Structured Data on Wikimedia Commons

GLAM pilot projects workshop

Wikimedia Conference 2018

Hello!

Pam Drouin
User:PDrouin (WMF) pdrouin@wikimedia.org
Sandra Fauconnier
User:SandraF (WMF) sandraf@wikimedia.org
Alex Stinson
User:Astinson (WMF) astinson@wikimedia.org
Ben Vershbow
User:BVershbow (WMF) bvershbow@wikimedia.org
Liam Wyatt
User:Wittylama liamwyatt@gmail.com

Goals

1. **You** learn about Structured Data on Commons (SDC);
2. The **SDC team** learns what you need in order to be able to work with SDC;
3. We **all** leave this workshop with a list of ideas and potential pilot projects!

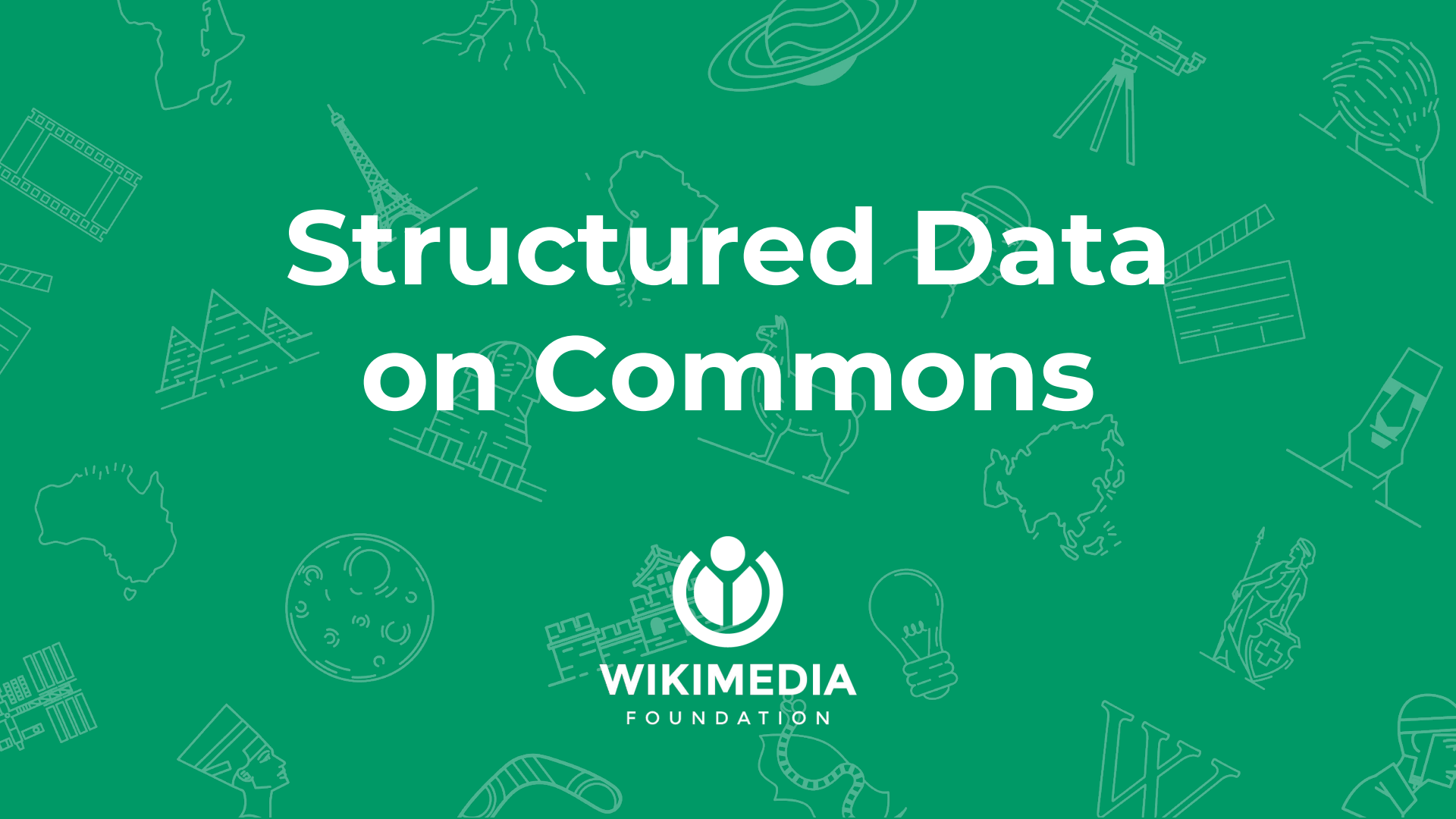
Schedule

Structured Commons and GLAM: intro	30 minutes
Workshop exercise 1: List your upcoming GLAM projects	5 minutes
Workshop exercise 2: Project highlights and issues	5 minutes
Workshop exercise 3: Changes and needs	60 minutes
Group summaries	10 minutes
Wrapping up, conclusions and follow-up	10 minutes

Parking lots

WMF Support for SDC Pilot Projects

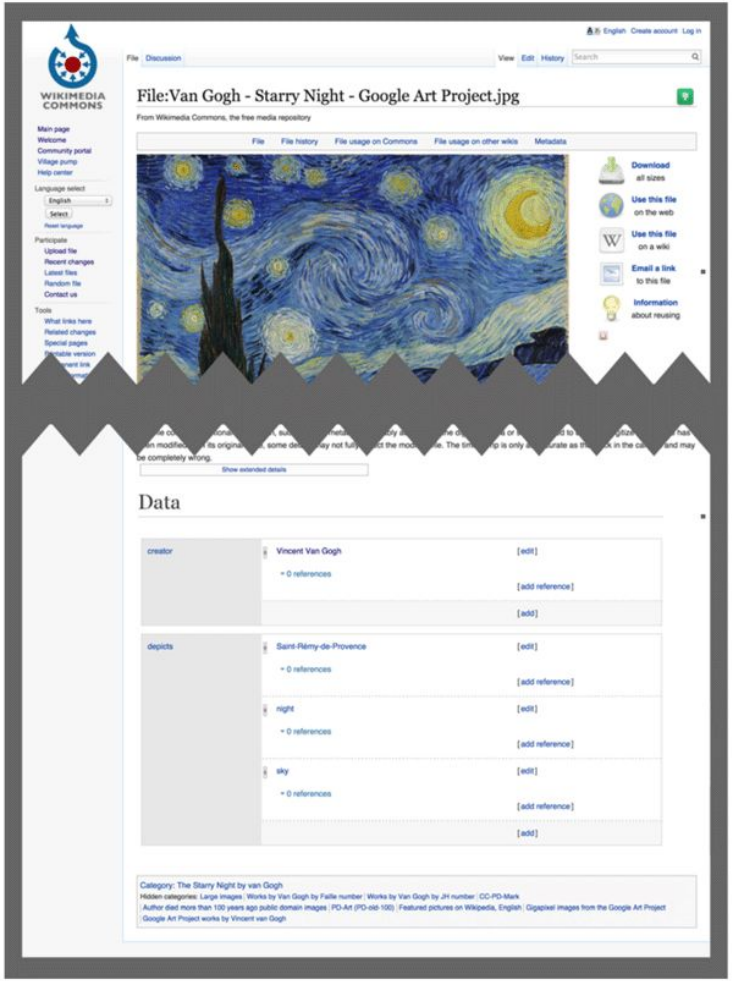
Structured Data on Commons



Structured Data on Commons

converts metadata on Commons
to a **structured** & **machine-readable** format
making Commons files easier to
view, **search**, **edit**, **organize** and **re-use**,
in many **languages**

WIKIMEDIA
FOUNDATION



File page

Data section

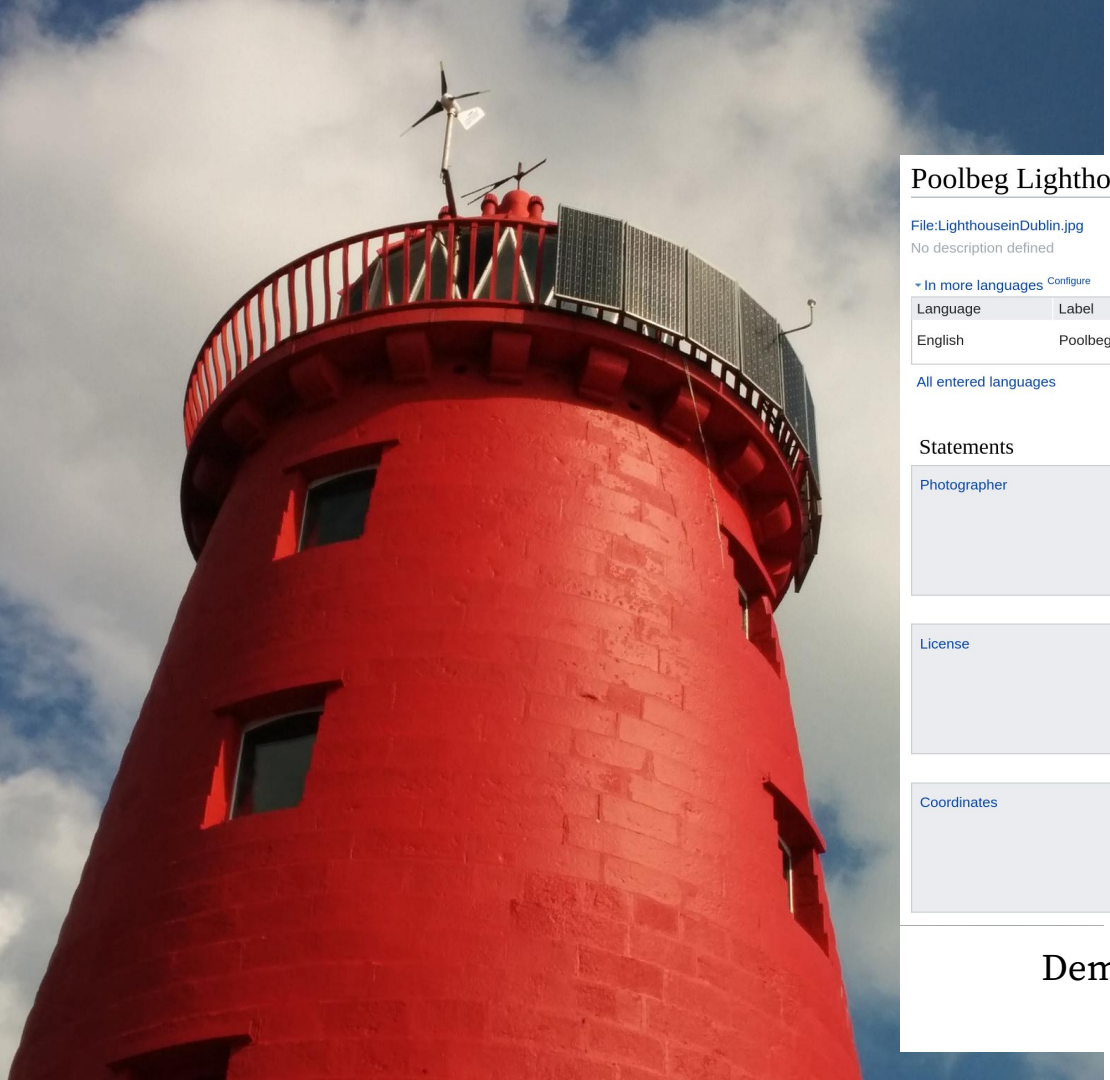
Starry Night by Vincent Van Gogh, from Google Art Project. Public domain. Via Wikimedia Commons.

5 outcomes

(and one corollary outcome)

1. Make editing metadata easier
   (and enable multilingual statements)
2. Make search more effective
3. **Satisfy GLAM use case**
4. Allow license compliant reuse
5. Support tool development

WIKIMEDIA
FOUNDATION



Federation

Poolbeg Lighthouse (M13)

File:LighthouseinDublin.jpg

No description defined edit

In more languages Configure

Language	Label	Description
English	Poolbeg Lighthouse	No description defined

All entered languages

Statements

Photographer	Lydia Pintscher 0 references add reference add	edit
License	CC-By 0 references add reference add	edit
Coordinates	53°20'31.56"N, 6°9'6.12"W 0 references add reference add	edit

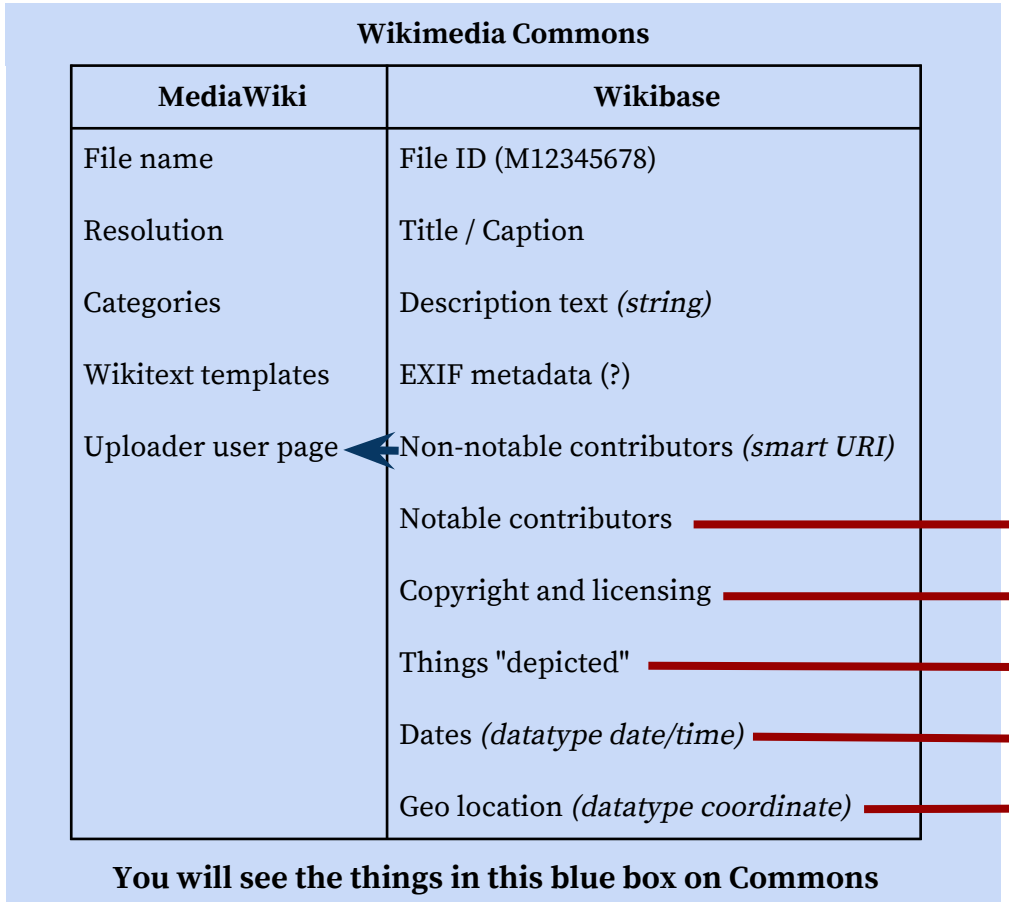
Demo: http://structured-commons.wmflabs.org/
and http://federated-wikidata.wmflabs.org/

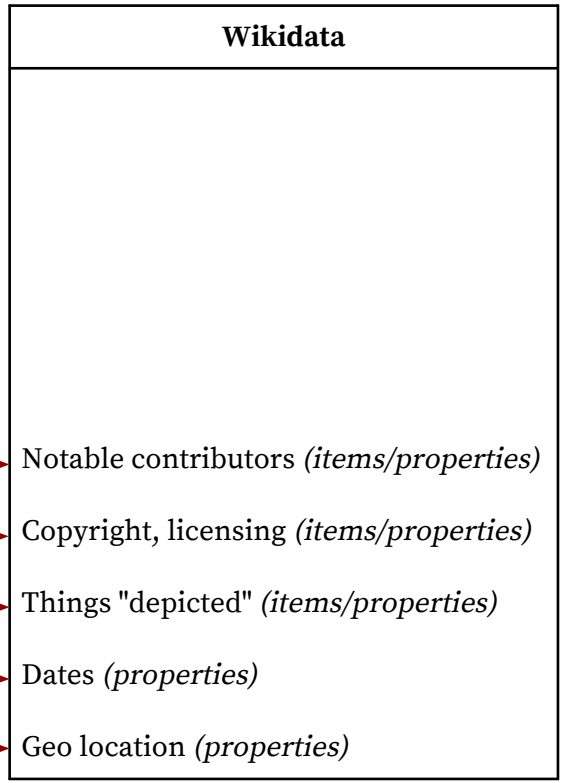
Structured Commons - What lives where?

(sample, doesn't cover everything)

Smart URI (new type of link to Wikimedia user page, Flickr user page, ...)

Uses Wikidata items/properties via federation

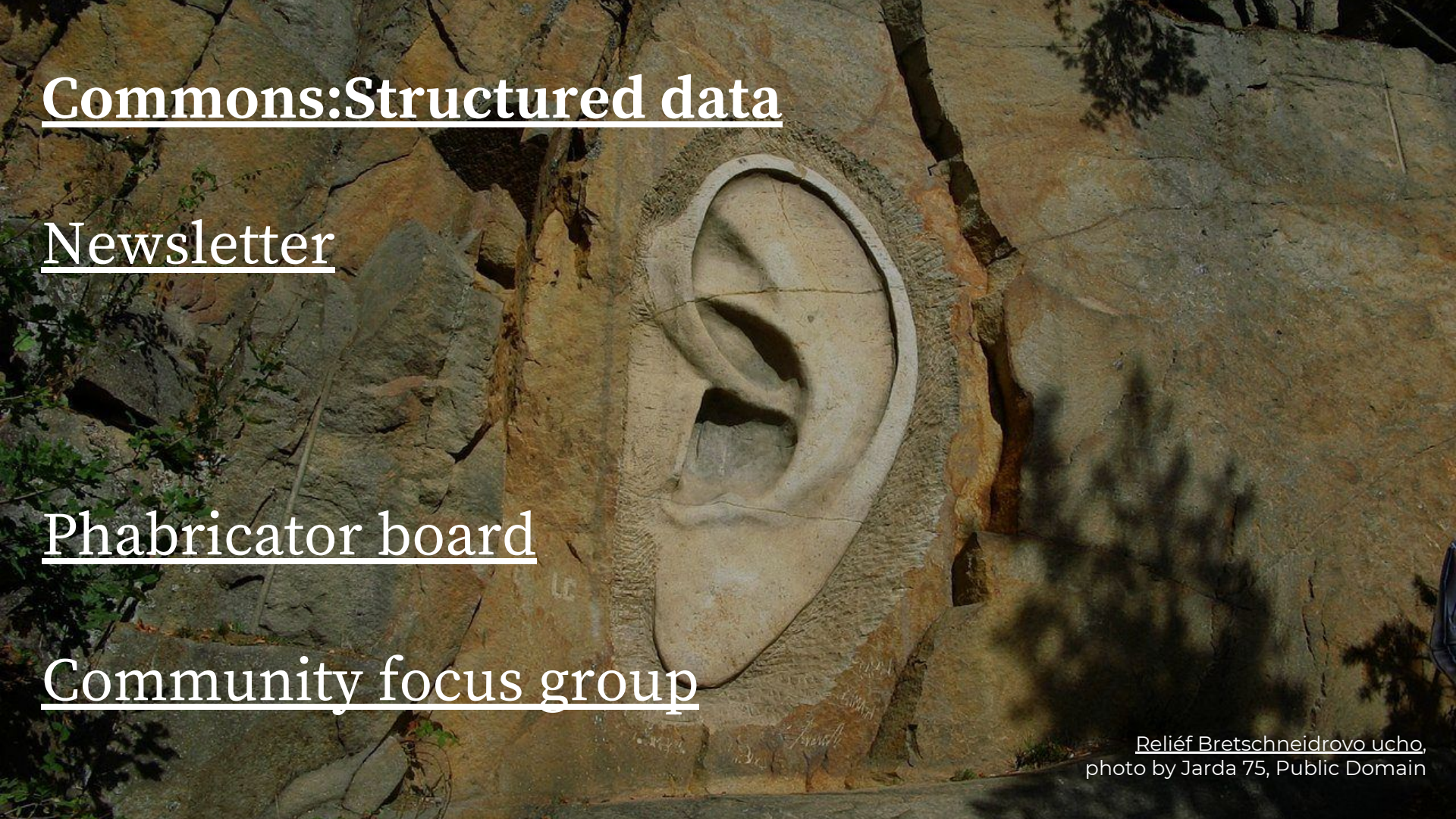




Upcoming

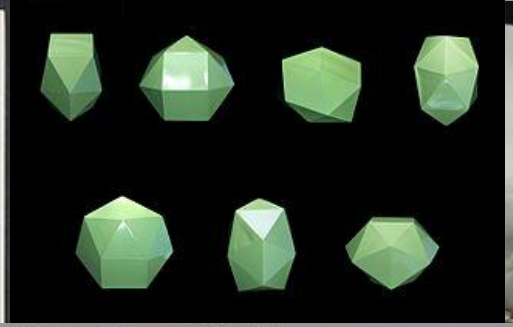
- **A first feature**
  - Multilingual captions (Summer 2018)
- **Full structured data rollout**
  - 'Depicts' in Autumn 2018
- **Search**
- **Upload**
- **New data modelling by Commons and Wikidata communities**
  - What is depicted in media
  - Contributors
  - Copyright and licensing

WIKIMEDIA FOUNDATION



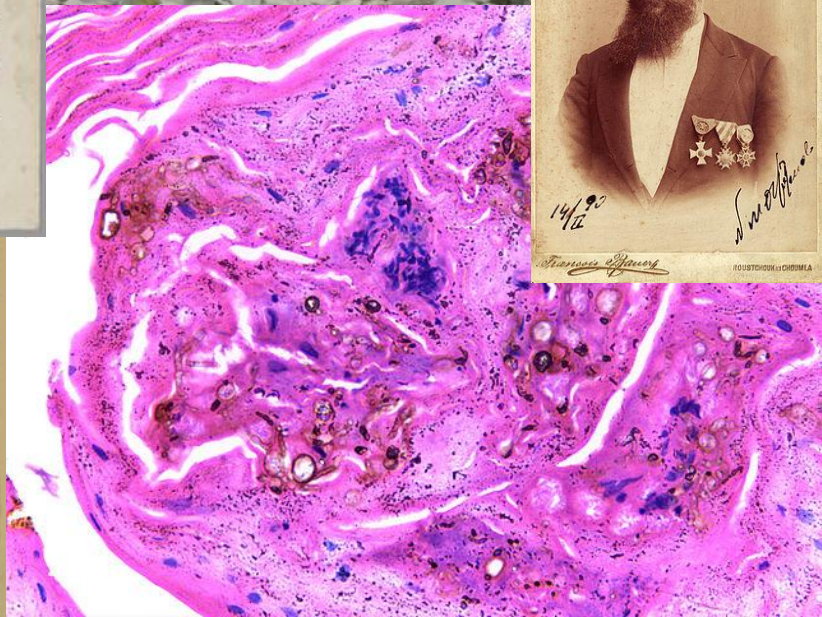
Reliéf Bretschneidrovo ucho,
photo by Jarda 75, Public Domain

SDC and GLAM



URSULA by het UURWERK.
Daar slaat het zes:
Nu aan mijn lea;
Als ik die ken,
Zegt moeder, dat ik vlijtig ben.
VINCENT met het VLANDEL.
Vincent met het vaantje
Vertoont u deze print;
Hij is al een baantje,
Die 't vaderland mint.

Coração de Suíno
Técnica: Diafanização
1 2 3 4 5 10 15cm



Research

- Focus group @ 2017 European GLAMwiki coordinators meeting
- Interviews w/ GLAMers from 5 continents
- Survey through GLAM monthly newsletter

[meta:Research:GLAM and Commons]

WIKIMEDIA FOUNDATION

Personas

Wikimedian in Residence

Sammy

Tech skills: Intermediate
Wikimedian: No
People involved: Sammy, GLAM Lawyer (for licensing)
Type of media: Historical books and pictures
Location: Canada

Problems

Limited knowledge of different templates and template syntax

Needs to manually correct errors after batch uploads

How to show impact of uploads?

"I don't understand Commons well and have a hard time optimizing my uploads."

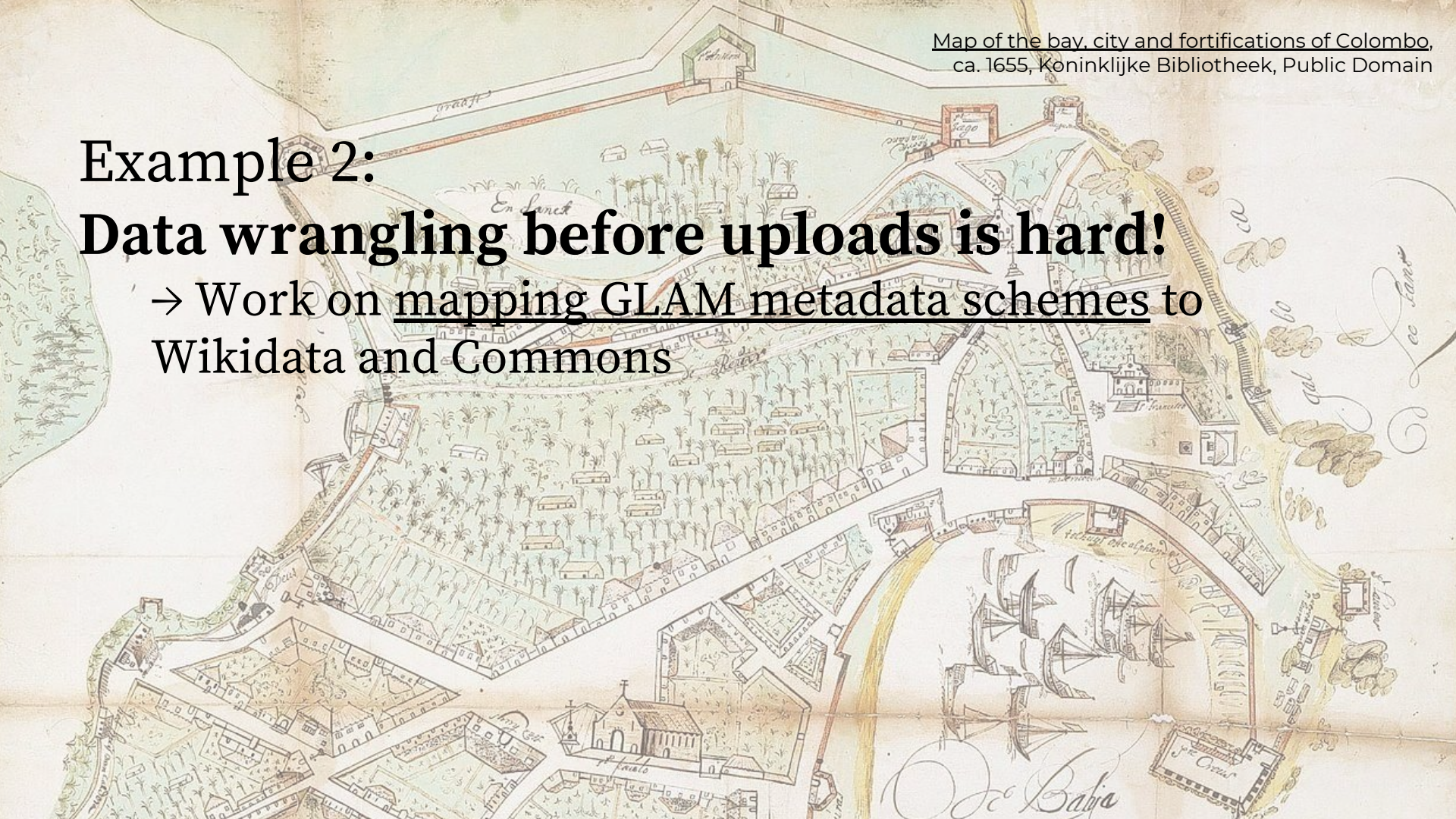
Example 1:

GLAMs and Wikimedians need good documentation

→ Embark on GLAM pilot projects, starting April 2018

- to be well documented

Second 'Fram' expedition, 1899, National Library of Norway, from Flickr the Commons (no known copyright restrictions)



Map of the bay, city and fortifications of Colombo, ca. 1655, Koninklijke Bibliotheek, Public Domain

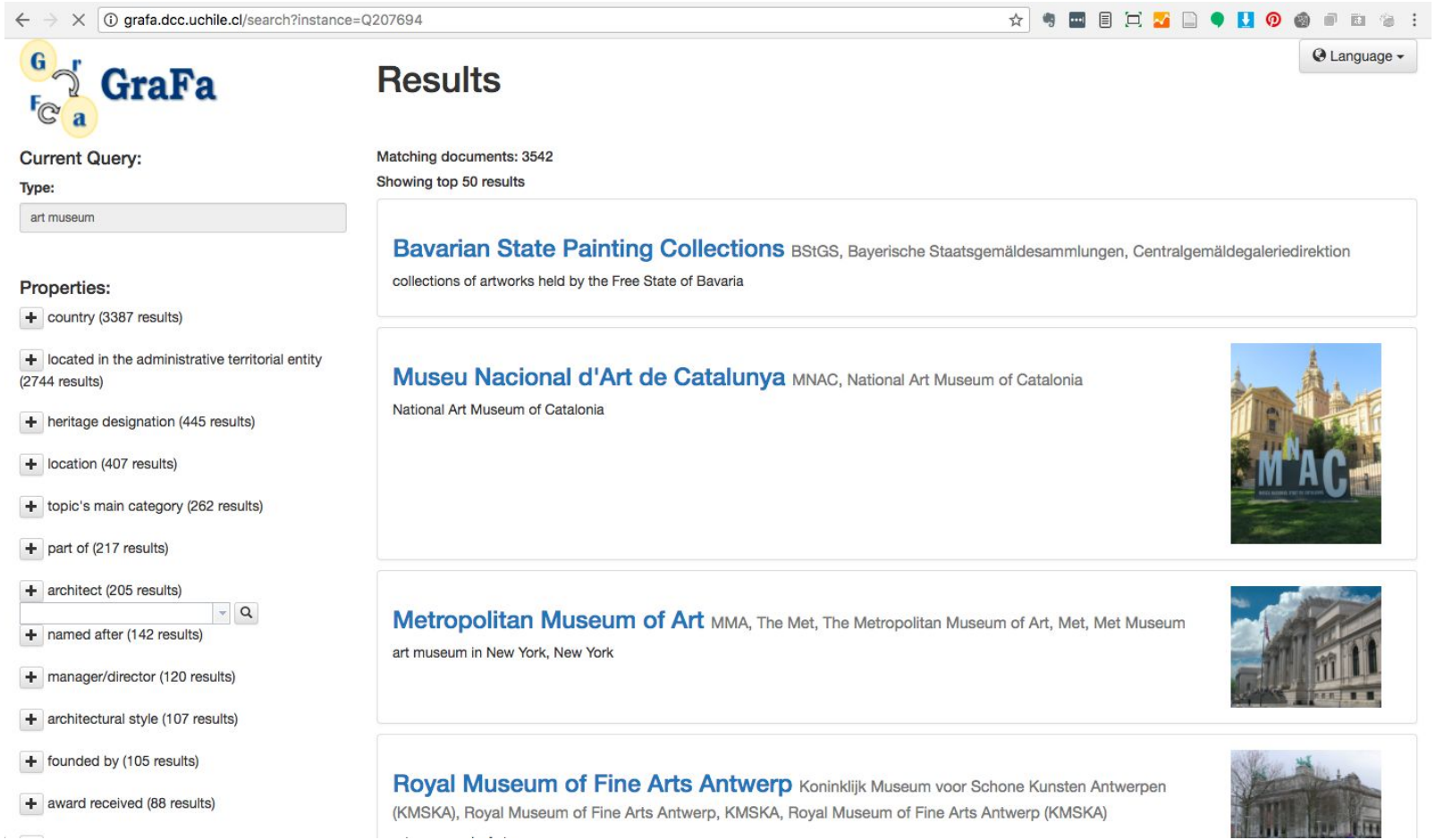
Example 2:

Data wrangling before uploads is hard!

→ Work on mapping GLAM metadata schemes to Wikidata and Commons

What does SDC make possible for new GLAM-Wiki projects?

WIKIMEDIA
FOUNDATION



New ways to search Commons and to discover interesting files

On Sunday October 4th, World Animal Day, the Rijksmuseum will host a birdwatching day in collaboration with Naturalis Biodiversity Center, Wikimedia Netherlands and the COMMIT SealincMedia project.

Rijksmuseum – Naturalis Biodiversity Center – Wikimedia Nederland & COMMIT/ SEALINCMedia present a unique birdwatching event

Birds are everywhere. In your own garden, in nature, and also in art. Among the Rijksmuseum's 1,2 million collection objects are many prints, paintings and artefacts that have bird species depicted on them. Among the 37 million objects in the Naturalis collection are many birds from all over the world that have been collected in the last 200 years, as well as historical drawings of plants and animals in which many birds are depicted.

Wikimedia Commons

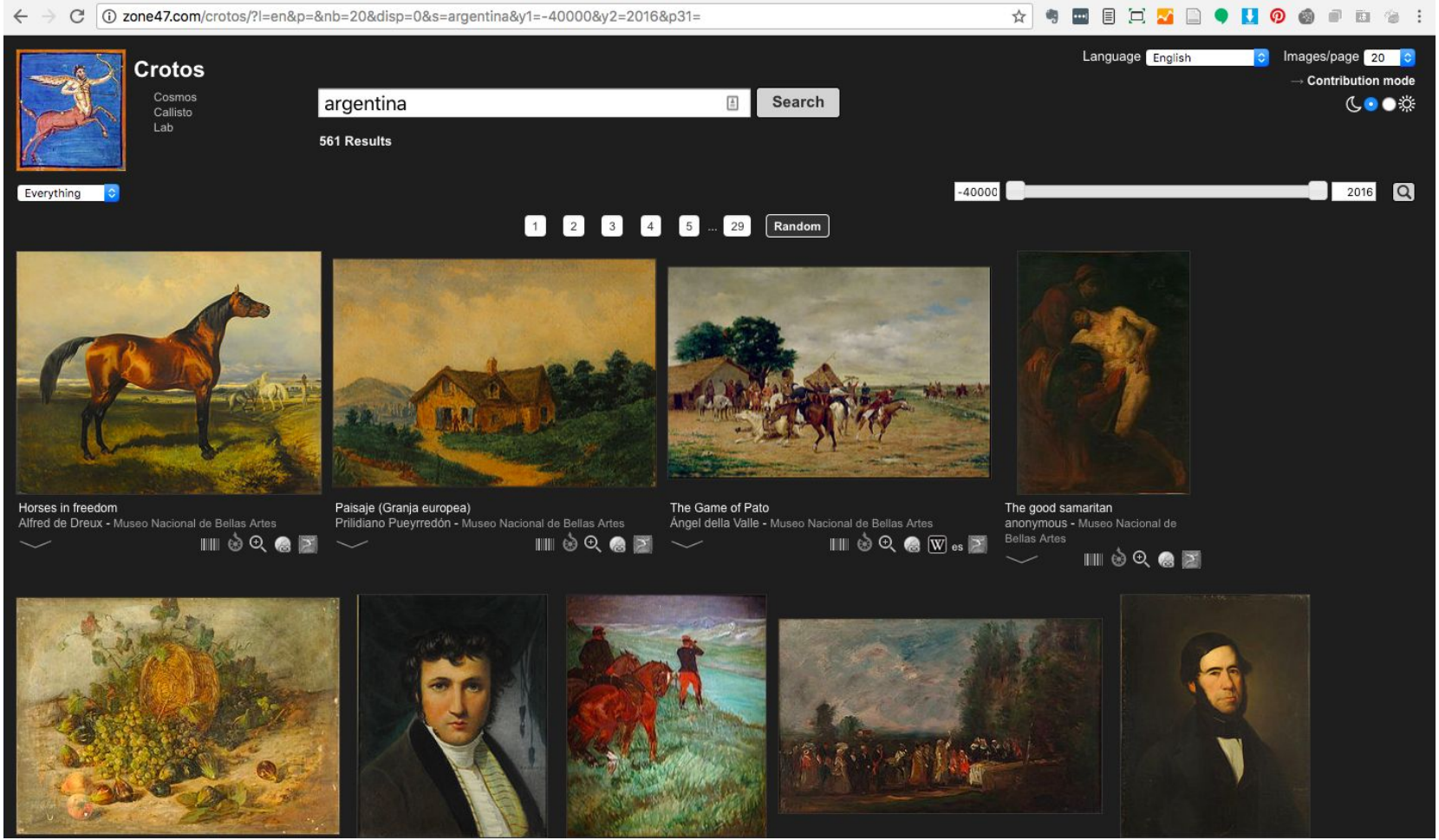
Some of the depicted birds are easily identified. Others require a trained eye to determine which species the artist has pictured. The Rijksmuseum and Naturalis are looking for experienced bird watchers and other avian enthusiasts to join an expedition through their digital collections and help the museums identify bird species in works of art.

The Rijksmuseum and Naturalis are currently in the process of donating large parts of their digitized collections of bird images to Wikimedia Commons, Wikipedia's open multimedia library. Participants of the birdwatching day are challenged to collaboratively identify as

Much better crowdsourcing

… with data that can be given back to an institution

WIKIMEDIA FOUNDATION



More advanced tools (based on a much improved Commons API)

Secure | https://commons.wikimedia.org/wiki/Self-portraits_of_women

Self-portrait *(1650)*
Louise Hollandine of the Palatinate
private collection

Self-portrait *(1650)*
Margaretha van Godewijk
Dordrecht

Self-portrait *(1653)*
Judith Leyster
private collection

Self-portrait *(1658)*
Elisabetta Sirani
Pushkin Museum of Fine Arts

Self-portrait *(1660)*
Elisabetta Sirani

Self-Portrait at Her Easel
(1670s)
Maria Schalcken
Rose-Marie and Eijk van Otterloo Collection
Museum of Fine Arts Boston

Self-portrait *(1670s)*
anonymous
Louise Hollandine of the Palatinate
Stedelijk Museum Zwolle

Self-portrait in the Louvre
(1672)
Élisabeth Sophie Chéron
Department of Paintings of the Louvre

Self-portrait *(1675)*
Mary Beale

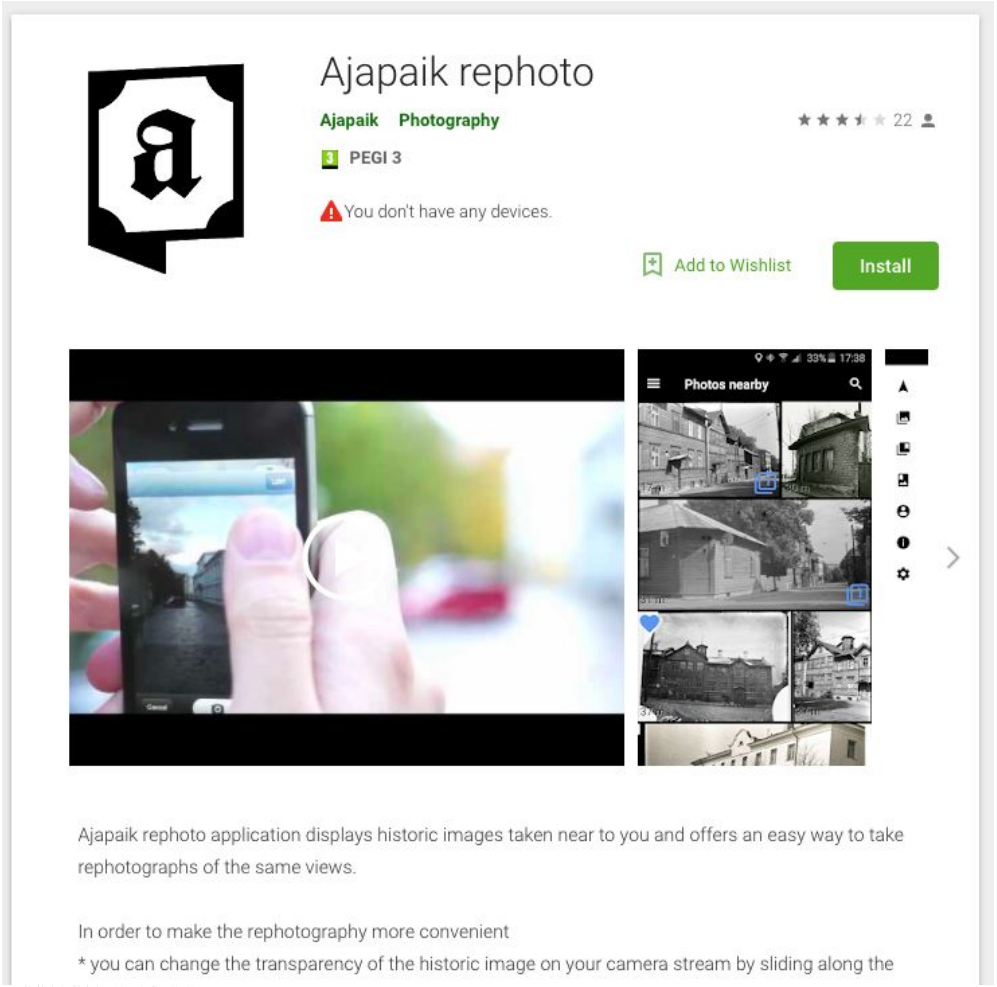
Self-portrait *(1678)*
Mary Beale
Moyses Hall

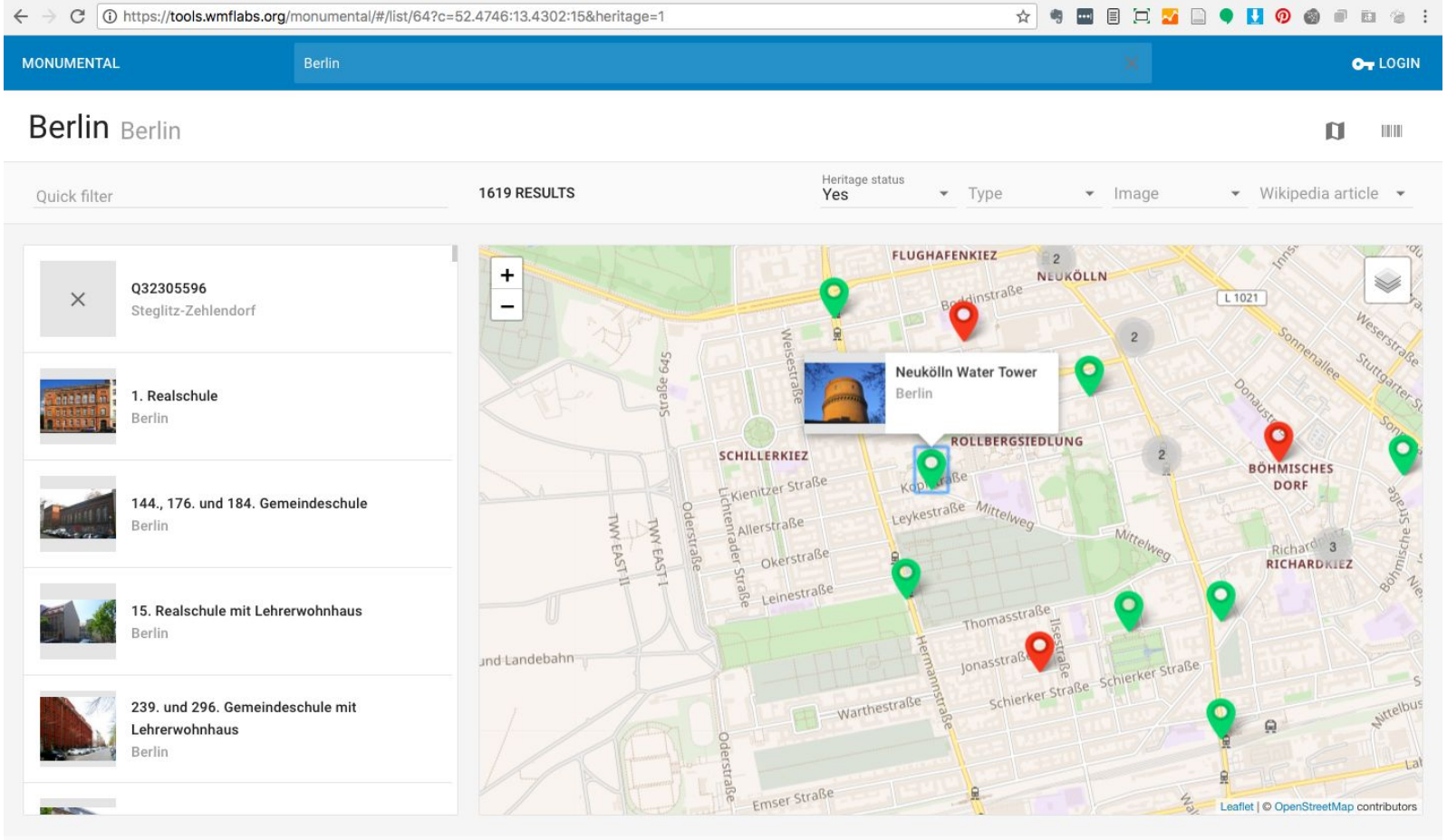
Self-portrait *(1691)*
Anna Waser
Kunsthaus Zürich

Self-portrait *(1701s)*
Rosalba Carriera

Bringing together information that has never been combined before
(across heritage collections!)

New types of campaigns





More value and potential of Commons' media files **beyond Wikipedia**

Who already uses Wikidata in GLAM projects?

How and why?

WIKIMEDIA
FOUNDATION

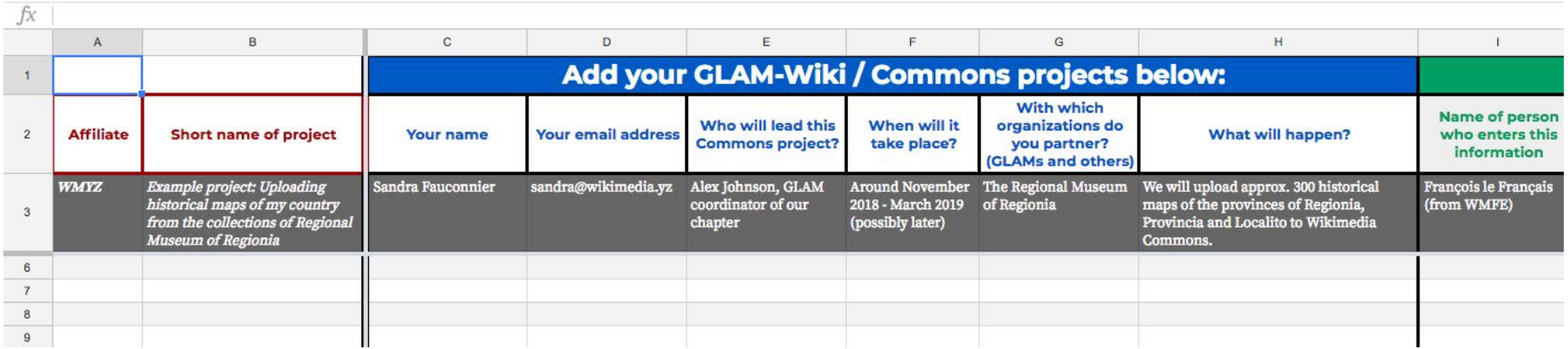
Workshop time!

List your affiliate's upcoming GLAM + Commons projects

1. Do this **individually**
2. In this spreadsheet:

   https://is.gd/glamcommons
3. 10 minutes maximum

What to do:



	A	B	C	D	E	F	G	H	I
1			Add your GLAM-Wiki / Commons projects below:						
2	Affiliate	Short name of project	Your name	Your email address	Who will lead this Commons project?	When will it take place?	With which organizations do you partner? (GLAMs and others)	What will happen?	Name of person who enters this information
3	*WMYZ*	*Example project: Uploading historical maps of my country from the collections of Regional Museum of Regionia*	Sandra Fauconnier	sandra@wikimedia.yz	Alex Johnson, GLAM coordinator of our chapter	Around November 2018 - March 2019 (possibly later)	The Regional Museum of Regionia	We will upload approx. 300 historical maps of the provinces of Regionia, Provincia and Localito to Wikimedia Commons.	François le Français (from WMFE)
6									
7									
8									
9									

only add your info in the red and blue columns (for now)

Project highlights and issues

1. Do this **in pairs**
2. In the same spreadsheet: **https://is.gd/glamcommons**
3. 10 minutes maximum

Gustav Köhler: Berlin, Frauen beim Selbststudium, Weiterbildung. 1952, Bundesarchiv, CC BY-SA 3.0 Germany. https://commons.wikimedia.org/wiki/File:Bundesarchiv_Bild_183-13800-0006,_Berlin,_Frauen_beim_Selbststudium,_Weiterbildung.jpg

What to do:

I	J	K	L
(Workshop only - Part 2)			
Name of person who enters this information	What is interesting/special about this project?	What is challenging/difficult about this project?	Do you feel ready to do this project with structured data? Why (not)?
François le Français (from WMFE)	We don't have many historical maps of these regions on Commons and Wikipedia yet. It is our first collaboration with this museum.	The museum doesn't have very extensive descriptions of the maps. We don't have many volunteers who do GLAM projects locally and who can help with the upload.	No. Our chapter and volunteers don't have a lot of experience with Wikidata yet and even Commons uploads are still quite new for us.

write info about your neighbor's project in the green area

3 x 15 minutes...

1. **Opportunities** of SDC
2. New working **processes** with SDC: what do you **need**?
3. **Ideas** for (new projects) with SDC

Service Civil International / Anthony Dubois, CC BY-SA 3.0
https://commons.wikimedia.org/wiki/File:SCI-formation.jpeg

Changes, needs, ideas!

1. **4 or 5 larger groups**
    a. Similar interests
    b. Mix of beginners and more experienced people
    c. Assign a **note-taker** and a **reporter**

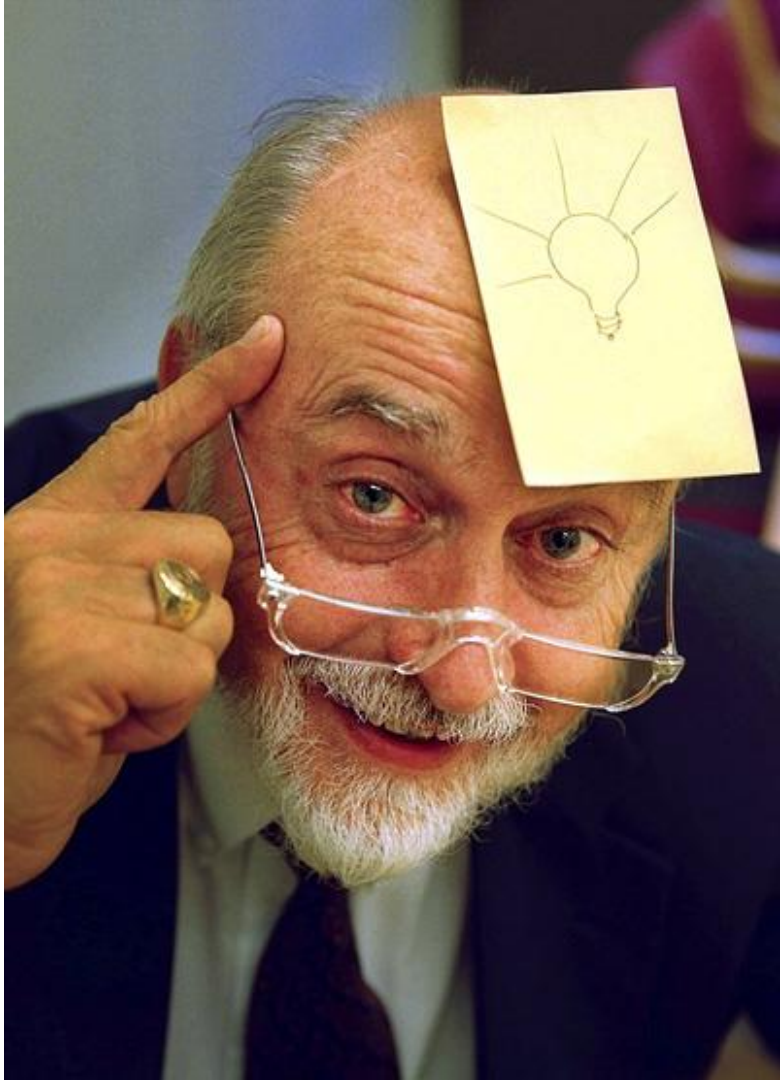
2. **15 minutes per question + break** - max. 60 min total
    a. Alex, Ben, Liam, Pam, Sandra keep time and structure

Notes here:

https://etherpad.wikimedia.org/p/sdc-pilots

Process

1. **Discuss** the questions **together** (take good notes!)
2. Remaining ideas, issues? **Post-it notes**



Arthur Fry, inventor of the Post-it Note, with one on his forehead bearing a with a picture of a lightbulb. Photo by Tinkeringbell, Public Domain https://commons.wikimedia.org/wiki/File:Fry-lightbulb-on-forehead1.jpg

Town crier on the island of Terschelling, the Netherlands. Nationaal Archief / Spaarnestad Photo; photographer unknown. No known copyright restrictions, from Flickr the Commons. https://commons.wikimedia.org/wiki/File:Dorpsomroeper_-_Town-crier_(3280639091).jpg

Reporters!

1 point per 'block' that stood out
(3 in total)

Opportunities (15 mins)

- Does SDC help reach new **audiences**? How?
- Does SDC help you to work with new **partners**, on new collaborations? How?
- Which **other opportunities** does SDC bring?

Think of examples from your own projects.

1. **Discuss** the questions **together** (take good notes!)
2. Remaining ideas, issues? **Post-it notes**

Process and needs (15 mins)

- **People** involved in your programs (Wikimedians, affiliate members/staff, GLAM staff…)
  - Are they interested in structured data?
  - What do they need to get 'ready' for such projects?
- Your **working process** for GLAM-Wiki projects
  - How will this change?
  - Do you need help? In which areas?

1. **Discuss** the questions **together** (take good notes!)
2. Remaining ideas, issues? **Post-it notes**

Grab a snack or a drink!

WMCON 2017 Sweets Table, by Camelia Boban, CC BY-SA 4.0
https://commons.wikimedia.org/wiki/File:WMCON_Berlin_2017_10.jpg

Ideas brainstorm (15 mins)

- Which of your own **planned** projects seem most **ready** for SDC?
- Which other **(new) projects** can you imagine with SDC?
    - By your own affiliate
    - Together with other affiliates

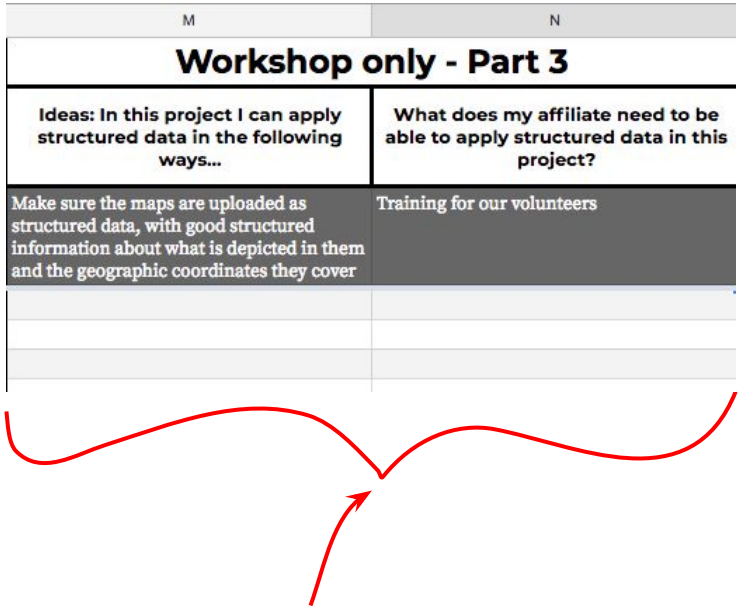
1. **Discuss** the questions **together** (take good notes!)
2. Remaining ideas, issues? **Post-it notes**

Town crier on the island of Terschelling, the Netherlands. Nationaal Archief / Spaarnestad Photo; photographer unknown. No known copyright restrictions, from Flickr the Commons. https://commons.wikimedia.org/wiki/File:Dorpsomroeper_-_Town-crier_(3280639091).jpg

Back to your own planned projects...
New insights on SDC applications

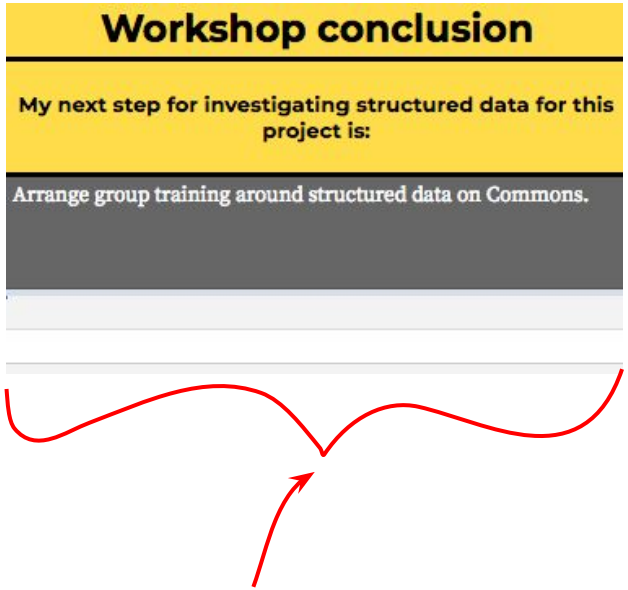
1. Do this **individually**
2. In the same spreadsheet:
   https://is.gd/glamcommons
3. 10 minutes maximum

What to do:



M	N
Workshop only - Part 3	
Ideas: In this project I can apply structured data in the following ways...	**What does my affiliate need to be able to apply structured data in this project?**
Make sure the maps are uploaded as structured data, with good structured information about what is depicted in them and the geographic coordinates they cover	Training for our volunteers

write new thoughts about your project in the white area



Workshop conclusion
My next step for investigating structured data for this project is:
Arrange group training around structured data on Commons.

write your own conclusion for follow-up in the yellow column !

Conclusion

Who thinks they already have a quite clear idea for an **implementable pilot project**?

Let's follow up!

What are your main takeaways?

Thank you!