**<https://www.beeldengeluid.nl/en/blogs/research-amp-development-en/201506/dutch-cultural-heritage-reaches-millions-every-month>**

**Dutch Cultural Heritage Reaches Millions Every Month**

Submitted by Maarten Brinkerink on 17 June 2015

***Image****: Een menigte aanschouwt een komeet door Jan Luyken (1698)****Collection****: Amsterdam Museum, CC-0.*

**The cultural sector increasingly makes its collections available as *open data* and *open content*. These types of initiatives bring along the growing need of measuring their impact. On either a national or international level, there currently is no single body that tracks this type of data across collections. In 2014, the [Open Culture Data](http://www.opencultuurdata.nl/english/%22%20%5Co%20%22Open%20Cultuur%20Data%20English%20info%22%20%5Ct%20%22_blank) network therefor started an exploratory research project on the (im-)possibilities of measuring the impact of open cultural data. The project was called *GLAMetrics* - metrics for gallery, library, archive and museum collections.**

This initiative meant the beginning of a quantitative analysis of the consequences of opening cultural data - an evolution that affects the entire sector, both nationally and on an international level. This blog post presents the initial outcomes of our research into the reach and reuse of culture heritage from The Netherlands through Wikimedia projects.

*Wikimedia projects* are the different projects that come out of the Wikimedia community. Among them we find the different language versions of Wikipedia - such as [en.wikipedia.org](http://en.wikipedia.org/%22%20%5Co%20%22Wikipedia%20English%22%20%5Ct%20%22_blank) and [nl.wikipedia.org](http://nl.wikipedia.org" \o "Wikipedia Nederlands" \t "_blank) - and projects such as [WikiSource](https://en.wikisource.org/%22%20%5Co%20%22WikiSource%22%20%5Ct%20%22_blank) and [WikiData](https://www.beeldengeluid.nl/en/WikiData%22%20%5Co%20%22WikiData%22%20%5Ct%20%22_blank).

**Methodology**

In order to be reusable within Wikimedia projects, open culture data sets need to be published as open content on the media repository [Wikimedia Commons](https://commons.wikimedia.org/%22%20%5Co%20%22Wikimedia%20Commons%22%20%5Ct%20%22_blank). In October 2014 we set up and distributed a survey to [all members of the Open Culture Data network](http://www.opencultuurdata.nl/netwerk/%22%20%5Co%20%22Open%20Cultuur%20Data%20netwerk%22%20%5Ct%20%22_blank) to inventorise which of their open culture data had been added to Wikimedia Commons.

Thirty representatives from institutions in the network filled out this survey. Eleven respondents currently have one or multiple open culture data sets on Wikimedia. Three institutions indicated they're currently working on their first publication.

Subsequentially, we collaborated with Wikimedia Netherlands to complete, as far as possible, the overview of Dutch cultural institutions on Wikimedia.

Wikimedia offers various publicly available instruments to gather data on the reach and reuse of materials within the various Wikimedia projects. From November 2014 onwards, Open Culture Data has applied these measurement instruments for Dutch institutions on Wikimedia Commons. More specifically, we used the tools BaGLAMa 2 and GLAMorous, both created by [Magnus Manske](https://twitter.com/magnusmanske%22%20%5Co%20%22Magnus%20Manske%20on%20Twitter%22%20%5Ct%20%22_blank).

* [BaGLAMa 2](https://tools.wmflabs.org/glamtools/baglama2/%22%20%5Co%20%22BaGLAMa%202%22%20%5Ct%20%22_blank) shows on which Wikimedia project pages content from Wikimedia Commons is being reused and how often these pages are requested.
* [GLAMorous](https://tools.wmflabs.org/glamtools/glamorous.php%20target%3D) shows per set or collection how much material is available for reuse and how often this happens.

As a sidenote to these instruments: Wikimedia doesn't currently measure mobile traffic well. Wikimedia also doesn't discern between page consultations by visitors or by machines - such as search engines that perform indexing. According to estimates this constitutes up to 15% of all traffic. Also, Open Culture Data was not able to de-duplicate Wikimedia project pages that use materials from more than one institution. Our assumption is that these two deviations cancel each other out and result in the numbers not being lower than what's mentioned below. It is our expectation that Wikimedia will share more data about reach and reuse in the future, such as anonimised data about user behaviour on the pages that use Dutch heritage content. This would give us a better insight into how much time and attention users spend on consulting specific heritage objects.

**Preliminary Outcomes**

From November 2014 onwards (the moment we started recording data) there were 23 Dutch heritage institutions who provided one or multiple collections for reuse in Wikimedia projects by publishing them on Wikimedia Commons.

Some institutions have had a presence on Wikimedia for only a few months: the [Catharijneconvent museum](https://commons.wikimedia.org/wiki/Category%3AMedia_contributed_by_Museum_Catharijneconvent%22%20%5Co%20%22Museum%20Catherijneconvent%20objects%20on%20Wikimedia%20Commons%22%20%5Ct%20%22_blank) joined in February as the 24th institution and the [Textielmuseum](https://commons.wikimedia.org/wiki/Category%3AMedia_contributed_by_Textielmuseum%22%20%5Co%20%22Textielmuseum%20objects%20on%20Wikimedia%20Commons%22%20%5Ct%20%22_blank) in April as the 25th. At the same time, the first Dutch institution on Wikimedia Commons, the [Tropenmuseum](https://commons.wikimedia.org/wiki/Category%3AImages%20from%20the%20Tropenmuseum%22%20%5Co%20%22Tropenmuseum%20objects%20on%20Wikimedia%20Commons%22%20%5Ct%20%22_blank), has been providing content for reuse for more than 56 months.

To date, close to 580,000 Dutch digital heritage objects have been added to Wikimedia Commons. This means that from the total collection of media items on Wikimedia Commons - [close to 24.5 million](http://commons.wikimedia.org/wiki/Commons%3AMIME_type_statisticshttp%3A//commons.wikimedia.org/wiki/Commons%3AMIME_type_statistics%22%20%5Co%20%22Wikimedia%20Commons%20statistics%20per%20MIME%20type%22%20%5Ct%20%22_blank) - around 2.4% consists of Dutch digital heritage. The large majority of this Dutch offer are images, but it also holds close to 2,000 audio recordings and 4,500 videos.

Thanks to GLAMetrics we now know quite a bit more about the reach and reuse of these materials:

* In the first quarter of 2015 the objects of these institutions were used on approximately 76,000 Wikimedia project pages. During the observed quarter, this number has grown by about 2.5%.
* In the first quarter of 2015, these pages were requested more than 200 million times, or approx. 67.5 million consultations per month. The Wikimedia projects together receive [approximately 20.5 billion consultations per month](http://reportcard.wmflabs.org/%22%20%5Co%20%22WMFLabs%20Report%20Card%22%20%5Ct%20%22_blank), so the portion of pages using Dutch heritage is approximately 0.3%.
* These pages together reuse close to 37.500 unique objects, or close to 7% of the total offer.
* In total, Dutch digital heritage objects have been reused close to 100.000 times on a Wikimedia project page.

For the entire measurement period, Wikimedia also offers data about the number of consultations for the pages that contain selected objects. Although not each and every Dutch heritage collection has been measured from its point of origin onwards (the difference ranges from just a few to an impressive 56 months), the outcome already is quite impressive: pages reusing Dutch digital heritage have been consulted 1.9 billion times in total!

**GLAM Expectations**

Around 7% of the total combined total of Dutch digital heritage objects on Wikipedia is currently being used on one or several Wikimedia project pages. Based on the assembled data we can pronounce a few preliminary statements for institutions that are considering opening up (part of) their collections via Wikimedia Commons.

* Reuse differs among collections. For some collections we see that up to 50% is being reused, while others experience no reuse at all. Especially in the initial phases a reuse total of 7% appears to be a realistic expectation for digital heritage.
* For these 7% of reused materials, for each digital object one can expect a reach of more than 2.100 consultations per month. On a yearly basis, this translates into 25.000 consultations per object.
* The exact impact is influenced by the extent to which the institution stimulates reuse by communicating with the community and organising activities.
* Based on the above, an institution can, with a donation of 1.000 objects, expect a monthly reach of up to 150.000 consultations of pages holding their materials.

**Follow-up**

As a follow-up on this first blog post, we intend to give quarterly updates on how the reach and reuse of Dutch cultural heritage materials on the various Wikimedia projects develops. We also hope to present increasingly broad outcomes as we gather more data along the way.

We'll investigate if we can gather data from older collections on Wikimedia retroactively to identify developments on the middle and long term. We also aim to compare the use on different Wikipedia language versions and other Wikimedia projects and to measure what percentage of the totality of Wikipedia is being enriched with Dutch digital heritage. Finally, we aim to study the influence of activities around content donations or heritage institutions to Wikimedia on reuse (as, for instance, organising [edit-a-thons](http://en.wikipedia.org/wiki/Edit-a-thon%22%20%5Co%20%22%20Edit-a-thon%20From%20Wikipedia%2C%20the%20free%20encyclopedia%22%20%5Ct%20%22_blank)).

In accordance with Open Culture Data's vision, [all the data assembled for this investigation](https://docs.google.com/spreadsheets/d/1v39zKUg0_BUuokjWzpVv42VnfJcSEPHYjptFM5KLAbI/edit?usp=sharing" \o "Google Spreadsheet containing the research data" \t "_blank) have been made available for reuse under a [CC-0 license](https://creativecommons.org/publicdomain/zero/1.0/deed.nl%22%20%5Co%20%22Creative%20Commons%20Public%20Domain%20deed%22%20%5Ct%20%22_blank).

We are highly interested to [hear your feedback, suggestions or further analyses](http://www.opencultuurdata.nl/contact/%22%20%5Co%20%22Open%20Cultuur%20Data%20contact%20form%22%20%5Ct%20%22_blank)!

**Colophon**

Written by Maarten Brinkerink (Netherlands Institute for Sound and Vision) with thanks to Lotte Belice Baltussen, Jesse de Vos, *Kennisland*'s Maarten Zeinstra and *Open State Foundation*'s Tom Kunzler for their suggestions.

This post originally appeared on the [Open Cultuur Data blog](http://www.opencultuurdata.nl/2015/05/glametrics-nederlands-cultureel-erfgoed-op-wikipedia-bereikt-maandelijks-miljoenenpubliek/%22%20%5Co%20%22GLAMetrics%3A%20Nederlands%20cultureel%20erfgoed%20op%20Wikipedia%20bereikt%20maandelijks%20miljoenenpubliek%22%20%5Ct%20%22_blank) in Dutch and was translated to English by Erwin Verbruggen.

**More info**

For more info about the Open Culture Data initiative, see:

* The [What, Where, How and Why](http://www.opencultuurdata.nl/english/%22%20%5Co%20%22English%20info%22%20%5Ct%20%22_blank) on the Open Culture Data blog
* *[Where Next for Open Culture Data?](http://www.audiovisualthinking.org/videos/issue-7-open-data/%22%20%5Co%20%22Where%20Next%20for%20Open%20Culture%20Data?" \t "_blank)* by Lotte Belice Baltussen & Maarten Brinkerink on Audiovisual Thinking, the journal of academic videos.

Tags:

[Open Cultuur Data](https://www.beeldengeluid.nl/en/tags/open-cultuur-data)

[wikipedia](https://www.beeldengeluid.nl/en/tags/wikipedia)

[open content](https://www.beeldengeluid.nl/en/tags/open-content)

[metrics](https://www.beeldengeluid.nl/en/tags/metrics)

[cultural analytics](https://www.beeldengeluid.nl/en/tags/cultural-analytics)

[GLAM](https://www.beeldengeluid.nl/en/tags/glam)

[GLAMetrics](https://www.beeldengeluid.nl/en/tags/glametrics)