Open Culture Data position paper

-

Open Data on the Web

Lotte Belice Baltussen - Netherlands Institute for Sound and Vision
Maarten Brinkerink - Netherlands Institute for Sound and Vision
Nikki Timmermans - Kennisland
Maarten Zeinstra - Kennisland

Abstract

Open Culture Data started as a grassroots movement at the end of 2011, with the aim to open up data in the cultural sector and stimulate (creative) reuse. In this context, we organised a hackathon, which resulted in the creation of 13 Open Culture Data apps. After this successful first half year, a solid network of cultural heritage professionals, copyright and open data experts and developers was formed. In April 2012, an Open Culture Data masterclass started in which 17 institutions get practical, technical and legal advice on how to open their data for re-use. Furthermore, we organised an app competition and three hackathons, in which developers were stimulated to re-use Open Cultural Datasets in new and innovative ways. These activities resulted in 27 more apps and 34 open datasets. In this paper we share lessons-learned, that will inform heritage institutions with real-life quantitative and qualitative experiences, best practices and guidelines of their peers with opening up data and the ways in which this data is reused. Since the open culture data field is still relatively young, this is highly relevant information needed to stimulate other to join the open data movement. To this end, we are already taking steps to cross the borders and let Europe know about the initiative, on both a practical and a policy level.

1. Introduction

Open data is data that can be accessed, distributed and reused by everyone, even for commercial purposes, without the need of explicitly asking the data owner for permission. Open data is also high on the digital agenda in Europe and the Vice-President for the Digital Agenda of the European Commission Neelie Kroes has even made the following call-to-action:

> I urge cultural institutions to open up control of their data...there is a wonderful opportunity to show how cultural material can contribute to innovation, how it can become a driver of new developments. Museums, archives and libraries should not miss it. (Kroes, 2011, p. 6)

GLAMs increasingly realise that open access to data helps drive users to online content, for instance by providing content for reuse on Wikipedia articles. Hence open data supports cultural institutions in the fulfilment of their public mission to open up access to our collective heritage, not just through their own channels, but outside these as well. Secondly, it stimulates collaboration in the GLAM world and beyond, which allows the creation of new services and supports creative reuse of material in new productions. Kennisland and the Netherlands Institute for Sound and Vision, together with Hack de Overheid (Hack the Government) launched the 'Open Culture Data' (Open Cultuur Data in Dutch) initiative in September 2011. We are a network of cultural professionals, developers, designers, copyright specialists and open data experts with the aim to make cultural datasets available under open conditions and stimulate the creation of useful and innovative applications in which these datasets are incorporated. The initiative promotes dialogue and the sharing of experiences on how to get more cultural data openly available. In this paper, we share Open Culture Data's lessons-learned, from developing a network of interested institutions, educating them in how to open up their collections and to connecting with a community of developers. Finally, we talk about our future plans, with a focus on how to gather informative metrics of the impact of opening up.

2. What Is Open Culture Data?

Open Culture Data started in September 2011 by defining guidelines in order to make clear to contributors what principles they should at least adhere to:

1   Open Culture Data is knowledge and information of cultural institutions, organisations or initiatives about their collections and/or works
2   Everyone can consult, use, spread and reuse Open Culture Data (through an open license or by making material available in the Public Domain).
3   Open Culture Data is available in a digital (standard) format that makes reuse possible.
4   The structure and possible applications of Open Culture Data are documented in a data blog.
5   The provider of the Open Culture Data is prepared to answer questions about the data from interested parties and respects the efforts that the open data community invests in developing new applications.

Open Culture data makes a clear distinction between content and metadata. All digitized cultural objects like scanned paintings, photographed objects and digital texts are defined as content. All descriptive information about the object are called metadata, e.g. the name of the creator, the year of creation, the size of the object, a description, etc. The accepted open licenses compliant with the rules above are the Creative Commons Public Domain Dedication (CC0)[1] for metadata and for content the Public Domain Mark[2] if copyright has expired; Creative Commons Attribution (CC BY)[3] or

---

[1] http://creativecommons.org/publicdomain/zero/1.0/

[2] http://creativecommons.org/publicdomain/mark/1.0/

[3] http://creativecommons.org/licenses/by/3.0/

Creative Commons Attribution-ShareAlike (CC BY-SA)[4] in cases where the organisation has (cleared) the rights. There are three reasons for opting to use these specific open licenses. First of all, Creative Commons licenses are ported and used worldwide. Furthermore, these choices are in line with rights requirements of other projects that support open data. Europeana has made CC0 for metadata a prerequisite for data providers (Europeana Office, 2012) and the board of directors of the Digital Public Library of America recommended to follow the CC0 policy for metadata last January (Pekel, 2013). Thirdly, regarding content, CC BY and CC BY-SA are the only two licenses that are compliant with the open content rules of Wikimedia Commons, the media archive of Wikipedia. They consider the other four Creative Commons licenses to be too restrictive, because they do not permit commercial reuse and / or making derivative works. (Wikimedia contributors, 2013)

3. Open Culture Data: From Ad-Hoc activity To A Solid Network
After a successful ad hoc and experimental start-up period from September 2011-December 2011 that resulted in 8 open datasets, 13 apps and a win of an app made with open culture data in the largest Dutch app competition (Baltussen et al., 2012), we concluded that we should continue and expand the Open Culture Data activities. In early 2012, the initiative was adopted by the digitization project Images for the Future[5] and Creative Commons Netherlands[6]. A common set of questions on how to open up kept emerging from professionals from various types of cultural institutions. Therefore, we decided to develop two activities that would connect data owners, app developers and policy makers: an open data masterclass and our own app competition.

3.1 The Masterclass: A Fast Track To Opening Up
The Open Culture Data masterclass started in April 2012 and ended in June 2012. A total of 24 Dutch cultural institutions signed up, of which seventeen took part in the end. Representatives from the institutions who already opened up data in 2011 acted as coaches in the masterclass. Various types of organisations - large and (very) small - signed up: five archives (regional and national), one library, six museums, four knowledge institutes, one sector institute and one Dutch tourism project. The participants had various levels of knowledge about open data, ranging from people whose organisations already had open data working groups, to those whose management was not yet fully convinced of the benefits of opening up. The masterclass touched upon the following topics which guided them through the process necessary to open datasets: Building blocks of copyright; Technology and tools (from open licensing to APIs); Reuse and applications; Benefits and risks and Hackathons.

As part of the course material for this masterclass three (openly licensed) white papers were written about legal and technological issues and useful tools to publish collections and metadata. (Zeinstra & Timmermans, 2012a, b and c) Each participant was assigned an experienced open data coach, to aid them with any questions or issues that would arise. As a direct result from the masterclass, 10 datasets from various institutions were opened up. Parallel to the masterclass, other institutions indicated that they wanted to join up as well, and at the time of writing, there are now 34 Open Culture Datasets.

4.3 Creative Reuse: Open Culture Data App Competition
In order to create interest in the open culture datasets and stimulate reuse, we launched our own app competition that ran from 16 June 2012 - 31 December 2012. We set out three main challenges for the developer community:

1   Create applications that expand audience reach (online, offline, on-site)
2   Create applications in which the audience is reached in novel ways.

---

[4] http://creativecommons.org/licenses/by-sa/3.0/
[5] http://imagesforthefuture.com/
[6] http://creativecommons.nl/

3   Create applications that connect different datasets.

There were four awards (Gold, Silver, Bronze, Dutch National Archives Prize) and a total prize pool of €7,500. The overall Gold winner could win €3,000. Three hackathons were held: two big ones in June 2012 and October 2012 and a smaller one specifically aimed at reuse of the data in games right before the December 31 deadline. Furthermore, many presentations to raise awareness of the competition and Open Culture Data itself were held between June and December, for instance at a developers conference, the Wikimedia Netherlands conference and Communication and Media Design students. In total 27 apps were submitted for the competition, a lot more than we had bargained for, and the quality of the apps exceeded our expectations. A jury of five (two GLAM and two open data professionals, and one developer) picked the winners. First prize went to Muse app made by Femke van der Ster, Peter Henkes, Jelle van der Ster. Muse app allows you to create your own work of art with cut-outs from world-famous old masters. You can bring the cut-outs to your own canvas, pinch, move, duplicate them to make a collage and share your masterpiece online.

Figure 1: Gold prize winner Muse app, developed by Femke van der Ster, Jelle van der Ster and Peter Henkes (*http://www.museapp.org/*).

The Silver went to Histagram[7] by Frontwise (Richard Jong), an app where you can make digital postcards based on historical pictures. The third prize went to SimMuseum[8] by Hay Kranen, a web game in which you can play a museum director, collect work of arts and build your own museum. Finally, the Dutch National Archives prize went to Tijdbalk.nl[9] made by Arjan den Boer. Users can make their own timeline with historical photos photos and add their own content as well.

4. Lessons-Learned
We learned many valuable lessons in our second year, but four main themes emerged:

1. Innovators lead the way: By gathering the right group of professionals in the cultural domain who believed in the (potential) power of open and were willing to experiment, we created a small but very powerful vanguard. For example: When the prestigious Dutch Rijksmuseum joined the initiative, this inspired other institutions like the National Museum of Antiquities to also participate.
2. Create practical examples: The fact that cultural institutions are hesitant to join the open data movement has a lot to do with either a lack of knowledge regarding copyright and IPR or a fear of the

---

[7] http://www.frontwise.nl/lab/histagram/

[8] http://simmuseum.haykranen.nl/

[9] http://tijdbalk.nl/

consequences for their current way of operating. Fear that their business model might be endangered and fear of people abusing their data, or re-using it for purposes they don't agree with, like misrepresenting the data. These fears are not per se grounded in fact and experience (see for instance Verwayen et al., 20) and it withholds institutions from what you can gain by opening up, like experimenting with innovative concepts for new services or applications. We have learned that by putting open culture data in practice and actively stimulating the reuse of the data, cultural institutions can be convinced to join the movement.

3. Thinking about open culture data requires a multidisciplinary perspective: Many cultural institutions have particular ideas about new applications and services for their data. But this is only one way of looking at it. We have learned that connecting cultural institutions with the 'outside world', the world of hackers, designers, students, but also other data providers and commercial companies is not only a lot of fun, but is also very helpful to institutions in finding new ways to make arts and culture meaningful in the digital era.

4. Re-use requires specific parameters: Developers do not just build an app, since their time is valuable and they need to feel a connection either with the data or the open data community to dedicate themselves. Therefore, during the competition period, we continuously raised awareness in the developer community and ask them what they wanted and needed, and were always available for questions and input. Also, we have seen in both the national and international open data movement that a majority of apps developed within competitions will disappear, because they cannot be sustained and often remain in the prototype stage. This is why Open Culture Data will focus on the development of sustainable business models for both the developers and the cultural institutions in the future. This issue is an important focal point in the projects Apps4EU[10] and Europeana Creative[11], in which the Netherlands Institute for Sound and Vision and Kennisland are participating. Finally, not all data is equally popular. Out of the 34 available datasets, 14 were actually used. The Rijksmuseum dataset, with its well-known collection, 125,000 high-res images and easy-to-use API, was the most popular and used in 9 out of 27 apps submitted. In general, collections that offer both metadata and content are more popular than open datasets that only contained metadata. When metadata-only data was used, this was usually done in combination with other datasets that did have images or videos.

5. Future Perspectives: To Measure Is To Know And Spreading The Word Internationally
Now that many cultural datasets from early adopters in the Dutch cultural heritage sector can be reused, a need arises to measure the effects of opening up culture. Although measuring online success is gaining ground in the GLAM sector, specific tools for measuring the effects of opening up data are far and few between (Finnis et al., 2011). The need for a measurement model for the impact of opening up culture was expressed by various stakeholders. Since the open culture data field is relatively young, this highly relevant information is still often in an anecdotal stage. To improve this situation and to contribute to the accumulation of data about the impact of opening up cultural data and content, Open Culture Data will develop a measurement model (GLAMetrics) to gather evidence about the effects of open distribution. Based on our experiences with testing the measurement model we will produce a survey that we can circulate widely among GLAMs that are openly distributing their data and/or content. It will be a challenge to gather this data from technically diverse sources (different databases, Google Analytics, customised web metric systems) and make sure the data is aligned so it can be properly compared. Also, not all institutions will want to share all their metrics openly, since this can contain sensitive information. Still, we already have already successfully tested a first version of the GLAMetrics model within the Open Culture Data network and as a result are confident great strides will be made in making the impact of opening up truly measurable.

The results of our open culture impact analysis is not the only outcome from the Open Culture Data activities in the Netherlands that we hope to share with the international open culture data movement. For the P2PU (*http://p2pu.org*) grassroots open education project we are developing free

---

[10] http://ec.europa.eu/information_society/apps/projects/factsheet/index.cfm?project_ref=325090/

[11] http://pro.europeana.eu/web/europeana-creative/

online courses for GLAM-professionals who are interested in opening up their data or content, based on our master class. The first actual European Open Culture Data spinoff was launched in Belgium September 2012 and will organize its first hackathon in the spring of 2013. Another spinoff in Finland is also in the works. (Pekel, 2012)

We believe that institutions from all sizes and with low and big budgets alike can join the open movement. The coming years, we hope to see a vast increase of freely sharable and reusable collections that inspires an ever-growing, connected GLAM community.

Acknowledgements

References

Baltussen, L.B. e.a., 2012. Open Culture Data position paper - Using Open Data workshop, Hilversum / Amsterdam: Netherlands Institute for Sound and Vision / Kennisland / Hack de Overheid. Available at: http://www.w3.org/2012/06/pmod/pmod2012_submission_22.pdf.

Europeana Office (2012). "Europeana's Huge Cultural Dataset Opens for Re-use." Published September 12, 2012. Consulted Consulted January 24, 2013. http://pro.europeana.eu/web/guest/press-release?p_p_id=itemdetailsportlet_WAR_europeanaportlet_INSTANCE_FX4c&p_p_lifecycle=1&p_p_state=normal&p_p_mode=view&p_p_col_id=column-2&p_p_col_count=1&_itemdetailsportlet_WAR_europeanaportlet_INSTANCE_FX4c_itemId=1284453&_itemdetailsportlet_WAR_europeanaportlet_INSTANCE_FX4c_javax.portlet.action=setItemId.

Finnis, J. et al., How to Evaluate Online Success? A New Piece of Action Research. In J. Trant and D. Bearman (eds). Museums and the Web 2011: Proceedings. Toronto: Archives & Museum Informatics. Published March 31, 2011. Consulted October 23, 2012. http://conference.archimuse.com/mw2011/papers/how_to_evaluate_online_success.

Kroes, N. (2011). "Foreword: Culture and Open Data: How Can Museums Get the Best from their Digital Assets?" Uncommon Culture: From Closed Doors to Open Gates, vol. 2, no. 1/2. Consulted December 10, 2012. http://www.firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/UC/article/view/3771/3053.

Pekel. J. (2012). OpenGLAM workshop at the OKFestival. OpenGLAM blog. Published September 27, 2012. Consulted February 28, 2013. http://openglam.org/2012/09/27/openglam-workshop-at-the-okfestival-2/

Pekel, J., (2013). Digital Public Library of America recommends CC0. OpenGLAM blog. Published January 21, 2013. Consulted February 3, 2013. Available at: http://openglam.org/2013/01/21/digital-public-library-of-america-recommends-cc0/.

Verwayen, H., M. Arnoldus, & P. B. Kaufman. (2011). "The problem of the yellow milkmaid. A business model perspective on open metadata." Den Haag: Europeana. Consulted December 10, 2012. http://version1.europeana.eu/web/europeana-project/whitepapers/

Wikimedia contributors. (2013). Commons:Licensing. Wikipedia, The Free Encyclopedia. Published January 14, 2013. Consulted January 21, 2013. http://commons.wikimedia.org/wiki/Licensing.

Zeinstra, M. & N. Timmermans (2012a). Auteursrecht en Open Data in de Culturele Sector, Amsterdam: Creative Commons Nederland. Available at: http://www.opencultuurdata.nl/wp-content/uploads/2012/07/Auteursrecht-en-Open-Data-in-de-Culturele-Sector.pdf.

Zeinstra, M. & N. Timmermans (2012b). Creative Commons-licenties en methoden voor het openstellen van culturele data, Amsterdam: Creative Commons Nederland. Available at: http://www.opencultuurdata.nl/wp-content/uploads/2012/07/Creative-Commons-licenties-en-methoden-voor-het-openstellen-van-culturele-data.pdf.

Zeinstra, M. & N. Timmermans (2012c). Open Data: Delen, Verbinden, Verrijken, Amsterdam: Creative Commons Nederland. Available at: http://www.opencultuurdata.nl/wp-content/uploads/2012/07/Open-Data-Delen-Verbinden-en-Verrijken.pdf.