

For information to the GLAMWiki-community

Two applications for GLAMWiki
development projects

Background and context

Like many Wikimedia chapters Wikimedia Sverige is often in pursuit of opportunities to gain funding for larger projects. In our case we've been especially interested in opportunities to fund software development related to GLAMWiki-collaboration and language technologies.

We identified the Swedish Innovation Agency's [call for new CivicTech](#) to be a very promising such opportunity to pursue. The purpose of the call is to fund the development of *prototypes* that can enhance trust between (Swedish) citizens and the public sector. In the call text crowdsourcing is emphasised as a good way to build such trust.

After some internal brainstorming and idea generation we agreed on two project ideas. Both ideas had a strong emphasis on crowdsourcing and we believed them to be good candidates for the call AND that they could serve as vehicles to develop software that would serve the GLAMWiki-community well. We identified potential partners for the projects and contacted them. They proved interested to partner with us and so we decided to pursue not only one, but two applications.

On August 26 we submitted two applications¹: “Crowdsourcing along the Silk Roads” and “The WikiMap - old maps made anew, with and for all”. Please note that as per the instructions from the Swedish Innovation Agency the emphasis on the applications follow their template and **does not** focus on the technical design of the prototype(s) to be developed in the project. Focus

¹ The applications were written in Swedish. Included in this document are translations into English. Please excuse the occasional Swenglish!

is rather on how the prototype is intended to be able to be used to build trust between citizens and the public sector.

On October 28 we were informed by the Swedish Innovation Agency that neither of the applications were selected for a grant. There were 89 applications in total with 7 ultimately selected for grants. Of our two applications Crowdsourcing along the Silk Roads came closest to being selected, having been among the 12 applications selected for a second round of evaluation.

Crowdsourcing along the Silk Roads

This project application was made in partnership with the Swedish [Museum of World Culture](#) and the cultural organisation [Re:Orient](#). It proposes that inviting persons with a connection to countries from which the Museum of World Cultures has objects in their collection to contribute information and translations of existing information to those objects will increase their trust in the museum and other Swedish heritage institutions.

From a strictly technological point of view the scope of the application is to develop:

1. A *prototypical extension* to [OpenRefine](#) making it capable of reading and writing Structured Data on Commons. Reading SDC is important, not only to edit the metadata of existing files on Commons, but to fetch user-added metadata for export from OpenRefine and later import into a GLAMs database, so-called [data roundtripping](#).
2. A *prototypical extension* to, or fork/version of, the [ISA-tool](#). Focus would be to extend it in terms of localisation, improved campaign statistics, and to allow users to add all available SDC statements, not just depicts and captions. Effort would also go into making the visual and interaction design of the ISA-tool more appealing to institutions and potential micro-contributors both. Perhaps something more in the style of [Zooniverse](#) and other such platforms for crowdsourcing campaigns.

The WikiMap - old maps made anew, with and for all

This project application was made in partnership with the [Swedish National Archives](#). It proposes that working with, interpreting and annotating, historical maps in the classroom is a good way to engage young people with their heritage, with local and global history, and for pupils to learn how humans for a long time have shaped the environment.

From a strictly technological point of view the scope of the application is to develop a new more technologically up to date *prototype* version of the [WikiMaps Warper](#). The prototype would also be extended with capabilities such as being able to load and display any map available via [IIIF](#) and annotating regions and points on maps with Wikidata and Commons SDC-statements, e.g. linking named places on the map with their corresponding Wikidata objects.

The applications

Crowdsourcing along the Silk Roads

[Wikimedia Sverige](#) (WMSE), the [National Museums of World Cultures](#) (*Statens museer för världskultur*, SMVK) and [Re:Orient](#) (Re:O) have launched the project *Crowdsourcing along the Silk Roads*. The goal is to develop prototypes for people with knowledge about the cultures and languages in the areas along the Silk Roads to contribute information to images in the museums' collections: enrich the details, translate the descriptions or identify where the photos were taken.

With crowdsourcing we will be able to engage new communities with both Swedish and global cultural heritage, making it more accessible, usable and relevant – as well as increasing the quality of SMVK's collections data. The prototypes are available under open licenses and developed so that they can be used with any openly licensed image collection, both in Sweden and internationally, not just SMVK's collections.

Who, when and where?

Studies show that several social groups are underrepresented in museum exhibitions.² This is considered a contributing factor to the fact that some members of those groups don't perceive museums and their exhibitions as relevant to them. One of the underrepresented groups are Swedes with backgrounds in other countries. It is safe to assume that the same applies not only to exhibitions, but also to online museum experiences.

At the same time, there are collections in Swedish museums – items collected by Swedish explorers, scientists and missionaries – coming from countries and language areas that several migrant groups in Sweden feel a connection to. Sweden's cultural heritage is both local and global – most of all, it is diverse. However, most of the international collections in Swedish museums are described only in Swedish. This can make it harder for new Swedes to access and appreciate these collections, whether online or in physical exhibitions. It also makes it harder for users who do not speak Swedish to discover, study and analyze the collections.

In our project, we will invite people with diverse cultural and linguistic backgrounds to apply their knowledge to enriching and translating selected parts of the photographic collections of the National Museums of World Cultures. The project also includes releasing parts of these collections online for the first time. Our hypothesis is that by actively involving citizens in improving museum collections they are knowledgeable about, their trust in the museum will increase. Other museums, archives and libraries will be able to use the methods and tools developed in this project; they could also be adapted to other collections and other groups of citizens.

The project will increase knowledge about and interest in SMVK's collection among the "crowdsourcers". Our goal is to build long-term relationships and engagement, leading to the volunteers contributing to the collections also after the project has been completed. Since the

² See e.g. the report "[Att vidga sin publik handlar om att vidga sig själv](#)" (in Swedish), esp. pp. 31–42.

improved collections data will be immediately available on the Wikimedia platforms, millions of people around the world will be able to enjoy it. At a later point, a curated selection will be ingested into SMVK's authoritative database.

Published on the Wikimedia platforms, the photos will reach a much larger and more diverse audience. The contributions made in this project can also enrich physical exhibitions, for example by creating multilingual digital guides. In the short term, during the lifetime of the project, this will apply only to a curated small part of the huge collections housed by SMVK. However, cultural heritage institutions and the Wikimedia movement alike work with long-term goals; we believe that our prototypes and methods will be scalable and will gradually help make the collections of SMVK, and other cultural heritage institutions, richer and more multilingual – that is, more accessible.

Method and technology

Despite the fact that (Swedish) cultural policy explicitly states that citizen participation is a goal, crowdsourcing is rarely used in the digitization of cultural heritage. When it is, it is to a limited extent only. Much of the digitized material never reaches the citizens, or cannot be enjoyed by them due to language barriers. When it does reach the citizens, it is often a case of one-way communication; it is very unusual that the public is invited to become co-creators. This is a lost opportunity for increased civic engagement in cultural heritage; engagement that could be channeled, in partnership with the institutions, into enriching digitized cultural heritage with additional information and more nuanced perspectives.

Crowdsourcing enables citizens to engage directly with collections. New perspectives are encouraged and personal connections to the cultural heritage institutions and their collections strengthened. As citizens are empowered and institutions increasingly opening up to collaboration with citizens, the conditions for long-term preservation and relevance of cultural heritage improve.

For these reasons we will design and develop two prototypes that will make it significantly easier for both SMVK and volunteers/citizens to publish photographs on Wikimedia Commons and together improve the information about them by adding and translating metadata, adding geographical coordinates and so on. The prototypes will be developed and hosted on Wikimedia Foundation's shared development platform, giving us easy access to the Wikimedia API's. By using these API's and the Wikimedia movement's own development environment, we reduce the resources we need to invest in technical infrastructure. Instead, we can focus on good design and key functionalities of the prototypes themselves.

Additionally, using existing technical infrastructure improves our ability to maintain and further develop the prototypes after the project has ended (as it lowers our costs of operation).

The prototypes will make it significantly easier for users to:

- Create or improve structured, descriptive and multilingual metadata in Wikimedia Commons files. That could include e.g. identifying specific buildings and other depicted objects or narrowing down where the photo was taken.

- Translate image captions, descriptions and metadata into other languages, in particular the languages that the volunteers know and are used in the areas where the photos were taken.

We are also going to develop a prototype for SMVK to:

- Create, publish and track crowdsourcing campaigns on Wikimedia Commons.
- Add and edit structured metadata to files uploaded to Wikimedia Commons as part of large-scale uploads.
- Export the information added by volunteers on Wikimedia Commons in standard formats³, enabling SMVK to import them to their collections database.

Together, the prototypes will make it possible to create and carry out crowdsourcing campaigns on Wikimedia Commons with a focus on structured, multilingual metadata, as well as to export selections of the data for import into the museum's collection management system.

Selection and examples of material from SMVK

The photos selected for the project were taken in areas along the Silk Roads of Central Asia and the Middle East. Most of them originate from scientific expeditions and missionary activities. Some of the photographers included are [John Törnquist](#), [Sven Hedin](#), [T. J. Arne](#) and [Johan Gunnar Andersson](#). Not only do the photos speak to the imagination, they are also valuable for research in fields such as climate science, geography and anthropology. The majority were taken in the first half of the 20th century. They depict both urban environments and the countryside, giving an insight into communities and landscapes that have undergone huge changes in the 20th century.

One photo – many improvements

The photo below is a good example of the sort of material we are going to work with in the project. However, the metadata is unusually sparse. The motif is unknown, so we don't know where in Iran the photo was taken. What can volunteers do to improve its description and make it more accessible to others?

³ Such as CSV.



Fig 1. [Photo from the collections of SMVK](#). Unknown photographer.

They could:

- Identify the ruins in the photo. A reverse image search [hit in Flickr](#) suggests that it might be the ruins of the [Blue Mosque](#) in [Tabriz, Iran](#). Photos of this mosque on Wikimedia Commons confirm that's the case. Using this information, volunteers can:
 - Update the description (in several languages) and add multilingual structured metadata with this new piece of information.
 - Add coordinates to the photo.
 - Create a Swedish Wikipedia article about the mosque and improve the very short [article about Tabriz](#).
 - Add the photo to the Wikipedia articles about the Blue Mosque in Tabriz (e.g. in [English](#) and [Farsi](#)).

Roundtripping – a virtuous data exchange circle between institutions and the Wikimedia platforms

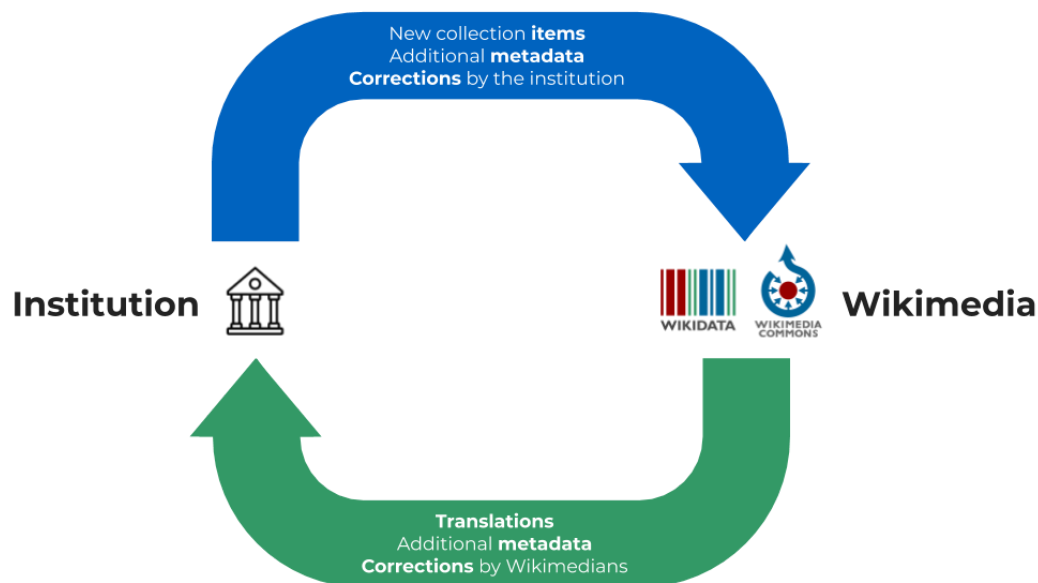


Fig 2. How data can flow in a “virtuous circle” between an institution’s collections and the Wikimedia platforms. By Sandra Fauconnier, [CC BY-SA 4.0](#).

Such additions and improvements of a photo’s metadata should be available for SMVK to export from Wikimedia Commons. SMVK should then be able to import them into their own collection management system. The Wikimedia movement calls this process “roundtripping”, and we believe it has the potential to create considerable added value for GLAM institutions that collaborate with the Wikimedia movement.⁴ Technical investigation of this possibility is very much in its infancy, and this project would be a major step forward in the development of roundtripping.

Innovation

This is the first time that the project participants collaborate on crowdsourcing of collections’ improvements. Because of this, we are developing new methods of collaboration. Our software will make it possible to involve new groups in learning from and developing the resources provided by the cultural heritage sector.

Citizen research as such is quite unexplored and underutilized⁵ in the Swedish cultural heritage sector. In addition, some of the methods and tools we are going to use are a trailblazing development in the field of crowdsourcing-based collaboration between cultural heritage institutions and the public. By this we refer to the development of tool support both for the “roundtripping” process and for creating and editing multilingual structured and machine-readable metadata for media files on a large scale.

There is currently no platform that is both open, free⁶ and available in Swedish⁷ that cultural heritage institutions and researchers could use for crowdsourcing or citizen science. By

⁴ See [this article](#) and the case studies it references.

⁵ See the [examples on Medborgarforskningsportalen](#) (in Swedish). Almost all are in the fields of natural science and nature conservation. The only cultural heritage project is by the Swedish Museum of Natural History.

⁶ Free to use, built on open source software and aimed at open data.

⁷ We know of one case where the Swedish Museum of Natural History ran a crowdsourcing campaign on Zooniverse, which has an English interface.

implementing the prototypes developed in this project, Wikimedia could become such a platform for institutions with collections of photographs or other images. Scaled up this has an enormous potential!

How?

Public sector openness to citizen participation results in increased trust in public institutions. In this project, we invite people who are knowledgeable about or have a personal connection to a specific cultural heritage collections to co-create digital information about them. In this way, we take openness one step further. Members of the minority groups that we invite to participate in the project will learn that their heritage and knowledge are highly appreciated, which will increase their trust in SMVK and, in the long run, other cultural heritage institutions.

This civic collaboration project will also make the objects in the museums' online collections more discoverable and useful. In the long run, this will increase the use of open cultural heritage collections, which in turn will strengthen the legitimacy of digitizing cultural heritage collections.

By having an open and inclusive proactive approach to the issue of problematic historic terms that can be found in the collections, we make it possible to systematically improve and contextualise such terms. This will give the public a deeper and more nuanced understanding of the collections. Our project and its crowdsourcing methodology is a natural way to involve persons most concerned with the issue in finding ways to resolve it.

Only a small part of these outcomes will be achieved in the lifetime of the project. Only after the project has been completed, when multiple cultural heritage institutions have established practices for collaboration with citizens and the Wikimedia movement, will far-reaching and sustainable results be achieved.

Ethics, safety and challenging norms

The project will have an equal gender distribution among the staff as well as strive towards an equal gender distribution among the test users. A targeted effort will be made to involve people living with disabilities and women belonging to the relevant minority groups in the development process.

The photographic collections we are going to work with come from the late 19th and early 20th centuries; a time marked by European imperialism and colonialism in the areas where the photos were taken. The motifs were selected by explorers, missionaries or visiting scientists and catalogued in Sweden by Swedish curators.

When we invite people who have a personal connection to the areas where the photographs were taken to describe them in their own languages, the photos are enriched with new perspectives. The dominating European perspective is challenged and problematized. We are going to develop the prototypes with internationalization in mind, including support for right-to-left languages such as Farsi and Arabic. These two languages will be especially prioritized as they are relevant to the selected collections at SMVK⁸. The user interface will be translated into more languages

⁸ This does not mean that other language groups won't be invited to participate, eg. Uighur- and Mongolian speakers will also be invited and localisation of our prototypes to those languages encouraged.

by Wikimedians, a community experienced in localizing volunteer developed software into multiple languages.

The project will also provide an opportunity to identify terms for ethnicities and religions that were used when the photos were taken or catalogued, but are problematic today. They will be contextualized and supplemented with modern terms.⁹

To ensure that the prototype for editing and translating file metadata is accessible to people living with disabilities, it will be designed and developed in accordance with the [WCAG 2.1 AA standards](#). We will evaluate the usability of the software based on heuristics of accessibility and usability.

Demonstration

We will demonstrate the prototypes and show how users can upload and edit multilingual structured metadata on Wikimedia Commons, one file at a time or in batch. We will also show how the data created by citizens can be exported out of Wikimedia Commons and made ready to be imported into SMVK's collection management system.

The demonstration will also include reporting the results of a usability analysis and present a suggested roadmap for future development. Furthermore, we will present the results of a qualitative study of the volunteers' attitudes to SMVK, an impact projection of the prototypes' if upscaled, and a set of good practices for how cultural heritage institutions with similar collections can collaborate with Wikimedians.

The prototypes will be available online, in beta versions, to everyone who wants to test them. Source code will be available from the first day of development.

Participants

Roles

The volunteers/citizen participants:

- Create or improve metadata about the photos on Wikimedia Commons.
- Create or improve Wikidata items in multiple languages.
- Create or improve Wikipedia articles in multiple languages.

WMSE:

- Coordinates the project.
- Designs, develops and tests the prototypes.
- Matches the volunteers who prefer to use languages other than Swedish with relevant Wikimedia chapters or user groups.
- Writes white papers/best practices based on the learnings from the project.

⁹ See the case study [EMPOWER](#) by WMSE which deals with such "problematic data".

SMVK:

- Makes a selection of images and other media files to be used in the project, including ensuring that they are free from copyright or can be released under an open license.
- Networks with groups that have a connection to or knowledge about the areas depicted on the photographs.
- Writes articles and presentations about the experiences from the project.

Re:Orient:

- Supports SMVK in reaching out to relevant minority groups.
- Supports SMVK and WMSE in organizing edit-athons.
- Participates in evaluating the prototypes with a specific focus on how to develop and scale up the collaboration with the minority groups after the project has ended.

WMSE and SMVK together:

- Upload images and other media, including structured metadata, to Wikimedia Commons.
- Test large-scale uploads of structured data to Wikimedia Commons.
- Test uploads of data created by citizens to the museum's collection management system.
- Evaluate the project's effects on the participants' trust in and attitude towards SMVK.
- Spreads the learnings from the project to relevant target groups.

Skills and capabilities**Wikimedia Sverige:**

- Has experience uploading photo collections to Wikimedia Commons,
- Has experience working with metadata from many projects.
- Has many partners who could be offered the prototypes.
- Has experience developing Mediawiki software and using Wikimedia APIs.
- Has access to a uniquely large and multilingual volunteer community.
- Can ensure that the values created by the project will be sustained after it has been completed.

National Museums of World Cultures:

- Has participated in projects with Wikimedia and contributed to Wikimedia Commons with its media collections
- Has experience collaborating with external actors (researchers) and importing enriched metadata to its own collection management system.

- Has worked together with several external organizations/associations, as well as schools, to enrich and give a more nuanced perspective on its collections.

Re:Orient:

- Has a network that enables us to reach out to the main target groups of this project.
- Has extensive experience of organizing interesting events focused on the culture and heritage of the Middle East and North Africa.

Utilization

WMSE [collaborates with several cultural heritage institutions that house collections of photographs and other media files](#). Our prototypes will enable us to further develop and strengthen those partnerships, as well as to initiate new ones. This project will lay the foundations to a generalized, common, open and free service where cultural heritage institutions can use crowdsourcing to, together with Wikimedians, enrich and internationalize their collections.

WMSE has access to national and international networks in which the prototypes can be disseminated and demonstrated. As we already have extensive experience of this type of work we expect the prototypes will be positively received by the global Wikimedia movement, leading to further collaborations between the world's cultural heritage institutions and Wikimedians. We're convinced that members of the Wikimedia community might be interested in contributing to the further development of the prototypes.

WMSE continuously and proactively works to inform, demonstrate and educate on how cultural heritage institutions can benefit from partnerships with the Wikimedia movement. Both during the project and immediately after its completion, we will intensify our efforts on this front, using the project as a driving case. See Activity A4: "Evaluation and knowledge sharing" below for more details about the utilization and dissemination of outcomes.

To make it as easy as possible for others to build upon our results, all the deliverables and project artefacts will be made available under open licenses. The source code will be licensed [GPLv3](#), [MIT](#), or [similar](#). All reports, evaluations, recorded seminars, and other documentation will be licensed [CC BY-SA](#). We and the users of the prototypes will waive all copyright for the metadata created or generated by the software so that it can be marked [CC0](#). The photos that will be uploaded to Wikimedia Commons as part of the project are either part of the [Public Domain](#) due to their age or will be released as [CC BY](#) by SMVK.

As for the technical infrastructure, the prototypes will be deployed in Wikimedia Foundation's [Toolforge](#) and [Wikimedia Cloud Services](#). Development will be managed and tracked in [Phabricator](#). This solution is much more cost effective than deploying the prototypes on platforms managed by ourselves or on commercial hosting services. Most importantly, the prototypes are clients to Wikimedia's official APIs – such as the Wikimedia Commons API – which are provided, managed and developed by WMF. This allows Wikimedia Sverige's own investment in the supporting infrastructure to be very limited. By managing the available resources effectively, we maximize our chances to see the prototypes being maintained and developed further after the project has been completed.

Activities, budget and planning

Title (of the activity)	A1. Planning and research
Start- and end date	M1-4
Short description	We create a project plan and do 2–3 Design Sprints as part of shaping it.
Cost and work hours	280h Project manager (WMSE); 160h Digitization coordinator (SMVK); 120h Image curator (SMVK); 120h Software developer (WMSE). Total: 680h 50 000 SEK consultancy costs (UX/UI-research and design) Budget: c. 250 000 SEK
Accountable	WMSE
Participating	SMVK
Expected results	A full project plan. It includes, apart from standard project artefacts, an initial campaign plan and a selection of photographic collections best suited for the project. UX/UI-design is at a stage where software development (of the functional prototypes) can begin.

Title (of the activity)	A2. Prototyping
Start- and end date	M5-14
Short description	Iterative prototype development with regular testing with users and stakeholder consultations. Tickers, documentation, and source code is made available from the beginning. We will also begin to batch upload media and metadata to Wikimedia Commons.
Cost and work hours	420h Project manager (WMSE); 160h Digitization coordinator (SMVK); 80h Image curator (SMVK); 2180h Software developers (WMSE). Totalt: 2840h. 50 000 SEK consultancy costs (technical design) 100 000 SEK consultancy costs (UX-research and UI-design) Budget: c. 1 200 000 SEK
Accountable	WMSE
Participating	SMVK
Expected results	One prototype to enable users to translate, add and improve metadata; One prototype to batch upload media files with

	associated linked open metadata; Significant selections from SMVK's collections published on Wikimedia Commons.
--	---

Title (of the activity)	A3. Full-scale prototype testing
Start- and end date	M10-14
Short description	In this period the prototypes will be available online for test use and we will systematically collect user feedback. Heuristic reviews of usability and accessibility will also be performed. As collaborative forms of work are central to the Wikimedia movement we will arrange "editathons", meetups where volunteers using the prototypes work together and with access to museum expertise, to enrich the photographic collections. At each editathon we will collect user feedback and perform user interviews.
Cost and work hours	120h Project manager (WMSE); 160h Digitization coordinator (SMVK); 110h Image curator (SMVK); 100h Software developer (WMSE); 120h Producer (Re:Orient). Total: 610h 50 000 SEK consultancy costs (UX-research) 50 000 SEK event expenses (Re:Orient) Total: c. 250 000 SEK
Accountable	WMSE
Participating	SMVK, Re:O
Expected results	We have arranged at least two major editathons. We have reports on usability and accessibility along with recommendations for future development.

Title (of the activity)	A4. Evaluation and knowledge sharing
Start- and end date	M13-15
Short description	We collect and collate results from the project and prototype testing in order to identify lessons learned, knowledge gained and future opportunities. We communicate the results and publish a good practice for GLAMWiki-collaborations of similar nature.
Cost and work hours	160h Project manager (WMSE); 160h Digitization coordinator (SMVK); 120h Image curator (SMVK); 160h Software developer (WMSE); 24h Producer (Re:Orient). Total: 624h 50 000 SEK consultancy costs (impact assessment) Budget: c. 300 000 SEK

Accountable	WMSE
Participating	SMVK, Re:O
Expected results	An analysis of how participation has affected the trust the volunteers feel towards SMVK and heritage institutions, An impact assessment of a scaled-up scenario of use for the prototypes; A published good practices document for heritage organisations and Wikimedians and an open seminar about the good practices with upwards of 50 participants.

The WikiMap - old maps made anew, with and for all

Who, when, where?

The Swedish landscape has, since the end of the ice age, been shaped by humans: forests have been burned and cut down, wetlands drained, rocks gathered into cairns, lakes lowered, rivers regulated, new forests planted, and wetlands recreated. Today, in the Anthropocene, the landscape is changing faster than ever.

Historical maps provide a unique opportunity to understand how the landscape has changed. They make it easier to understand, visualize and analyze the landscape. In addition, place names in the maps are important for understanding an area's cultural- and natural history. Maps also illustrate how the state over time has tried to gain control of the country and its inhabitants including how indigenous peoples have been displaced, linguistically and physically.

Sweden has a world-unique collection of historical maps from the 17th century onwards. Here, a nationwide mapping took place before the modern agricultural reforms and urbanization. From Tornio in the north to Trelleborg in the south, villages, farms and arable land were measured and depicted on detailed village maps. In addition, there are a large number of city maps and maps of larger administrative divisions, such as parishes, counties and provinces. In total, for the period 1630–1930, there are over 800,000 maps in the national archives.

Requirements analyses carried out by the Swedish National Heritage Board¹⁰ show that historical maps are highly valued by physical planners at the county and municipal levels. However, the analyses also show that access to the map images alone is not enough. In addition to simply access to the maps, physical planners want them to be:

- Warped (geocoded), i.e. they can be overlaid on top of modern maps.
- Easily readable with place names and other texts on the maps interpreted, transcribed, and searchable.
- Free to use without license fees.

Proven technology is available to improve the maps in accordance with these requirements. But the country's map collections are extensive and the manual processing necessary is time consuming and thus costly. This applies in particular to warping historical maps against today's modern maps. It would also be easier to find and use historical maps if they were better classified and enriched with more metadata.

In this project, we will lay the foundations for an open technical platform prototype where users together can process, enrich and interpret historical maps. Full-scale tests of our prototype will be carried out with several target groups: people with an interest in maps, family and local history, minority languages and history, together with school children (at upper primary and secondary school levels). The focus will be on maps of areas with minority populations, maps of former provinces and colonies, as well as maps that can show drastic environmental changes as a result of human activity. This provides an opportunity to bridge generational gaps, engage young

¹⁰ See [Behov av kulturmiljöinformation hos Riksantikvarieämbetets användare](#) (PDF, in Swedish)

people in cultural heritage work, and work with central and topical content in the [national curriculum for the history](#), e.g.:

- History use, source criticism, and the understanding how archives in general, and maps in particular, are not neutral historical sources.
- Historical perspectives on the situation of the indigenous Sami and other national minorities in Sweden.
- Sweden of the great power era as well as European dominance, imperialism and colonialism.

The long-term vision is to establish a dynamic link between the map archives in the National Archives (NA) and the Wikimedia platforms. Every named place in Sweden, even those that have been abandoned, should have a Wikipedia article where historical maps can be added to the article. Hand-drawn maps and printed map series will be made available as Wikidata objects. Thus geocoded, warped, with identified place names, and added rich descriptive metadata the maps' usefulness and value for amateur research, academic research, physical planning, and school education will be greatly increased.

Method and technology

Despite great public interest in cultural heritage and national cultural policy encouraging increased citizen participation, crowdsourcing is rarely used in digitization of cultural heritage. Large parts of historical maps that are digitized don't reach the citizens because the materials are too difficult to find. This is a lost opportunity for increased citizen participation in cultural heritage work, participation that could also be directed to making digitized map collections richer and more discoverable.

We will design, develop, and evaluate a prototype that makes it easy for volunteers-citizens to warp, describe, and annotate historical maps. The prototype will be developed and deployed on the Wikimedia movement's technical platforms, primarily Wikimedia Commons. By relying on open source code and by using [Wikimedia's infrastructure](#), we save a great deal of development work. The prototype is intended to enable users to:

- Warp and geocode historical maps. Warping of historical maps means that they are "fitted in" with contemporary maps and thus coordinates are given.
- Geotag and annotate places and individual objects on maps, such as churches and other buildings, with Wikidata identities. This enables the addition of place names in several languages and linking them to gazetteer entries as well as its corresponding Wikipedia articles and media representations in Wikimedia Commons.
- Classify different types of maps and supplement or correct information, e.g. with information about year, surveyor, scale, title, etc. when such metadata is missing.
- Import and export of historical maps with associated metadata "in batch" from/to Wikimedia Commons and the National Archives systems.

Through crowdsourcing, citizens come into direct contact with their heritage. Increased participation, and increased institutional openness towards digital civic cooperation, strengthens the foundation for the long-term preservation of cultural heritage. Specifically in this project, the participants' involvement in the preservation of cultural heritage in archives as well as landscape history and political history will be strengthened. By directly connecting historical landscapes,

borders, domestic and external colonization, with topical events today, we believe that (young people) can become interested in cultural heritage and thereby develop a stronger trust in National Archives (and by extension the entire cultural heritage sector).

Selection of map collections

To develop the prototype, this project will focus on historical maps from two major collections at the National Archives.

1. [Sweden's oldest large-scale village maps from the 17th century](#)

With about 15,000 17th-century maps of villages and farms this is the largest collection in the world of its kind and time period. All the maps have been digitized, transcribed and geocoded in research projects at National Archives. Existing metadata and geodata can therefore form the basis for the prototype that this project develops. Published on the Wikimedia platforms, the material can be enriched with links to place names for villages, parishes, etc. These maps were a direct result of the growing Swedish state of the time and its efforts to control people, resources and landscapes. They are therefore well suited for analyses of geopolitics, taxation, and the development of nation states and empires.

2. [Military Archives maps and drawings](#)

This collection contains maps of many different types: city plans, sketches, printed maps, charts, world maps, etc. In addition to maps of Sweden and Finland, it also contains a large number of foreign maps. About 35,000 maps in the collection have been digitized. The associated metadata includes archivists, authors, and years, but is otherwise sparse. The maps are also not classified based by type of map and nor is there any associated geodata. A large number of maps from this collection are therefore well suited for warping, geocoding and addition of descriptive metadata.

An example of how the information on a historical map could be improved

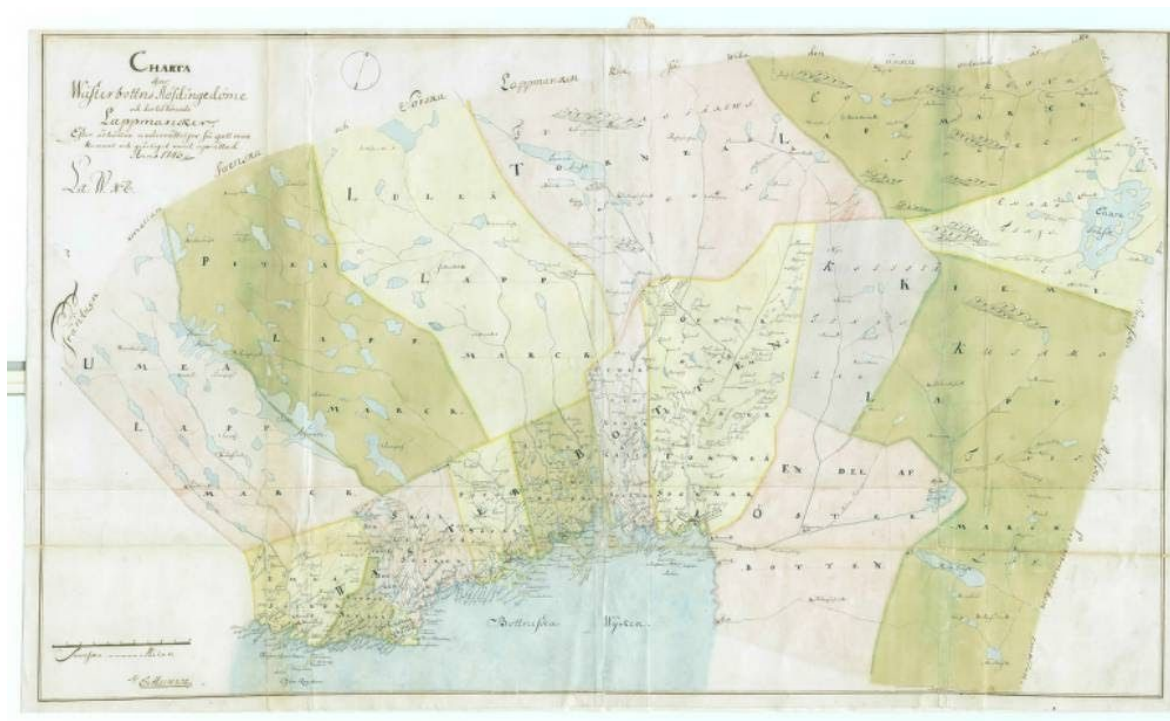


Fig 1. "[Charta öfwer Wästerbottens Höfdingedöme och dertil hörande Lappmarker](#)". From the collections of the National Archives.

How can volunteers make this map more discoverable and useful?

They can:

- Warp it so that it can be overlaid on modern maps of Sweden and Finland.
- Annotate it, identifying settlements, but also other places (eg. lakes and rivers) and buildings (churches, fortifications, etc.), linking them to their corresponding Wikidata objects (or creating them if they do not already exist), and adding labels in several languages, eg. Lule Sámi, Northern Sámi, and Meänkieli.
- Transcribe the map's title and description (top left) and thus make the contents of them indexed for search. The title of the map can also be translated into different languages for better indexing.
- Research and identify who created the map, and link them to person-objects in Wikidata (or create one). In this case, the cartographer/surveyor, whose signature appears on the map, can be identified as [Esaias Hackzell](#). Information about the author, context and purpose for which maps have been created facilitates interpretation and enables new dimensions of discovery.

In a next step we would export the warped map, the annotations, and additions of other metadata from Wikimedia Commons and Wikidata to import into or link to the NA's system. Within the Wikimedia movement this is called "data roundtripping" and is considered to have great potential to create added value¹¹ for [GLAMs](#). This, because their collections data can then be improved at source through collaboration with citizen-volunteers. Technical solutions for

¹¹ See e.g. [Data Roundtripping: A new frontier for GLAM-Wiki collaborations](#) and the analyses and reports the article references.

roundtripping have only just begun to be investigated. This project would represent a significant step forward in the development of such capabilities, in this case for the first time applied to a very specific collections that also requires specialised technical solutions.

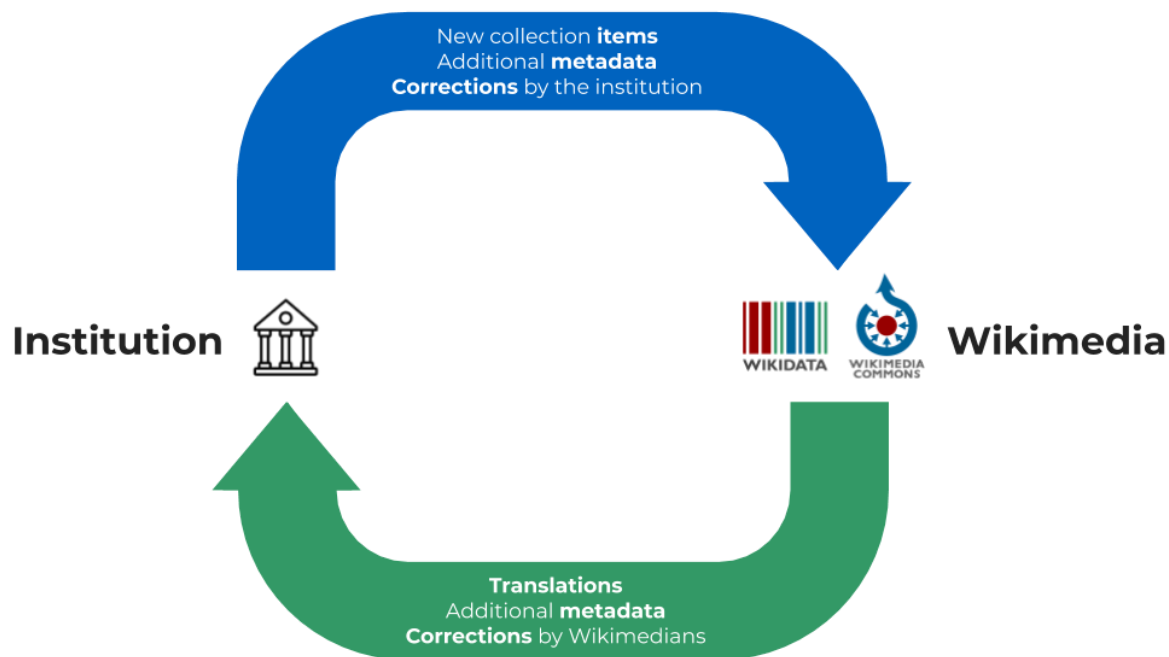


Fig 2. Sketch of how media files and associated metadata can flow, in a “roundtrip”, from institution to Wikimedia and back again. Author: Sandra Fauconnier, [Some rights reserved](#).

News value

Citizen research, crowdsourcing, and GLAMwiki collaboration are in themselves not new. However, the vast majority of cultural heritage institutions in Sweden, and globally, do not collaborate with either citizens or the Wikimedia movement in their digitization activities.

There is currently no open, shared and free platform on which Swedish cultural heritage- or research institutions can conduct crowdsourcing with maps as material. Wikimedia's platforms, with the prototype developed in this project, can become this open and free platform, in Sweden and globally.

Some specific methods and techniques we intend to use in the project are cutting edge. This refers to both warping maps in the browser and semantically aware geographical annotation of places with Wikidata and its “knowledge graph”.

Today, only a relatively few Swedish historical maps are available on Wikimedia Commons, Wikipedia, and Wikidata. As part of this project, large amounts of maps will be made available on those platforms and a foundation laid for the long-term publication of all NA's maps suitable for the purpose. This means that large amounts of formally openly licensed maps become more usable and accessible to new and different target groups, also internationally, than the NA can reach by itself.

How?

In our project, people will be invited to coordinate and enrich (annotate, add metadata) historical maps. Our hypothesis is that by inviting citizens to collaborate with the purpose to improve the collections of NA, citizens' trust in this government agency can be increased. The method, and the prototype, is applicable and useful to other institutions with collections of maps (including National Land Survey-organisations), both nationally and internationally.

We see opportunities for young people to use crowdsourcing and digital tools to understand, get involved in and engage with their cultural heritage. Young people are otherwise an underrepresented group in local history associations. Our project contributes to a solution by opening up for digital engagement with maps that are topical, such as domestic and external colonization. We believe this can increase engagement, create interest and act as a catalyst for deeper investigations into Swedish history.

To ensure that as many young people as possible discover the tool and the collections of the National Archives, we will ensure the prototype is suitable for both educators and historical associations. In this way, engagement with and trust in our cultural heritage institutions can be built over time.

If the prototype is successful, targeted efforts towards various local communities and minority groups are planned to be implemented, thereby increasing the confidence of these, often ignored, groups in cultural heritage institutions.

Ethics, safety, and critique of pervasive norms

We will ensure an even gender distribution in the project and also strive for an even gender distribution among the test participants. We will also involve people with disabilities in the testing.

Historical maps are not only a source of what the landscape looked like, but also show how governments at different times viewed and shaped our societies. In our map editathons we will therefore work with a selection of maps from areas where national minorities live (or have lived) or which Sweden has ruled over or colonized. This will problematize the "Swedishness norm" in a constructive manner. As women are under-represented in the Wikimedia community, we will make targeted efforts to reach and engage women, both online and in writing offices. By working with school classes, we can also even out and compensate for the over-representation of men, middle-aged and older, in local history societies.

To make the prototype usable for people with disabilities, we will develop it to be WCAG AA-compatible. We will test its usability for users with motor disabilities and less severe visual impairments.

Demonstration

We will demonstrate the prototype and show how users can warp, geocode, and annotate historical maps and add metadata to them. We will also show how citizen-created data can be imported into and enrich NA's source database. Source code and documentation will be available online under open licences. A usability analysis and a roadmap for desired future development will also be made available.

Furthermore, we will present the results of a qualitative investigation into the volunteers' attitudes to and confidence in the National Archives. An impact assessment of the prototype's potential for once scaled up will also be included.

Partners and stakeholders

Roles

Volunteer-citizens

- Geocode and warp maps
- Identifies and annotates named places in the maps
- Improves or translates the metadata of the maps
- Creates or improves Wikidata objects for cartographies, settlements, and historical administrative divisions

WMSE

- Coordinates this project
- Designs and develops the prototype and tests it with users
- Writes white papers and best practices for how the Wikimedia movement and institutions can work together to enrich and publish map collections

NA

- Selects suitable maps for the volunteers to work with
- Defines information/metadata templates for the maps
- Writes papers and communicates in other ways their experience from the project
- Establishes contact and collaboration with other institutions and research groups who work with map

WMSE and NA together

- Upload the maps to Wikimedia Commons
- Develops an import function to NA's source database
- Plans and executes map editathons, "mapathons"
- Creates guides for teachers to create learning resources based on maps
- Evaluates the results of the "mapathons" and the prototype and their effect on citizens' trust in the NA

Reference group

We intend to put together an advisory reference group of people from cultural heritage institutions, the Wikimedia movement, teachers, and representatives of organizations with collections of historical maps, or who simply have a great interest in historical maps.

Skills and capacity

WMSE

- Has run multiple projects with a focus on geodata and historical maps

- Has good experience with metadata management in a large number of projects and has worked with structured data on Wikimedia Commons
- Has a large number of partners in cultural heritage which we could invite to also test the prototype
- Has uploaded large amounts of media files to Wikimedia Commons and metadata to Wikidata
- Has experience developing software clients to the Wikimedia APIs
- Has access to a large international community of volunteers in the global Wikimedia movement
- Will secure the sustainability of the project outcomes

NA

- Has been responsible for many research projects with a focus on historical maps
- Has digitised large amounts of maps
- Makes available [TORA](#), a register of villages and farms in Sweden. The service is built on linked open data and provides geo-spatial search and discovery.
- Makes available the [National Database of Archives](#), which contains information about archives in Sweden as well as 65 million pieces of archive imagery.
- Has a unit for digital experience with a mission to develop digital services, ideally in collaboration with external partners
- Has deep and broad archival knowledge, research and academic publications capabilities and expertise in older sources, including historical maps.
- Has a good network and working relationship with academic research, genealogical societies and software companies, and local history societies.

Utilization

The platforms supporting the prototype are managed and developed by the Wikimedia Foundation and volunteers within the Wikimedia movement. Tools / applications developed on top of the platforms are managed by the Wikimedia community and can be deployed on [Wikimedia Cloud Services](#).

The prototype developed in this project will be managed by WMSE and volunteers in the Wikimedia movement with an interest in maps. The movement is global and works continuously and proactively to create content partnerships together with institutions. We expect that the prototype will be used by the Wikimedia movement and institutions with maps not only in Sweden, but all around the world!

All Wikimedia platforms are developed with open source code, have open APIs, and contain only open data - as does our prototype. This makes it possible for anyone to download, or via the APIs integrate, the historical maps into their own systems - something we will also proactively communicate to e.g. the county administrative boards, with their many cultural and natural environment officers, and various university departments, e.g. cultural geography, history, and languages.

WMSE works continuously to inform about and demonstrate how institutions, often in partnership with schools and using open learning resources (OERs), can work together with the

Wikimedia movement. During the project, and immediately after, we will, together with NA, intensify this work and based on experience and lessons learned from the project publish two “good practices”. One on how institutions with collections of maps can work with the Wikimedia movement, local and local history associations and another on how teachers can create learning resources centred on historical maps.

Activities, budget and time plan

Title	A1. Planning and research
Starting- and end-date	M1–4
Short description	This may also include early prototyping, user research and technical design.
Cost and work hours	120h Project Manager (WMSE); 40h Business Developer (WMSE); 80h Research Manager (NA); 200h Software Developer (WMSE); 40h Systems Developer (NA). Total: 480h 100 000 SEK consultancy costs (UX-research and UI-design) Budget: c. 250 000 SEK
Accountable	WMSE
Participants	NA
Expected results	A detailed project plan. It includes, apart from standard project artefacts, an initial campaign plan as well as a selection of maps most suitable for the campaign. UX-design and technical design are at a level where prototypy development can begin.

Title	A2. Prototype development
Starting- and end-date	M5–14
Short description	Iterative software development with regular feedback from testing with users and stakeholder dialogues. Tickets, documentation and source code will be publically available from the first beginning.
Cost and work hours	360h Project Manager (WMSE); 40h Business Developer (WMSE); 80h Research Manager (NA); 2 280h Software Developer (WMSE); 240h Systems Developer (NA). Total: 3 000h 100 000 SEK consultancy costs (UI-design and UX-testing) Budget: c. 1 250 000 SEK
Accountable	WMSE
Participants	NA
Expected results	A working prototype enabling users to warp, geocode, and annotate maps on Wikimedia Commons as well as add metadata to them.

Title	A3. Full scale prototype tests
Starting- and end-date	M9–12
Short description	<p>We will organize "mapathons", occasions where users together and with access to guidance from WMSE and NA, work to use the prototype in different ways to enrich maps. They serve as full-scale tests of the prototype where we can observe and collect participants' feedback and also conduct participant interviews. This feedback will be integrated in both the prototype development and the final evaluation (i.e. the previous and next project activities).</p> <p>The “mapathons” will be arranged at events such as Kartdagarna, Släktforskardagarna, Arkivens dag och Hack for Heritage, which allows us to disseminate our results and lessons learned to institutions with map collections of their own.</p>
Cost and work hours	<p>120h Project Manager (WMSE); 220h Business Developer (WMSE); 80h Research Manager (RA); 160h Software Developer (WMSE). Total: 580h.</p> <p>25 000 SEK consultancy costs (User research) 25 000 SEK expenses for “mapathons” and tests in school context</p> <p>Budget: c 300 000 SEK</p>
Accountable	WMSE
Participants	NA
Expected results	We have arranged a minimum of 4 mapathons (of which half online) with participants from schools. We have also arranged at least 1 mapathon at a major event.

Title	A4. Evaluation and knowledge-sharing
Starting- and end-date	M13–15
Short description	We collect and analyse the results of the prototype testing and other project activities in order to identify lessons learned and future opportunities. We publish and communicate these results and based on them we create a good practice for crowdsourcing improvements to historical maps.
Cost and work hours	<p>80h Project Manager (WMSE); 40h Business Developer; 80h Research Manager (RA); 80h Software Developer (WMSE); 40h Systems Developer (RA). Total: 320h</p> <p>75 000 consultancy costs (Impact analysis and User research)</p> <p>Budget: c. 200 000 SEK</p>
Accountable	WMSE

Participants	NA
Expected results	An analysis of how the effects on the participants level of trusts in the NA; A projection of the impact of the the prototype if scaled-up; A good practice aimed at institutions and GLAM-Wikimedians; A good practice aimed at teachers; An open seminar about the good practices with at least 50 participants; A usability- and accessibility analysis of the prototype.