



Swiss Open Cultural Data Hackathon 2017: Final Report

Beat Estermann, 26 October 2017

Contents

1. Introduction.....	2
2. Main Objectives of the Hackathon.....	3
3. Achievement of Project Goals	3
3.1. Opening up cultural data and content for reuse and making them available at a central location....	3
3.2. Improving the visibility of Swiss heritage data and content at an international level.....	5
3.3. Promoting the re-use of cultural data / content (with a special focus on sustainability)	7
3.4. Fostering the exchange and cooperation among stakeholders from various backgrounds	7
3.5. Propagating the OpenGLAM principles within the Swiss heritage sector	9
3.6. Promoting the public visibility of OpenGLAM	10
4. Positive Developments	11
5. Potential for Improvement	11
6. Outlook	12
7. Project Resources	12
7.1. Financial resources.....	13
7.2. In-kind contributions.....	14
8. Concluding Remarks	15

Appendix A: Project Goals

1. Introduction

On 15-16 September 2017, the OpenGLAM Working Group of the Opendata.ch association and its partners organized the third edition of the Swiss Open Cultural Data Hackathon, preceded by two half-day workshops on the afternoon of the 14 September. This year's hackathon was kindly hosted by the University of Lausanne, represented by the Plateforme de cultures et humanités digitales (LaDHUL) and by the Bibliothèque cantonale et universitaire Lausanne (BCUL). Further partners included infoclio.ch, the Basel University Library, docuteam, and the E-Government Institute of the Bern University of Applied Science. In addition, the event was supported by several volunteers.

The hackathon was preceded by several pre-events in spring 2017:

- [Geneva Open Libraries](#) in the context of the [Open Geneva Hackathons](#) (12-13 May);
- [Pre-event "Digital Humanities"](#) at the University of Lausanne (7 June);
- [Zurich Open Archival Data Hackday](#) in cooperation with the Association of Swiss Archivists (9 June).

The present report provides a summary of event results as well as some insights with regard to future hackathons. It is based on an assessment of project goals, the results of an internal evaluation meeting, the past years' participants' survey, as well as a workshop on the future orientation of the hackathon held during this year's hackathon. An overview of the financial result is also provided. This year's participants' survey will again be carried out roughly half a year after the event, similarly to the survey of previous years.



Photo: Laurent Dubois, BCU Lausanne, CC BY-SA 4.0.

2. Main Objectives of the Hackathon

The main objective of the hackathon was to bring different stakeholder groups together, for them to interact around specific topics in order to share experiences, and to develop concepts and software prototypes. The event was to be used as a jumping board for the creation of software and other forms of data/content re-use that get some public visibility or have some other outside impact (e.g. by their use for research, in the context of Wikipedia/Wikimedia, or by facilitating the crowdsourcing of certain tasks). Last but not least, the hackathon was to be used as an opportunity to encourage Swiss heritage institutions to open up their data and content and to spread the word about the principles of OpenGLAM¹.

3. Achievement of Project Goals

The table in appendix A gives an overview of the goals that were set for this year's hackathon, the level of their achievement, the achievements in the previous years for comparison, as well as suggested targets for next year. Please note that not all targets could be assessed yet, as dissemination activities take more time and this year's participants survey has not been carried out yet.

The documentation of individual hackathon projects and an overview of the media coverage can be found on the event website².

Specific goals were set in six areas. In the following, we will shortly discuss the level of achievement of these goals in the context of the hackathon.

3.1. Opening up cultural data and content for reuse and making them available at a central location

The specified targets were fully achieved: 116 open datasets / collections from 60 Swiss institutions³ have so far been made available through the make.opendata.ch website (compared to 70 open datasets / collections from 40 institutions in the previous year). This has only been possible thanks to continuous and relentless efforts in this area, as convincing institutions to open up their data and content often takes considerable time, with efforts often bearing their fruits only after one or several years of dialogue with the institution.

Three indicators show that we are on a good track in this regard:

- The organizing team has been able to reach out to about as many new institutions as last year.
- Several institutions who have provided data/content in 2015 or 2016 have done so also in 2017, indicating their satisfaction with the hackathon and creating a positive feedback loop.
- Over 50% of the participants of the 2015 and 2016 hackathons indicated that the hackathon was (very) effective in getting/promoting access to cultural data.

3.1.1. Data catalogue platform strategy

One of the big changes this year regarding the presentation of datasets consisted in the migration of the OpenGLAM CH data catalogue to the opendata.swiss platform⁴ where data providers can manage the description of their data from their own user account. Alternatively, institutions are still offered the possibility to (temporarily) include their descriptions on the OpenGLAM CH account. The collaboration with the Swiss

¹ <http://openglam.org/principles/>

² <http://make.opendata.ch/wiki/event:2017-09>

³ These numbers do not include all the collections from the e-rara and e-manuscripta platforms, which were officially marked as Public Domain material in the course of this year; these platforms have been counted just as one dataset.

⁴ <http://opendata.swiss>

Federal Archives, who are hosting the portal, has so far been positive, and the organizing team is pursuing the discussions with regard to the improvement of the platform. Concretely, the hackathon team would like to see improvements to the platform in four areas:

- The ability to add references to standard licenses (e.g. Creative Commons licenses) in the metadata.
- The ability to add a picture to the metadata record (which is useful especially in the case of collections of content, like photographs, prints, scans of paintings, etc.).
- Adaptations to the functionality allowing the integration of data from opendata.swiss on third party websites (such as widgets showing up-to-date metadata from the portal).
- The ability to tag datasets at the level of the catalogue. At present, tags for datasets are managed exclusively by the organizations, which makes it awkward to tag a specific subset of datasets for a specific event (e.g. a hackathon). Furthermore, the automatic harvesting of some organizations is set up in a way that prevents the attribution of tags in language versions not covered by the harvesting routine.

3.1.2. Harmonizing data publication practices

Over the coming year, the organizing team should not only encourage the remaining institutions to move their catalogue entries from the OpenGLAM CH account to their own account, but also lead a conversation with the data providers on how to handle the heterogeneity of datasets and in order to develop best practices with regard to their publication.

Datasets in the heritage sector are rather diverse; they typically comprise one of the following types of data or a combination thereof:

- structured data describing or documenting some cultural or historical phenomena;
- collections of digital artefacts (photographs, scans of prints, paintings, etc.);
- metadata pertaining to collections of digital or physical artefacts;
- geodata services (e.g. historical maps available through a web map service);



Photo: Laurent Dubois, BCU Lausanne, CC BY-SA 4.0.

- ontologies, thesauri, or vocabularies;
- registers of ontologies, thesauri, or vocabularies;
- inventories or registers of heritage institutions or collections;
- administrative data from cultural institutions.

Another source of heterogeneity among datasets lies in their varying granularity:

- While some datasets comprise entire catalogues, covering the collections of many institutions (e.g. catalogues of platforms such as Swissbib or e-codices), others pertain to the collection of a single institution (sometimes also available through one of the platforms).
- While some digital collections comprise many thousands of artefacts, others only comprise a few dozens of them. Whether or not to break one collection up into many smaller ones is usually at the discretion of the data owner.

A further source of heterogeneity of datasets in the context of OpenGLAM lies in the fact that some collections may comprise both artefacts that are in the public domain or freely licensed and artefacts whose copyright status is unknown or the use of which is restricted. Here again, it is up to the data owner to decide whether or not to split up existing collections into smaller ones based on the licensing criterion. From a thematic point of view, splitting collections up based on this criterion may often not make much sense.

3.2. Improving the visibility of Swiss heritage data and content at an international level

The catalogue entries on the opendata.swiss platform are automatically harvested and made available on the European Data Portal⁵. While in the past, digital collections have been systematically made available also through the Open Collections Portal⁶ of the Open Knowledge network, the organizing team has stopped doing so, because the portal has been used by rather few organizations and because it is unclear to what extent the platform will be maintained in the longer term.

3.2.1. The role of Wikimedia Commons and Wikidata

Wikipedia is a very effective channel to increase the visibility of heritage data and content at an international level. In order to be used within Wikipedia, content needs to be uploaded to Wikimedia Commons, and data needs to be made available through Wikidata. Apart from the use of the data within Wikipedia and its sister projects, Wikidata is more and more also been used by heritage institutions worldwide as a hub for authority files and as a collaborative lab for data integration.

With a certain regularity, the annual hackathon brings to the fore new collections and datasets that are potentially valuable to the Wikimedia community. As in the previous years, there are about ten photo collections among the newly released datasets that are worthy of uploading to Wikimedia Commons. In some cases, the institutions are taking care of the upload themselves, in others they need to be provided with some guidance on how to go about it. In contrast to previous years, only a fraction of eligible collections has been uploaded to Wikimedia Commons. This issue should be addressed over the coming months.

For the first time, the organising team has carried out a census of datasets that would be good candidates for the ingestion into Wikidata: The approx. 15 datasets (out of 115) that would be eligible for ingestion into Wikidata contain data about people (actors, dancers, writers, photographers, painters, politicians, etc.), about historical monuments, heritage institutions, theatrical productions, or the world's animal and plant species. Three datasets have already been ingested into Wikidata, and 1-2 more are in the process of being ingested.

⁵ <https://www.europeandataportal.eu>

⁶ <https://openglam.org/open-collections/>

It should be noted however that visibility of heritage data and content does not so much arise directly from their publication on Wikimedia Commons or their ingestion into Wikidata, but from their use in Wikipedia or in other contexts. Thanks to the Open Cultural Data Hackathon there is a regular flow of new data and content that is released under free copyright licenses; however, in order to foster their use, communities need to be built around them. This is one of the preoccupations of the Swiss GLAM-Wiki Contact Group bringing together representatives of several heritage institutions as well as Wikimedia CH and opendata.ch.

There has been a radical change in this area over the past five years: While the main challenge a few years ago has been to convince institutions to open up their data and content (with online communities only waiting for their release), the challenge nowadays lies in growing the online communities and in attracting their attention to newly released datasets. The reasons for this mainly lie in the large quantities of heritage data and content that have been made available at a global level, in the modest community growth, as well as in platform interfaces and functionalities that – almost by definition – are lagging behind such developments.

3.2.2. The role of data portals in exposing heritage data and content at an international level

Apart from the measures described above, other efforts may be undertaken directly by the heritage institutions themselves in order to enhance the visibility of their data and content at an international level. The impact of the sole hackathon may be limited in this area, but the organizers can certainly sensitize the institutions for this issue. In this context, conversations should also be held with providers of Swiss heritage data portals, as to which are the best avenues to promote the use of Swiss heritage data and content at an international level, and how they could support this effort.



Photo: Laurent Dubois, BCU Lausanne, CC BY-SA 4.0.

3.3. Promoting the re-use of cultural data / content (with a special focus on sustainability)

This is a goal that is generally well achieved by means of the hackathon, as the number of hackathon projects and the number of responses to the 2015 and 2016 participants' surveys indicate: For almost 60% of the participants, the hackathon was (very) effective in helping them find out how data/content can be used, while more than 70% found that it was (very) effective in getting new inspiration or ideas.

However, over the years, a decrease in the number of hackathon projects that are pursued during the event can be observed (the average group size tends to increase). Also, compared to the first cultural data hackathon in 2015, which had also seen installations, performances, and offline games among the projects that were pursued, this year's and last year's editions saw a smaller variety of projects, mostly focusing on app development or data integration and enhancement. One could therefore easily get the impression this year that there have been fewer groundbreaking ideas than in the previous years. Also, a trend towards more data management and data enhancement projects could be observed, along with a tendency among hackathon participants to pursue the same or similar projects from one hackathon to the other (a development that was encouraged by holding pre-events in form of hackdays). These developments may be interpreted as a sign that projects are pursued with a longer-term focus; at the same time, the smaller variety of projects may indicate a certain lack of creativity.

The sustainability of hackathon projects is an aspect that has been an object of discussions among the hackathon organizers for several years. In last year's final report⁷, it was noted that cases in which a new team forms at the hackathon, starts developing a project, and keeps pursuing exactly that project after the end of the hackathon are rather the exception than the rule. The report also identified various alternative patterns by which the hackathon has a sustainable effect. Discussions around this topic boil down to three possible avenues that could be pursued in the future (some form of combination may also be possible):

1. The focus should not be laid so much on the sustainability of the projects themselves; the sustainable effect of the hackathon lies in the networking and in the exchange of knowhow and ideas.
2. The hackathon organizers could encourage institutions to commit (financially) to the further development of selected projects or set up a mentorship program that comes along with some financial support for selected projects.
3. The organizational set-up of the hackathon could be changed in a way that encourages teams to further pursue their projects after the hackathon – e.g. by organizing a "post-production" follow-up event, where the various teams present their completed projects.

In the past, the hackathon has performed rather well according to the logic of point 1, whereas developments resembling points 2 and 3 have occasionally happened on a voluntary and rather spontaneous basis – without much coordination effort by the organizing team.

3.4. Fostering the exchange and cooperation among stakeholders from various backgrounds

This is an area where the hackathon is doing very well, as is exemplified by the fact that over 80% of participants of the 2015 and the 2016 edition appreciated the hackathon's effectiveness in terms of meeting interesting people and fostering networking. There is also strong episodic evidence that this exchange and cooperation is actually happening within hackathon project teams. Table 1 below shows that we have again been able to attract a good mix of participants from various backgrounds and profiles. It should be noted however, that the representation of artists and designers among the participants has been significantly lower this year than in previous years; thus, the diversity of participants has somewhat declined.

⁷ https://glam.opendata.ch/wordpress/files/2015/12/GLAMHack2016_Final_Report_20160912.pdf

Table 1: Different categories of participants

Participant category	2015 edition (N = 49 of 107)	2016 edition (N = 94 of 105)	2017 edition (N = 94 of 98)
Female	19%	33% ↑ **	37%
Male	81%	67% ↓ **	63%
Data provider	35%	–	–
Data provider or expert	–	28%	32%
Software programmer	35%	25%	34%
Ideator	27%	25%	22%
Researcher	22%	31%	21%
Wikipedia editor	12%	11%	5%
Wikidata editor	–	–	5%
Artist	8%	7%	1% ↓ **
Designer	4%	13% ↑ *	3% ↓ **
Organizer	25%	–	–
Hackathon organizer	–	11%	11%
Other	12%	20%	18%

Changes marked * are significant at the 0.10 level; those marked ** are significant at the 0.05 level.



Photo: Laurent Dubois, BCU Lausanne, CC BY-SA 4.0.

3.5. Propagating the OpenGLAM principles within the Swiss heritage sector

The immediate targets in this area have been fully achieved. The hackathon provides a good occasion to reach out to heritage institutions who aren't committed to the OpenGLAM principles yet. Over the past year, we have also been noting an increased interest among Swiss heritage institutions to host (smaller) hackathon events themselves, indicating that the idea starts to stick. Furthermore, in the course of this year we have been able to present the concept of the hackathon and some hackathon projects at an international digital humanities conference (Bern, February 2017), at a public domain event, targeted at representatives of heritage institutions and other interested persons (Basel, April 2017), and at the information session of the Geneva Open Libraries event, targeted at heritage institutions and potential hackathon participants (Geneva, May 2017). Thus, the hackathon provided a good basis to promote the idea of OpenGLAM within the Swiss heritage sector.

This said, the hackathon is obviously not the only effort undertaken in this regard. The activities of the members of the Swiss GLAM-Wiki Contact Group and of the Swiss Association of Archivists clearly also contribute towards this goal – sometimes with important synergy effects with the hackathon.

3.5.1. The role of data portals

In the longer term, heritage data portal providers are important allies when it comes to promoting the OpenGLAM Principles⁸, notably by implementing the following policies and/or functionalities on their platforms:

- Do not apply any (legal or technical) usage restrictions on public domain content.
- Declare the copyright status (and possibly other usage restrictions) of the content provided or referenced on the portal (this should be part of the platform's metadata schema).
- Publish the metadata provided through the portal under a CC-0 Waiver⁹ and make it available for download.

In addition, portal providers may actively encourage and empower their data providers to freely license content and to increase its usability and visibility at an international level.

The main heritage portals with a geographical scope covering the whole of Switzerland are:

- e-codices (<http://www.e-codices.unifr.ch/>)
- e-manuscripta (<http://www.e-manuscripta.ch>)
- e-rara (<http://www.e-rara.ch/>)
- Memobase (<http://www.memobase.ch/>)
- fotoCH (<https://de.foto-ch.ch>)
- Swissbib (<https://www.swissbib.ch/>)
- Archives Online (<http://www.archivesonline.org>)
- Archives Quickaccess (<https://archives-quickaccess.ch/>)
- Museums Online (<http://www.museums-online.ch/>)
- HAN (<http://www.ub.unibas.ch/han/>)
- kirchen.ch/archive (<https://www.kirchen.ch/archive/>)

Apart from these, there are further portals with a regional scope (e.g. www.notrehistoire.ch), platforms that are still in the concept phase (e.g. a platform about Swiss paintings), portals focusing on a specific theme

⁸ <https://openglam.org/principles/>

⁹ <https://creativecommons.org/publicdomain/zero/1.0/>

(e.g. www.histoierurale.ch), as well as institutional platforms that also host data/content from third parties (e.g. www.e-pics.ethz.ch).

Over the past years, the organizing team has been in contact with several portal providers in order to sensitize them for the cause of OpenGLAM. There have been some notable advancements in this area, but overall progress in terms of the adoption of the OpenGLAM Principles by the portals remains rather slow. Thus, continued effort will be needed also in the future to promote OpenGLAM among portal providers.

3.6. Promoting the public visibility of OpenGLAM

The coverage in online media and blogs has been higher than in previous years. Also, thanks to the coverage of the pre-events, the topic of OpenGLAM has been covered over a longer period of time. As in 2016, coverage by traditional media has been low, reflecting the limited efforts that have been made in this area.

We also continued our activities on Facebook and Twitter. On Facebook, we were able to double the number of followers from 67 in March 2017 to 130 in October 2017. The absolute number of followers is however still rather low; more could be done in the area of social media if more resources were allocated to specifically promoting hackathon- and OpenGLAM-related content. On the other hand, the impact is probably higher if the various heritage institutions involved in OpenGLAM promote such content through their own channels. There is episodic evidence that such activities are indeed taking place; we have however not gathered any quantitative data to report such activities in a systematic fashion.

While the hackathon is a good means to promote the idea of OpenGLAM among a specialized public (heritage institutions, digital humanities, interested software programmers), it does not easily lend itself to reaching a wider audience. Individual hackathon projects as well as the concept of the hackathon have however been presented at several occasions where a more varied public could be reached. The same is probably true for some of the social media activities of heritage institutions.

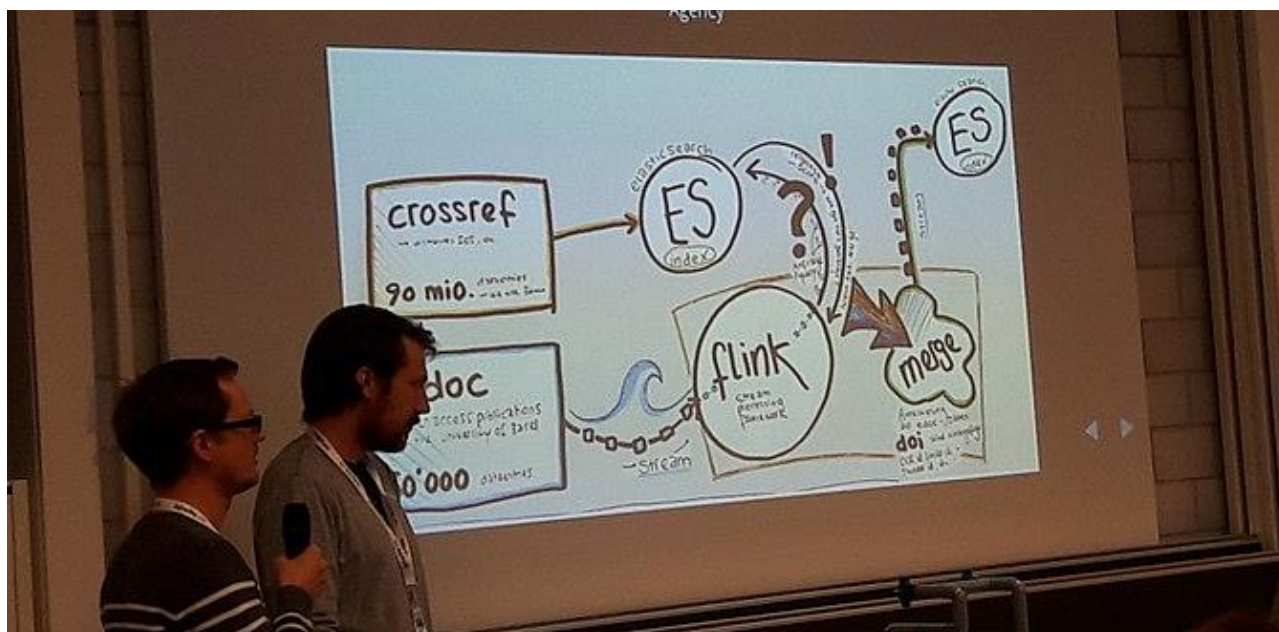


Photo: Beat Estermann, CC BY-SA 4.0.

4. Positive Developments

As has been noted at the internal evaluation meeting and at the future workshop, several positive developments could be observed at this year's hackathon:

- The organization of pre-events has started to develop its own dynamics: the Zurich Archival Hackaday was initiated by the Association of Swiss Archivists and was organized by several archives in the Canton of Zurich, while the Geneva Open Libraries event was made possible thanks to the strong involvement of the UN Library and Bodmer Lab. The organizing team welcomes this development and hopes that more such events will be organized in the future at the initiative of other members of the GLAM community; in fact, these different events reinforce each other both by attracting more people and by encouraging more institutions to open up their datasets.
- Several organizations have started to contribute staff time to the organization of OpenGLAM events.
- Several people attended the hackathon in groups, and from several institutions that had been involved in the organization of earlier hackathons, new people got involved in the hackathon. Thus, there seems to be a strong word-of-mouth effect in some communities and organizations.
- We were able to significantly cut catering costs while maintaining a high level of quality.
- The extension of the side programme by adding half-day workshops on Thursday afternoon was very well received. The workshops not only attracted more participants than expected (almost 40% of hackathon participants attended one of the Thursday workshops), but they also seemed to have a strong impact on the projects that were pursued during the hackathon itself.

5. Potential for Improvement

Potential for improvement has been identified in several areas. The most notable points that were brought up during the internal evaluation meeting were:

- Someone among the team should be in charge of the communication activities (substantial resources should be foreseen for this task).
- The way hackathon groups are managed should be re-considered. This year, two hackathon groups grew rather big. Maybe some group coaching would help to keep all the participants "on board".
- We should try to get more variety among the projects that are pursued during the hackathon: This year, fewer hackathon projects were worked on than in the previous years; there was again a smaller variety of projects than during the first edition of the hackathon in Bern, and there have been fewer groundbreaking ideas. Maybe it would be helpful to have a more structured idea pitching session.
- This year, only few members from the host institutions (staff and students) participated in the hackathon; in future editions, it would be nice to see stronger involvement by members of the host institutions and local communities.

6. Outlook

During the hackathon, a workshop was held to discuss the future orientation of the hackathon. Thereby the following points were discussed:

- The fact that next year's hackathon will be hosted by the Swiss National Museum in Zurich should be used to especially reach out to museums, which so far have been clearly under-represented within the OpenGLAM community (and among data providers).
- The focus of the hackathon could be widened (again) to include not only apps running on standard hardware, but also museum installations and other forms of data use that allow the engagement of the public.
- We could encourage projects that involve user engagement / visitor experience.
- A follow-up event could be organized where hackathon artefacts are presented to the public. This may also give the participants an extra incentive to further pursue their projects.
- In this context, a future cooperation with Museomix¹⁰ was also discussed and generally welcomed.
- At the same time, it was stressed that it is important not to estrange the existing community of data providers (mainly libraries and archives) and hackathon participants.

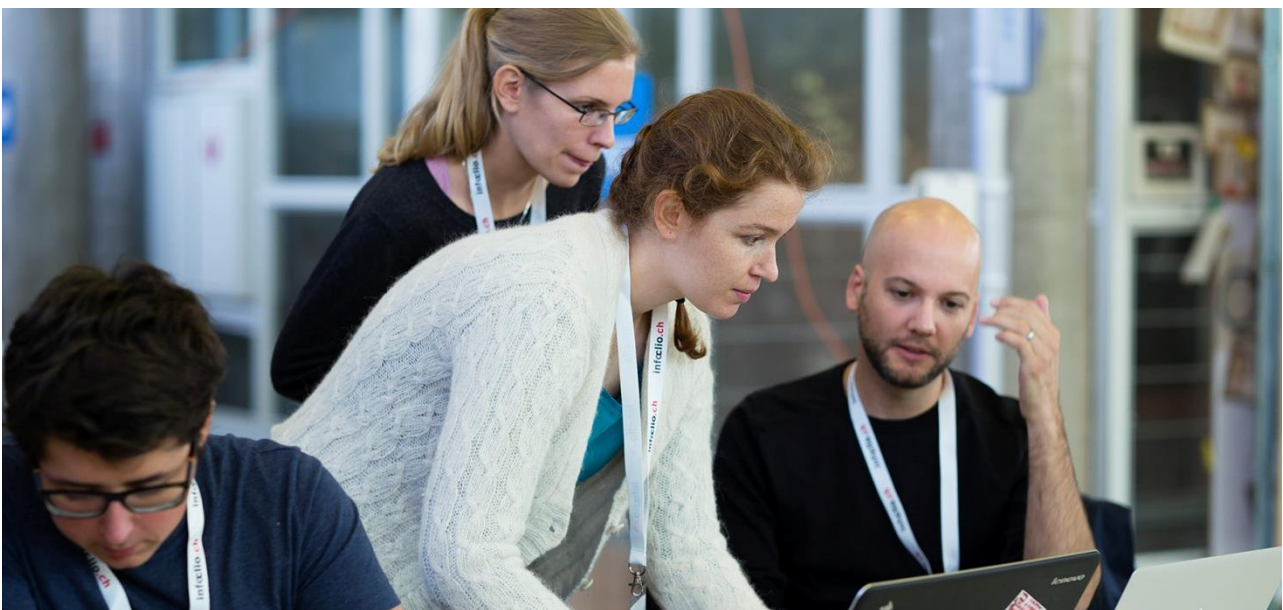


Photo: Laurent Dubois, BCU Lausanne, CC BY-SA 4.0.

¹⁰ <http://www.museomix.ch/>

7. Project Resources

7.1. Financial resources

Table 2 below gives an overview of the financial resources of the project:

Table 2: Overview of financial resources

	Budget	Effective Costs / Revenues
Expenses	- 55'000	- 34'936
Catering	18'000	7'503
Accommodation (Youth Hostel)	5'000	3'578
Project Coordination	20'000	20'000
Association opendata.ch (10% of revenues)	5'500	3'855
Varia / Incidentals	6'500	0
Revenues	+ 55'000	+ 38'550
Internal Sponsors	18'000	18'000
External Sponsors	37'000	20'550
Balance	0	+ 3'613

The project was kindly supported by the following sponsors:

- Bibliothèque cantonale et universitaire Lausanne BCUL (10'000 CHF)
- Infoclio.ch (6'000 CHF)
- Université de Lausanne, Institut des sciences sociales (2'000 CHF)
- Ernst Göhner Stiftung (10'000 CHF)
- Migros Kulturprozent (5'000 CHF)
- Wikimedia CH (2'000 CHF)
- Bern University of Applied Sciences, E-Government Institute (1'550 CHF)
- docuteam (1'000 CHF)
- Swiss National Library (1'000 CHF).

Note that the effective costs for catering and accommodation have been lower than expected given the fact that the effective number of participants was lower (98) than the one budgeted for (150) and that the catering prices in general were lower in Lausanne than in Basel.

The resulting benefit remains on the account of the opendata.ch association and is earmarked for future hackathons or similar events related to cultural heritage.

7.2. In-kind contributions

The organization of the event would not have been possible without the substantial in-kind contributions made by several individuals and organizations, most notably by contributing their volunteer and/or staff time:

- Marianna Schismenou, Dominique Vinck (Université de Lausanne)
- Ramona Fritschi (BCUL)
- Isabelle Lucas, Jan Baumann (infoclio.ch)
- Frédéric Noyer (docuteam)
- Lionel Walter (Basel University Library)
- Oliver Waddell, Oleg Lavrovsky, Martin Grandjean (opendata.ch)
- Beat Estermann (opendata.ch / Bern University of Applied Sciences)

but also by allowing us to use their infrastructure and/or equipment:

- University of Lausanne (hosting of the hackathon)
- UN Library (hosting of pre-event)
- Zurich Cantonal Archives (hosting of pre-event)
- Opendata.ch (online and hardware infrastructure, financial administration).

In addition, we would like to thank the countless people involved in the organization of the pre-events.



University of Lausanne, “Genopode” building. Photo: Rama, CC BY-SA 3.0 France.

8. Concluding Remarks

This year's hackathon has again been a great success, and both the participants and the organizers are looking forward to the next edition of the event.

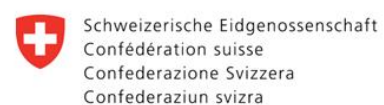
As the present report shows, there is room for improvement in various areas and there are rather concrete ideas of how to reach out to new target groups and to preserve the innovative spirit of the hackathon in the future.

We would like to thank all our sponsors and partners for supporting the event and are looking forward to pursuing further opportunities for cooperation.

Our Sponsors:



ERNST GÖHNER STIFTUNG



Eidgenössisches Departement des Innern EDI
Schweizerische Nationalbibliothek NB