else you'd like to tell us about your answer?

RDF/XML is not an exchange protocol. Nor is SPARQL (though it can be used to fetch data). XML and CSV are the most common formats we receive from the GLAMs, in very few cases does the XML serialize RDF.

Who were the main participants in crowdsourcing?

Please tick all that apply

General population

Children/school

s



Higher
education/research



Members/frequent
users



Professional
s



Retired people

Self-employed/casual
(paid)



Volunteer
s

Other:

Is there anything else you'd like to tell us about your answer?

For privacy reasons the Wikimedia movement does a minimum of tracking users. From surveys its known that Wikimedia volunteers skew towards male, white and academic. Paid editing and original research is discouraged,. GLAM-professionals are encouraged to

contribute based on their expertise, but discouraged from using the Wikimedia platform to "market" their own institution.

How many participants contributed?

Please tick one option

0-50

51-100

101-500

501-1000

1001-2000

more than 2000

Is there anything else you'd like to tell us about your answer?

Answer limited to Wikimedia Sverige GLAMWiki-collaborations only and very approximate.

What were the main intended motivations for participation?

Please tick all that apply



Contributing to knowledge
production

Identity/self-definitio
n

Interest/curiosit
y

Learning about heritage and history

Promoting locality and
place

Skill/career
development

Supporting an institution or
cause



Other: The motivation of individual Wikimedia volunteers vary greatly - all of the
above included.

Is there anything else you'd like to tell us about your answer?

Contributing free knowledge, in all languages, is the goal of the Wikimedia community as a whole. The language aspect is an important motivator for many Wikimedians, perhaps especially so for minority language communities.

What forms of evaluation were carried out?

Add question text

Validation of user
contributions

Formative evaluation

Impact
Assessment

Interface
heuristics

Outcome harvesting



Output
metrics



Summative evaluation

Is there anything else you'd like to tell us about your answer?

We also publish white papers, or case studies if you will. Not sure where they fit in the examples above. Example white papers here,

https://meta.wikimedia.org/wiki/FindingGLAMs/White_Paper

Under which licenses was the crowdsourced material made available for use?

Add question text



CC-BY



Other CC

GPL

None

Other:

Is there anything else you'd like to tell us about your answer?

Wikipedia articles are CC-BY-SA. Wikidata metadata is CC0. Wikimedia Commons media files vary in their licensing, but must be Open - so CC0, CC-BY, CC-BY-SA are the only supported Creative Commons licences. The PD-mark can be used to complement a

statement that there is no copyright in the media file, but the claim that something is in the public domain must be explicitly motivated (and based on the relevant jurisdictions).

Additional information

Please provide any additional information that you feel is relevant to your project. We would be interested to hear about challenges faced, sustainability plans etc.

Just a reminder that this response is based on our GLAMWiki-collaboration work in general. However, if you want to learn more about a specific GLAMwiki-project do check out FindingGLAMs, https://meta.wikimedia.org/wiki/FindingGLAMs/White_Paper It's recent and also comprises a wide spectrum of GLAMWiki-activities and different types of GLAM-data/collections.

A reiteration that the GLAMWiki-movement is global. This survey response is representative of the work carried out by Wikimedia Sverige only.

[Create your own Google Form](#)