

Tools for partnerships
developing the technology
our partners need
Wikimedia Sverige
and Wikimedia Foundation
2019-2020



Session info on the Wikimania wiki:

https://wikimania.wikimedia.org/wiki/2019:Technology_outreach_%26_innovation/Tools_for_partnerships_%E2%80%93_developing_the_technology_our_partners_need

Etherpad with notes:

https://etherpad.wikimedia.org/p/Tools_for_partnerships_developing_technology



John Andersson Wikimedia Sverige
john.andersson@wikimedia.se

Sandra Fauconnier Wikimedia Foundation
sandraf@wikimedia.org

John Andersson is executive director of Wikimedia Sverige.

Sandra Fauconnier is Program Officer at the GLAM team at the Wikimedia Foundation.



Introduction

Welcome everyone!

This is a presentation that targets people that have been active in the movement for a while.



A new initiative

- What is the situation in technology for content partnerships / the GLAM-Wiki community of practice today?
- How do we want to address this situation?
 - Governance
 - Technical
- Plenary discussion

4

Overview of this presentation

- We start with an introduction of the **context**. What is the situation in content partnerships in the Wikimedia movement today, and the technology built for it?
- Next, we'll explain our plans on how to **address** this situation.
 - First, we'll look at how we want to address this from a **governance** perspective and how we want to respond to the Wikimedia **movement strategy**
 - Second, we'll give more insight in our current thoughts about **technical scope and timeline**.
- Then we open the floor to your responses, questions, and comments.

Content partnerships in the Wikimedia movement

Demand from Wikimedia communities:

- Universal interest in **cultural heritage** across languages and borders.
- A well-established **GLAM-Wiki community of practice** (~10% of Commons, large portions of Wikidata/Wikisource)

Demand from content partners around the world:

- Wikimedia as **access and engagement platform** for digital collections

6

Content partnerships are **popular in the Wikimedia movement**.

Around the world, we see that our **Wikimedia communities** are very interested in them.

- The interest in **cultural heritage** is universal.
- Around the world, we have an established GLAM-Wiki community of practice. We can safely say that around 10% of the 50 million files on Wikimedia Commons have been contributed through content partnerships. Large parts of Wikidata and Wikisource have also been contributed through content partnerships.

Content partners and institutions around the world are also very interested in using Wikimedia platforms as **access and engagement platforms** for their digital collections.

But... technical infrastructure?

Pattypan
PetScan BaGLAMa
PagePile
GLAMwiki Toolset
GLAMorous
OpenRefine
Pywikibot
QuickStatements
GLAMorgan
Mix'n'match
VisualFileChange

Not sustainably built and supported

- *Ad hoc*
- Often volunteer-built
- Single point of failure
- High learning curves
- Not well integrated into Wikimedia platforms; difficult to find

However, the technical infrastructure for these partnerships is not ideal.

Technology for large-scale / batch contributions to Wikimedia projects is usually not sustainably built and supported.

Some of the pitfalls include:

- They were built ad hoc - as prototypes or custom solutions for single organizations or goals, not with the long term in mind
- They are often built by volunteers. If only one person maintains a crucial tool, there will be a single point of failure
- Many tools are poorly documented and have high learning curves.
- Most are standalone or external tools, not integrated into the software of Wikimedia Commons, Wikidata and other Wikimedia projects. Probably relatedly, many community members and content partners don't even know that certain tools exist. They are sometimes only known by expert users.

**An one year investigation:
July 2019 till June 2020**
Wikimedia Sverige and the
Wikimedia Foundation
team up, to start
addressing this issue.
How?

In order to address this issue, Wikimedia Sverige and the Wikimedia Foundation are teaming up.

We will work together for one year - July 2019 till June 2020 - to start working towards a better solution for this issue.

How will we go about this?

First, we will look at the context in which we will do this work - the Wikimedia movement strategy - and next, we will look into our first thoughts on technology and planning.



Movement Strategy observations

We are observing an interesting polarity between desire for **decentralization** and a call for more robust **common frameworks** to support movement work.

There are now a number of working groups that have been tasked on developing recommendations for how we can reach the Strategic direction. And how we can become the infrastructure and ensure that anyone can join us.

There are two trends within the current set of **recommendations** that are interesting to observe and that will have major implications on *who* will do *what* in our movement. We see that a strong trend is to underscore the need for **decentralization**, that not all should be handled by one central body, which is currently the Wikimedia Foundation, but that ownership should be **shared** across organizations across the world.

The other trend is the need for a **common framework**, and that support is given through capacity building and joint structures. This might sound like there is a **conflict** with decentralization at first, but what is envisioned is that the **power and independence** will increase amongst many actors so that they can develop this common framework within a given area, develop the **supporting** material, methods and structures for all of the movement.

Many of the working groups are suggesting to develop organizations that have a specific expertise and that will help move Wikimedia's work forward within a given thematic area. This initiative aim to pilot the implementation of the recommendations as they now are becoming much more concrete.

Testing a new model for growth

Creating a blueprint for a **larger role for affiliates** to:

- support **community capacity**
- increase and diversify **content**
- partner with, grow and engage **free knowledge actors** (chapters, user groups, thematic organizations, other organizations)

The long-term vision: **a network of affiliates around the world** serving as nodes of expertise, community engagement, and provide specialization

12

So what could this look like?

We want to start investigating how this could be structured, financed, coordinated and governed. What we learn will be documented and shared so that the knowledge gained allows other organizations to follow suite. The intent is to create kind of a blueprint that others can use to follow without having to invest as much.

Can Wikimedia Sverige find a new role in the movement that is accepted and supported by other movement actors? A role where we can help to develop local community capacity across the world so that not everyone has to start from scratch.

The goal is that the work that we do ensures that more content can be added to our platforms and that it become more diversified. The investments we make, should reduce barriers so that more people can take part. This underpins this approach.

An interesting part of the work during the coming year is to try to find a **decision making system** so that different free knowledge actors can help decide what should be prioritized. We believe that this co-ownership is crucial and we will put in a lot of effort to achieve it.

What we hope is that the work we do the coming year will allow for the creation of a **network of affiliates around the world** serving as nodes of expertise, allowing for community engagement, and providing specialized knowledge and capacities.

Need for support

The GLAM-Wiki community of practice needs a **technosocial service environment for collaboration:**

- infrastructure and tools for large-scale content contribution
- community training and documentation
- sustained financial resources
- networking support

13

There have been a lot of great work by the GLAM-wiki community over the years. “GLAM” is here broadly defined; besides collaborations with 'typical' cultural and knowledge institutions such as galleries, libraries, archives and museums, it also includes content such as private collections, bottom-up initiatives, community archives and so forth.

But we believe that there are still a lot of things that need to be improved to be able to scale across the world in the way we all envision. We believe that there is a need to develop infrastructure and tools for large-scale content contribution, which is what we will focus on during today's presentation and discussion.

But we also think that this need must be coupled with community training, with sustained financial resources, with documentation and with networking support. This is what we call a **technosocial service environment for collaboration**.

Developing this technosocial service environment is something we will integrate in our planning this year.

Why Wikimedia Sverige?

- **Experience** leading GLAM content partnerships **locally** and internationally (small and large GLAMs)
- Established **global GLAM content partnerships** (e.g. Europeana, UNESCO)
- **Technical capability** in tools and workflows for GLAM
- Close working relationship with WMF around **Structured Data on Commons**
- **FindingGLAMs** project has given us insights
- Possibility to find **external funding** to scale in a sustainable way

14

So why are Wikimedia Sverige doing this?

We have a lot of experience leading GLAM partnerships locally and internationally and have been working in the area for 12 years. That includes a number of partnerships with established global GLAM content partners, such as Europeana and UNESCO.

The last few years we have developed a small team of skilled developers that have the technical capability to support GLAMs with tools and workflows. They know MediaWiki development, how to build upload tools, develop APIs and much more. Most importantly we have invested heavily to ensure that they have a very deep understanding of how the Wikimedia movement works and of the many spoken and unspoken rules in our community.

The last couple of years we have developed a close working relationship with WMF around SDC as we have identified a number of opportunities together and overlapping interests, so we already know each other really well.

As part of the FindingGLAMs project we have developed a number of case studies outlining areas where there are specific needs due to lack in technology, a lack of documentation or no clear agreement on how to organize the work.

We also think that we will be able to find **external funding** to scale in a **sustainable** way. Sweden is a rich country with the possibility for us to get funds both from grants

from within the country or from the European Union, or to get donations from individuals or organizations.

Why Wikimedia Foundation?

- **GLAM team in the Community Engagement department**, with international connections and community expertise
 - GLAM pilots for the new **Structured Data on Commons** technology
- WMF is responsible for the **MediaWiki software**, so coordination is needed

15

So why the Wikimedia Foundation?

Well, in the Community engagement department, the Foundation has a skilled and knowledgeable GLAM team that has both international connections and community expertise. Their involvement in investigating what the needs are amongst GLAM actors across the world will help guide the work. But that's not all.

The SDC project, that is spearheaded by the Wikimedia Foundation, will have great value and impact of this initiative. Structured Data on Commons lays the technical foundation for working systematically with GLAM collections.

Wikimedia Foundation is also responsible for the MediaWiki software, so coordination is needed if Wikimedia Sverige starts building MediaWiki software. There need to an understanding within the Foundation why this is a priority.

Case studies

- **Structured Data on Commons (WMF):**
typical GLAM-Wiki collaborations. E.g.
 - An art museum upload to Wikimedia Commons
 - Digitization and upload of a private book collection
- **FindingGLAMs (WMSE):**
experimental / tricky GLAM-Wiki collaborations. E.g.
 - Connecting audio files with images and note sheets
 - Handling old racist content in historical collections
 - Uploading old interpreter dictionaries as lexemes

16

We can build upon the experience gained from a number of case studies we have created or are in the process of developing as part of the Structured Data on Commons project and the FindingGLAMs project.

Structured Data on Commons has run pilot projects for various 'typical' and often occurring GLAM-Wiki collaborations;

- for instance a typical upload of files from the collections of an art museum
- or the upload of a digitized private book collection

FindingGLAMs on the other hand focuses on tricky cases/experimental GLAM-Wiki collaborations where we describe the process and what is missing to do it better in the future. E.g.

- How we can connect audio files with images and note sheets on the different Wikimedia platforms
- How we can handle old racist content in historical collections that are made available on an open platform such as ours, or
- Uploading old interpreter dictionaries as lexemes on Wikidata

The findings in these case studies is an important foundation for this work that we are now aiming to do, and it should allow us to start working as early as possible, even though we know that there is a lot of research still to do. So even though this initiative has just started, we are certainly not starting from scratch.



Technology Research What will happen

Next, I will give an overview of

- our current thinking about the technical scope of this initiative
- the research that we plan to do this year
- what we intend to produce by June 2020



Making sense of scope and audiences

First we'll circle back on the statement of scope we made at the beginning of this presentation.

What is our current thinking on

- The scope of this initiative? What will it cover?
- For which audiences will we work?

Three levels in the communities of practice

1. Aspiring '**systemic**' partners (e.g. The Met, Europeana)
2. **Smaller** institutions (with limited digital capability)
3. '**Non-institutional**' (community archives, private collections etc.)

This covers partnerships that happen **around the world**.

The GLAM-Wiki community of practice also includes content partnerships with organisations **outside of the traditional GLAM definition** (academia, NGOs, IGOs, NASA etc.)

19

In order to make sense of the kind of content partnerships we are talking about, we have outlined **three categories** of them:

1. **Large-scale** organizations like the Met, Europeana, that usually have advanced digital infrastructures and who have expressed the need to exchange content with Wikimedia projects in a **continuous and systematic** way
2. **Smaller-scale** partner organizations - probably the majority of our movement's partners - that often have limited digital capacity and will have less frequent engagement with Wikimedia projects and communities
3. A category that we call '**non-institutional**' and that covers a lot of the bottom-up work by our communities around the world: work with small-scale community archives, private collections...

With these three categories, and especially the latter two, we want to emphasize that our focus is on the entire world; we want to be able to support content partnerships with Wikimedia projects wherever they take place.

We talk about the GLAM-Wiki community of practice with a generous mindset: the infrastructure we want to build, will also support partnerships that are not strictly considered GLAM, such as partnerships with universities, research institutions, NGOs, governments...

Users of the infrastructure

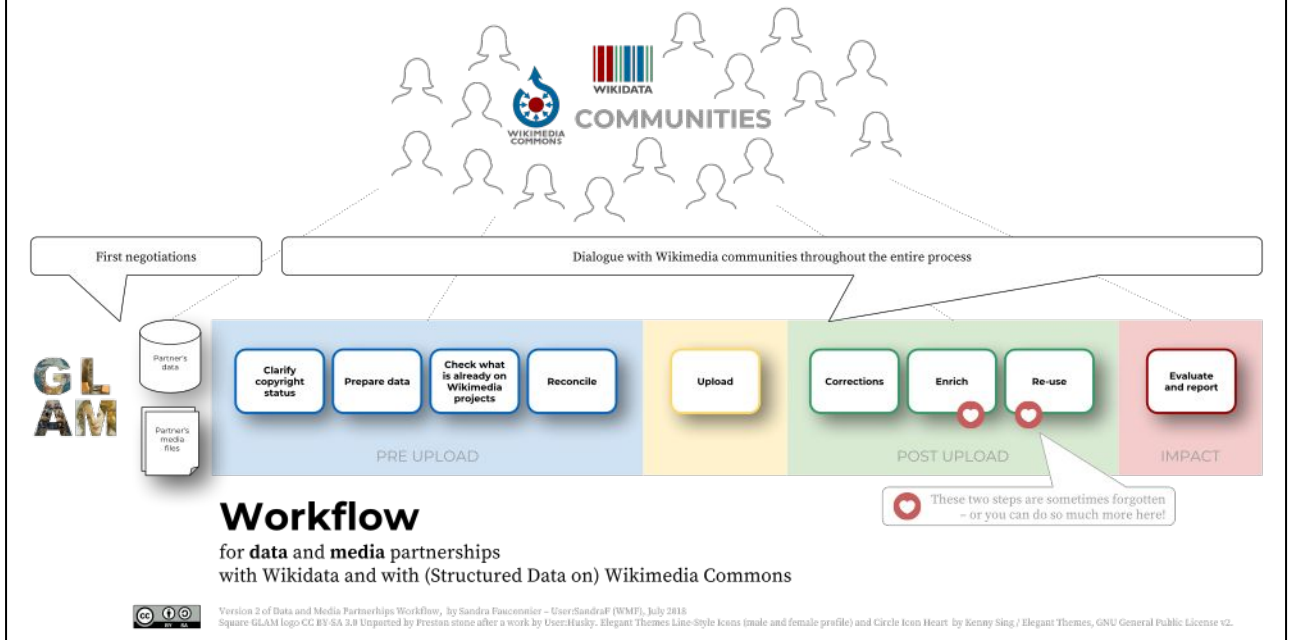
- **GLAM-Wiki volunteers and affiliate staff members** working on partnerships
- **Wikimedia community members** interacting with the content
- **Content partners** themselves
- **Downstream re-users of content**

20

For whom will we build? Who will be the users of partnerships infrastructure?

- The **volunteers and affiliate staff members** who do GLAM-Wiki partnerships
- The **community members and editors of Wikimedia projects** who don't work on the partnerships themselves, but work with the content, use it, enhance it...
- The **content partners** themselves, giving them more agency with their content on our projects where appropriate
- And re-users of the content, providing them with better means to do that re-use

What applications do we already know need work?



Both Wikimedia Sverige and the Wikimedia Foundation have looked at many existing GLAM-Wiki projects (in the broadest sense) and we already have a **general idea** what the current tools do, and how they fit into a general **workflow**.

Here you see a generic workflow for data and media partnerships built for Structured Data on Commons. It distinguishes between phases before upload, the upload of content, enhancement after upload, and impact measurement.

1 - Increase access to material

Thousands of organizations are interested to share material but are not able to do so, not even the big organizations

- Preparation of the content
- Support for new types of content contributions
- Across multiple Wikimedia platforms (Wikidata, Wikimedia Commons, Wikisource, Wikipedia)
- Facilitated re-use across Wikimedia platforms

22

A first problem space we see, is the phase of **increasing and improving access to materials** provided by a content partner. How can we help content partners to make their materials available on our platforms?

Partnerships are already encountering difficulties at the beginning of the process.

- How can their materials be selected and be made suitable for Wikimedia platforms?
- How is the material uploaded on Wikimedia platforms?
- How should we deal with materials that are 'new' to our platforms in terms of content and format?
- Can we envision more easy ways to work across Wikimedia projects, especially now with SDC, where it is necessary to do both work on Wikidata and Commons
- How can we make it easier for anyone to find and re-use the materials in a good way, both on and outside of Wikimedia platforms?

2 - Enrichment of the material

Once material is made available on our platforms our community only has the capacity to enrich a fraction of it.

- Volunteers want to help, but it is hard to find tasks
 - No way to easily discover what material needs attention
- Volunteers want to help, but they have limited time to spare
 - Currently most improvements are entirely manual. Can we help with gamification and some automation?

23

A second large problem space we see, is **enrichment**.

Our global communities hold a lot of expertise and knowledge that can help improve information about materials donated by partners. But it's not happening often enough, also because of technical constraints.

1. In the big sea of millions of files and pieces of data on Wikimedia platforms, you have to be in the know to find those materials that need work. Can we think of ways to help volunteers to better discover tasks to work on, and to find materials that need attention?
2. Volunteers have limited time, and there may be aspects of the enrichment process that can be made much more efficient. We can think of designing games to make contribution quick and fast, and some automation where that's suitable and appropriate.

3 - Showing partners the large impact of working with Wikimedia platforms

- Measuring impact and re-use, by providing statistics
- Visualizations and showcases of (enrichment and re-use of) partner contributions

The third problem space is impact measurement.

This is a crucial element that is needed to actually convince partners to work with Wikimedia, and to keep them around.

Our content partners contribute to our platforms because it helps them to fulfill their mission. They need to be able to see the impact of the work they do with us.

For this purpose, we know we need better statistics for showing impact and re-use of content partners' materials, also through visualizations and showcases of partner contributions.

Research

- Mapping the existing tools landscape (WMSE)
- Mapping the movement's past GLAM-Wiki work, with the technologies used (WMF)
- Researching needs in regions we don't know very well yet (WMF)
- How to build technical capacity among movement actors (WMSE)

25

For the upcoming year, we have planned several strands of research:

- We are going to closely look at the existing tools landscape, to learn which tools and features are needed and used most
- We are going to analyze and map all the past GLAM-Wiki work documented in This Month in GLAM and other places, with the technologies used there
- We know that there are large parts of the world that we don't know about yet. So we are also going to dive more deeply in the needs of various regions in the world that we don't know well yet.
- Wikimedia Sverige is itself also researching how to build out the technical capacity needed for this initiative, both for its own team and among all the movement actors.

Output by June 2020

- **Technical direction** for content partnerships
- **Research report** on international institutional needs and use cases from across the world
- **Growth, governance, evaluation, and sustainability plan**
- **Formalized agreement** between WMF and WMSE for FY20/21 and beyond
- WMSE **team growth** and first pilots

26

What will we end up with in a year?

- We plan to produce a technical direction, with a prioritized list of things that need to be done in the upcoming years
- We create a research report that outlines the results of the research we just told about
- We will also present a growth, governance, evaluation and sustainability plan
- By mid 2020 we want to have a formalized agreement for financial support between WMF and WMSE that covers the first years of work at least
- By that time, the team at Wikimedia Sverige will be extended and will continue to grow. Depending on funding, we may also have done first pilots.

How to get involved

Talk to us in person!

Email us:

John Andersson (john.andersson@wikimedia.se)

Sandra Fauconnier (sfauconnier@wikimedia.org)

Project and talk page on meta.wikimedia.org
upcoming

How can you get involved in this initiative?

We are of course together here at the conference - please find us and talk to us.

Feel free to send us an email with any questions or remarks.

After Wikimania, we will announce this further on mailing lists and also make information available on Meta. The talk page there will also be a first place of interaction.





What are your first thoughts?

- **Turn to the person next to you and discuss for 3 minutes.**
- **Share three things that come to mind (if you'd like).**

So what we would like to do now is to collect your initial thoughts. We would like you to turn to your neighbour next to you and discuss for 3 minutes. Try to identify three things together that you think is important for us to keep in mind in the continuous planning of this initiative.

After the 3 minutes, we would ask you to share your thoughts to the group. If you feel uncomfortable doing that you can email your thoughts directly to us or just write it in the Etherpad.

Milestones

-  **August 2019** – Wikimania
 - Presentation about this initiative
-  **October** – WikidataCon
 - First draft technical direction
-  **November** – Seek funding for first pilot development
-  **January 2020** – Start of WMF annual planning FY 2020-21
 - Technical direction
 - Start of first piloting
- **Q4** – MoU drafting

Very rough timeline

						2020					
Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
	Wikimania		WikidataCon			WMF Annual planning					
Research: existing tools and case studies			Research: the GLAM-Wiki ecosystem and (new) user needs								
Draft technical direction				Technical direction Draft governance, etc plan			Adjust plans based on research				
HR engagement			Start hiring new staff								
				Opportunity Fund application		Start piloting?					

WMSE – WMF GLAM team





WIKIMANIA
STOCKHOLM