

Wikidocumentaries

Short paper at DHN2018

Susanna Ânäs

background

Wikidocumentaries is a concept for a collaborative online space for gathering, researching and remediating cultural heritage items from memory institutions, open platforms and the users. The setup brings together communities of interest and of expertise to work together on shared topics with open online tools.

The platform will make available collections of open content – images, written documents, maps, data, sounds, print – through querying open repositories, such as the national content aggregator *finna.fi*, *Wikimedia Commons* or *Flickr Commons*. Wikidocumentaries uses open structured data from *Wikidata* as a means of organizing and filtering the content. Open tools on the platform are used to identify, arrange and enrich the content, but also to remix and reuse it creatively. The resulting data is recirculated to open platforms and made universally available.

For the memory organization, Wikidocumentaries offers a platform for crowdsourcing, for the amateur and expert researchers it provides a community of peers and audiences as well as tools for discovery and interpretation, and from the point of view of the open environments, it acts as an intermediate phase of curation.

Current environments fall short in serving this purpose. Content aggregators focus on gathering, harmonizing and serving the content metadata. Commercial services fail to take into account the open and connected environment in the search for profit. Research environments do not prioritize public access and broad participation. Many participatory projects live short lives from enthusiastic engagement to oblivion due to lack of planning for the sustainability of the results. Wikidocumentaries will try to battle these challenges.

The project has been prepared with a network of participants from Finnish memory institutions, universities and NGOs, who focus on different aspects of the project. Some provide content, while others advance the platform through research or as partners for community projects. This year Wikidocumentaries is supported by the Kone Foundation for co-designing and developing the platform and launching the first community projects.

This short paper will create an initial inventory of research topics that this environment surfaces.

A public domain kaleidoscope

Wikidocumentaries launches with Finnish places, people and events that are stored in Wikidata. Each topic is displayed as a page that gathers all open content about the topic from a variety of services. New topics can be created and they will be stored in Wikidata whenever possible.

Technologically Wikidocumentaries uses linked open data and especially Wikidata in structuring cultural heritage content. It creates a proof of concept of creating a federated Wikibase instance, i.e. a Wikidata-like repository that communicates with Wikidata proper. Best practises will be researched in collaboration with other Wikibase projects, such as the Rhizome.org or the Finnish Names Archive project (Kotus), which has been created by members of the project team.

Public & personal

For Wikidocumentaries everyone is notable and differing views are allowed. Each participant can focus on their own viewpoint and research, adding their own pictures and testimonies to topics of their interest.

As the 20th century in Europe is already a “black hole” in the online world for copyright reasons (e.g. Boyle, 2009) a new restricting challenge in bringing cultural heritage online arises with the forthcoming General Data Protection Regulation (European Union). When combining data and images about people in Wikidocumentaries, it must be legally and ethically appropriate. Publicly funded memory organizations’ mission to store and display information about historical subjects will most probably be secured as public or legitimate interest (Voutilainen, 2017), but the situation may be trickier for an amateur online project aggregating cultural heritage content, as it may not fit any specified exception. The museum sector is calling out for the “right to be remembered” as opposed to the right to be forgotten as a means for minorities to be represented in collections and, consequently, in recorded history (Levä, 2017).

In the project we will participate in discussions about technologies of consent, conducted in the MyData community (Conway, 2017). This project may take initiative in innovating novel ways of obtaining permission from living people to display images depicting them.

Other legal challenges can also arise from the European copyright reform. Especially, the proposal to impose neighbouring rights to publishers may complicate aggregating and displaying content from different sources (Communia, 2018).

Content permaculture

New images are saved into Wikimedia Commons and data is stored in Wikidata. Images can also be saved in other partnering projects when saving to Wikimedia projects can not be considered.

New opportunities arise with the advent of structured data for Wikimedia Commons. The world’s largest repository of open content will be remodelled to utilize another federated Wikibase instance, turning all metadata into a structured and machine-readable format. This

will enable systems to take advantage of and contribute to the repository in unprecedented efficiency (Morgan, Fauconnier, 2017).

More generally, Wikidocumentaries will continually revise how it sits as a part of the open de-centralized web. It keeps the tools, the content and the storage independent of one another to allow content to flow freely and prevent vendor lock-in.

Meaningful participation

Wikidocumentaries hosts open source tools to work with the materials. Participants identify and classify, locate, date and transcribe, put the pieces in context and order. Work is done reciprocally, by giving and getting. GLAMs get wide participation in their content, and the content gets curated before being saved in open platforms.

There is a growing body of research on crowdsourcing initiatives within Arts and Humanities projects (e.g. Ridge 2015, Dunn 2012). Research has identified modes of engagement from task-based workflows of crowdsourcing projects (Sprinks et al. 2017) to more open Commons-based peer production of Wikimedia projects (Benkler 2006).

This project aims to position itself as an inclusive environment that is driven by the desire of the participants to collaboratively produce information and make it available in the Cultural Commons. For the participants, it presents a scaffolding approach to participation (Ridge 2015, 160), facilitating moving between microcontributions and a completely open plan.

Wikidocumentaries explores ways to reach out to underrepresented topics and communities by investigating models of live participation, or “public engagement”, as Susan Schreibman prefers to call participation in the Letters of 1916 project (Schreibman et al. 2017).

Discovery, creativity, play

Wikidocumentaries can tell stories dynamically, for example in automated timelines or Twitter feeds, or you can use the material to remix your own stories. As tools are added, you can create story maps, narrated slideshows or perhaps collectively performed songs. Anyone can add new tools.

Ideally, the platform would grow to become a trusted space to bring together any Open Source tools for enriching distributed cultural heritage assets. Developers can create apps that reuse data produced in other contexts and tools. With this goal in mind, we are inviting the community of stakeholders: content providers, developers, domain experts and more, to play and innovate together in workshops, field trips and hackathons.

Cultural Commons

The work is done together. Amateur researchers, local communities, birds-of-a-feather, relatives, friends and families come together with expert members: museums, libraries, archives and scholars.

Both scholarly and amateur researchers share similar interests in knowledge discovery and can benefit from work and expertise of one another. Strikingly, the proportion of amateur

researcher visitors in the National Library newspapers service is over 80% (Hölttä 2016), and the results from other platforms also reveal the general interest in historical material. Thus, as a key strategy for further development, we aim to bring amateur and scholarly researchers together to discover historical content and to share and contribute to the Cultural Commons together (e.g. Madison 2009).

References

“Wikidata:Main Page – Wikidata.” https://www.wikidata.org/wiki/Wikidata:Main_Page.

“Wikidocumentaries.” <https://wikidocumentaries.github.io/>.

Benkler, Yochai, and Helen Nissenbaum. “Commons-Based Peer Production and Virtue.” *Journal of Political Philosophy* 14, no. 4 (December 2006): 394–419.
<https://doi.org/10.1111/j.1467-9760.2006.00235.x>.

Boyle, James. “Google Books and the Escape from the Black Hole | The Public Domain |,” September 6, 2009.
<http://www.thepublicdomain.org/2009/09/06/google-books-and-the-escape-from-the-black-hole/>.

Communia Association. “Seven Ways to Save the EU Copyright Reform Effort in 2018 - International Communia Association.” Accessed March 7, 2018.
<https://www.communia-association.org/2018/01/08/seven-ways-save-eu-copyright-reform-effort-2018/>.

Conway, Shaun. “How to Operationalise Consent – MyData – Medium.” Accessed March 6, 2018. <https://medium.com/mydata/how-to-operationalise-consent-7d6d6357d52b>.

Dunn, Stuart, and Mark Hedges. “Crowd-Sourcing Scoping Study : Engaging the Crowd with Humanities Research.” Centre for e-Research, Department of Digital Humanities, King’s College London, 2012.
<http://crowds.cerch.kcl.ac.uk/wp-content/uploads/2012/12/Crowdsourcing-connected-communities.pdf>.

European Union. “Home Page of EU GDPR.” <https://www.eugdpr.org/>.

Hölttä, Tiina. “Digitoitujen kulttuuriperintöaineistojen tutkimuskäyttö ja tutkijat,” 2016.
<http://tampub.uta.fi/handle/10024/98714>.

Kotus. “Nimiarkisto.” <https://nimiarkisto.fi/>.

Levä, Kimmo. “Suomen Museoliiton P.S.-Blogi: Oikeus Tulla Muistetuksi,” March 2016.
<http://museoliitto.blogspot.fi/2016/03/oikeus-tulla-muistetuksi.html>.

Madison, Michael J., Brett M. Frischmann, and Katherine J. Strandburg. “Constructing Commons in the Cultural Environment.” *Cornell L. Rev.* 95 (2009): 657.

Morgan, Jonathan, and Sandra Fauconnier. "What Galleries, Libraries, Archives, and Museums Can Teach Us about Multimedia Metadata on Wikimedia Commons – Wikimedia Blog," January 29, 2018.

<https://blog.wikimedia.org/2018/01/29/glam-multimedia-metadata-commons/>.

Morgan, Jonathan. "Research:Supporting Commons Contribution by GLAM Institutions – Meta," 2017.

https://meta.wikimedia.org/wiki/Research:Supporting_Commons_contribution_by_GLAM_institutions.

Open Knowledge Foundation. "OpenGLAM Principles | OpenGLAM." Accessed February 7, 2018. <https://openglam.org/principles/>.

Ridge, Mia, ed. *Crowdsourcing Our Cultural Heritage*. Digital Research in the Arts and Humanities. Farnham, Surrey, England: Ashgate, 2014.

———. "Making Digital History. The Impact of Digitality on Public Participation and Scholarly Practices in Historical Research." The Open University, 2015.

Schreibman, Susan, Vinayak Das Gupta, and Neale Rooney. "Notes from the Transcription Desk: Visualising Public Engagement." *English Studies* 98, no. 5 (July 4, 2017): 506–25. <https://doi.org/10.1080/0013838X.2017.1333754>.

Sprinks, James, Jessica Wardlaw, Robert Houghton, Steven Bamford, and Jeremy Morley. "Task Workflow Design and Its Impact on Performance and Volunteers' Subjective Preference in Virtual Citizen Science." *International Journal of Human-Computer Studies* 104 (August 1, 2017): 50–63. <https://doi.org/10.1016/j.ijhcs.2017.03.003>.

Voutilainen, Tomi. "EU:n uusi tietosuojasetus – Mikä oikeasti muuttuu? : Säilyttää vai arkistoida – Siinä kysymys." presented at Digitalia Summer School 2017.

