



≡ Menu

Museum Collections on Wikipedia: Opening Up to Open Data Initiatives

Elena Villaespesa, Pratt Institute, USA, [Trilce Navarrete](#), Erasmus University Rotterdam, The Netherlands

Abstract

The Web has become an important source of information, made possible by structured data. Open linked data enables ubiquitous presence as machines increasingly filter our views—via preferred search engines, the knowledge graph, or Siri—particularly of content found in Wikidata. In this paper, we identify paintings in Wikidata and analyze their usage in the English Wikipedia to find substantial impact. Our results provide evidence that publication of collections as open data facilitate an increase in views, enriched data, automatic translation, and magnified visibility. We find that the usage of paintings and views present a long-tail structure with an underrepresentation of contemporary paintings. Collaborations between museums and Wikimedia yield increased impact, yet projects are unsustainable. We propose an adjusted work-flow to accommodate for Wikimedia projects and amplify the impact of opening museum collections data.

Keywords: Wikipedia, Wikidata, paintings, linked open data, metrics, open access

Introduction: Open Data

In the early 1990s, the World Wide Web was presented as a new information communications technology which could allow universal access to content. In order to achieve this, documents needed to be published following a series of standards that enabled access through hyperlinks (Berners-Lee and Cailliau, 1992). Documents were indexed by search engines delivering relevant results accessible through Web browsers. As technologies and user needs have become more sophisticated, and we increasingly wish to query data semantically—also by our machines, data standards have further developed into open linked structures. It is not only documents that can be linked but also data elements. One such large dataset contains all information forming Wikipedia: DBpedia. The non-profit, user-based, free online encyclopedia makes use of the open software as well as open data to further distribute information, including content of museums, libraries, and archives.

As the Web became a ubiquitous source of information, a duality emerged: some firms made the web a profitable source of revenue and *closed* their content, requiring passwords or payments, while other firms continued to develop the open standards that enable access to structured content by people and machines. One leading open initiative is the World Wide Web Consortium (W3C), active in developing protocols and guidelines that support open standards. Its guiding principles recognize the value of sharing knowledge (as access to human knowledge leads to further knowledge generation) and the variety of access points (there are many devices which require flexible publication schemes) (W3C website: mission). Current legal frameworks are not always supportive of open data policies, and alternative business models are expected to imitate a revenue model based on property of objects, rather than on access to information. Open data, in fact, urges institutions to rethink their position in the information economy where strong network effects are present. The main benefit of joining the open data network is having access to its content while giving access within the linked structures that, in turn, enrich all participating members.

In order to use the Web as a single global database, some technical and legal solutions are still being sought out. However, an important challenge remains the participation of heritage institutions as important banks of vast, authoritative, structured, millenary information. From a 2013 survey on the perception of open data, Swiss institutions reported finding open data important and desirable (up to 81%), but they are rarely ready to make content available for “free” (7%) or for commercial use by others (1%) (Estermann, 2013).

In this paper, we will show the potential impact on museums who join the Wikipedia network, particularly when a long-term strategy is in place. For this, we will look at the current network of paintings from museums present in Wikipedia, and analyze the impact. We will then discuss the current paths and tools available for museums and propose an action plan to open up collections for museums to tap into the linked open data network.

Museum Collections Online

How are Internet users encountering museum collections online? Marty (2007 and 2008) found that accessing collections on museum websites is short-lived, while Europeana is mostly used by heritage staff (Clough, et al., 2017). On the other hand, Google and Wikipedia are highly used environments which easily give access to collections, independently of where these may be located. There are different ways Internet users may come across museums’ content and resources online beyond their website. Starting from a search, which is the most common method of finding information online, here are a few scenarios using the keyword “mermaid” in a Google search, Google Images search, and with Siri.

A search on Google will often generate a knowledge panel with an image and a few lines taken directly from a Wikipedia article, with a link to read more. Figure 1 shows the knowledge panel for “mermaid” shows one painting from the Royal Academy of the Arts and one illustration from the Smithsonian. If one were to follow the hyperlink to Wikipedia, the infobox on the Wikipedia article for “mermaid” shows the painting by John William Waterhouse from the Royal Academy of the Arts (UK) (see Figure 2). A search on Google Images can return images from the museum that have been uploaded to Wikimedia Commons. Figure 3 is a screenshot of Google image results, filtering by paintings and images labeled for non-commercial reuse. It includes several paintings from various museums. And a search using a voice assistant, such as Siri or Alexa, can return a response using the museum’s data on Wikipedia and Wikidata. Figure 4 shows the virtual assistants Siri’s response to “mermaid.”

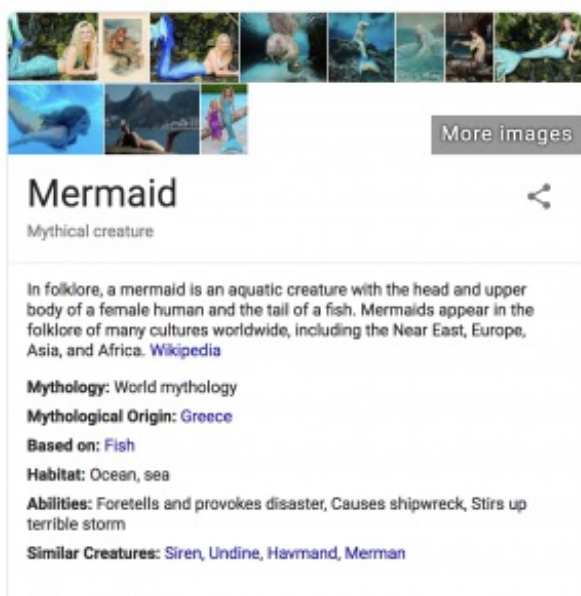


Figure 1: Google knowledge panel

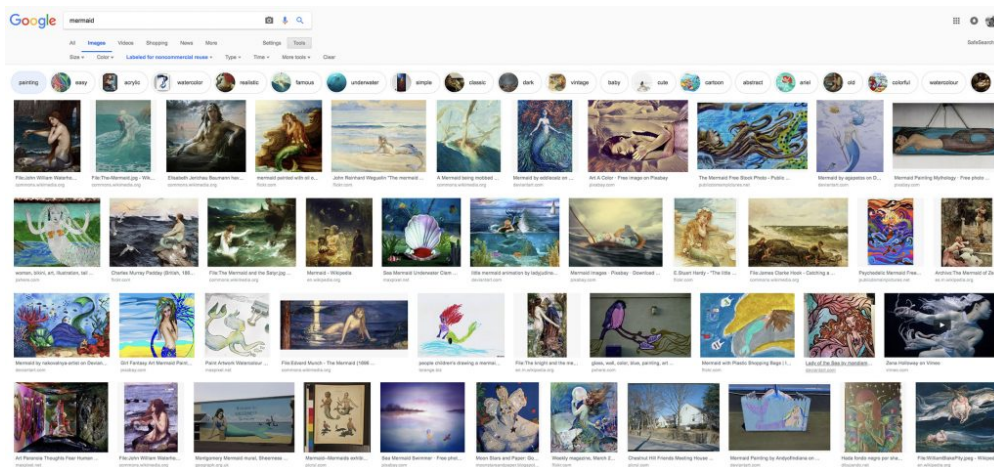
Mermaid



John William Waterhouse, A Mermaid (1900).

Grouping	Mythological
Sub grouping	Water spirit
Similar creatures	Merman Siren Undine
Mythology	World mythology
Country	Worldwide
Habitat	Ocean, sea

Figure 2: Wikipedia article



Google mermaid

Images Videos Shopping News Maps

Size Color Labelled for accessibility More Type More tools Clear

Settings

Mermaid images from various artists and styles, including classical paintings, modern illustrations, and digital art.

Figure 3: Google images

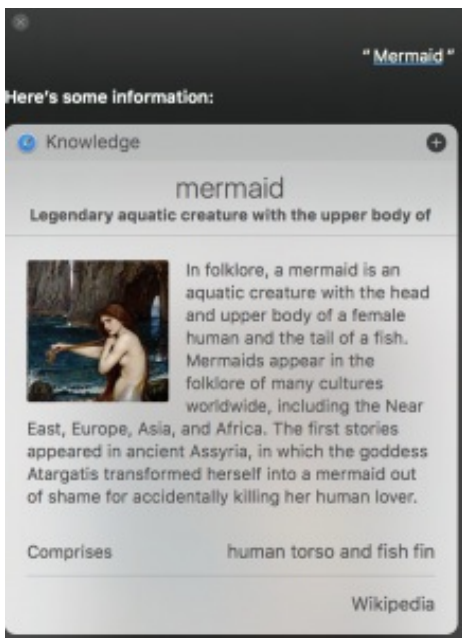


Figure 4: Siri

It is clear that with our current information-seeking behavior and search engine algorithms, being part of the global online structured networks of data, can position museum content in a prevalent position as quality trusted data. Finding a museum painting on one click represents being high in the hierarchy of online content, thanks to the structured environment provided by the Wiki-platforms strongly influencing the quality of the search results.

Methodology

Data for this study was gathered from three Wikimedia projects: Wikidata, Wikimedia Commons, and Wikipedia in English. All three projects have a public SPARQL or API (Application Programming Interface) endpoint. The process of data collection and analysis was conducted as follows:

1. List of paintings images used on Wikipedia articles. We started by identifying the “paintings” available in Wikidata with a query using the SPARQL language (Figure 5), resulting in 224,374 items. Then, we gathered the data of those paintings containing basic metadata of author, date of creation, and location that have an image representation and that are used in a Wikipedia page in the English edition, resulting in 10,054 items.

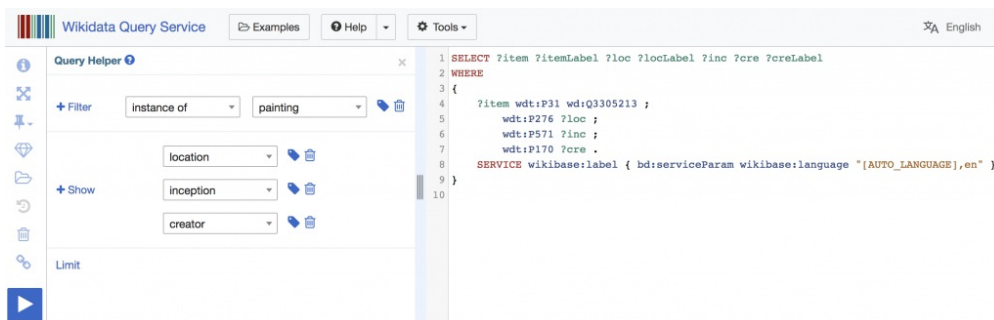


Figure 5: SPARQL query

2. Article coding. We manually assigned a code following a modified version of the topic category ontology proposed by Spoerri (2007): Entertainment, Politics and History, Geography, Sexuality, Science, Art, Religion and mythology,

Literature, Performing Arts, and Drugs. We added a category of Wikipedia to exclude pages that are not properly an encyclopedia article such as “did you know” and “features” of articles or images, file names, templates, and lists (e.g. paintings, years, recent additions). This left a dataset of 8,104 paintings (3% of paintings in Wikidata) that were used in 10,008 Wikipedia articles.

3. **Article views.** Using a Python script that connected to the Wikipedia page views API, we harvested the total page views in all identified Wikipedia articles for 2017.

4. **Location and country data clean-up.** The location data is, in the majority of cases, museums and galleries, but it also includes private collections, monuments, embassies, and historic places. The data was cleaned as there were duplications due to misspellings or similar instances of the location (e.g. Wallace/Wallace Collection), due to the inclusion of the gallery’s names instead of the whole museum (e.g. location of paintings from the Louvre Museum were listed by gallery name so they were grouped all together), in other cases there are several instances for the same museum collection (eg. Tate was divided by gallery: Tate Britain, Tate Modern, Tate Liverpool). In some cases, the collection could not be classified due to lack of information (e.g. “storage space” or “private collection”), so this cell was left black. Finally, the country was added to each collection.



Figure 6: Number of paintings data in each of the data collection and coding steps

Results and Findings

Access and Reach of Painting Images on Wikipedia

The dataset for this research includes 8,104 paintings that are used in 10,008 articles in the English Wikipedia. Using the site statistics, we can calculate that .17% of all the English Wikipedia articles include a painting image. The reach of these articles is substantial; Wikipedia articles including a painting received a total of 94.2M monthly views on average during 2017.

Item	Number
Articles	10,008
Paintings	8,104
Article views (avg. monthly in 2017)	94,217,895
Museums/Collections	785
Countries	59

Table 1: Summary of the data

The distribution of the number of views of those articles follows a long-tail shape. About 2% of those articles receive more than 100K monthly views, which means 38% of the total views (see histogram of the article views in Figure 7).

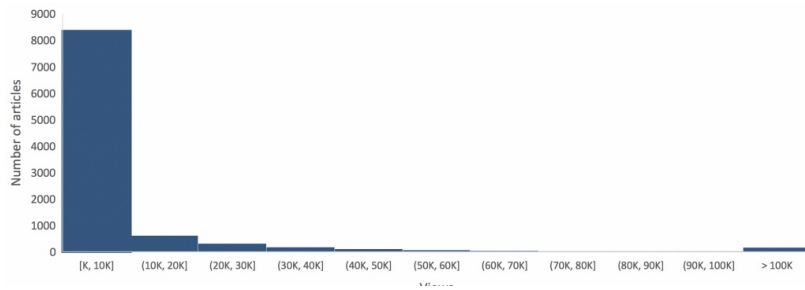


Figure 7: Histogram of the article views

The analysis of all article categories shows that approximately 33% of the articles where a painting image has been added are related to art, including pages about artists, artworks, art movements, art techniques, or art museums. The rest of the articles (67%) are not directly related to art but instead cover typical encyclopedia articles on topics of history, religion, and geography. In terms of views, art-related articles represent 12% of the total views, while 88% of the views of articles including a painting happen in non-art related articles (tables 2 and 3 show the top 10 articles in each of the categories).

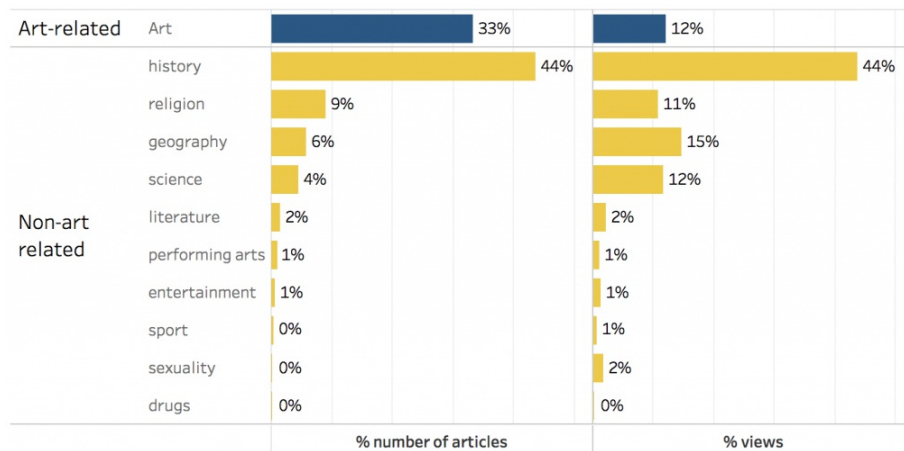


Figure 8: Percentage of articles and views by category

Article	Category	Views
Queen Victoria	history	1,189,122
Charles Darwin	science	859,127
Canada	geography	758,040
Mary, Queen of Scots	history	639,455
Abraham Lincoln	history	568,828
Russia	geography	555,588
France	geography	518,598

Alexander Hamilton	history	499,887
Henry VIII of England	history	490,664
American Civil War	history	481,034

Table 2: Top 10 non-art related articles based on monthly views (2017)

Article	Views
Leonardo da Vinci	359,195
Vincent van Gogh	272,892
Mona Lisa	265,626
Michelangelo	162,605
Romanticism	135,136
Art	105,672
The Starry Night	97,641
David (Michelangelo)	93,108
Claude Monet	90,716
Louvre	90,678

Table 3: Top 10 art related articles based on monthly views (2017)

There is a significant finding coming out of this analysis: painting images are used to illustrate a range of topics in Wikipedia including, for example, historical facts, political figures, cities, films, scientists, musical instruments, mythology, etc. (See some examples in figure 9.)

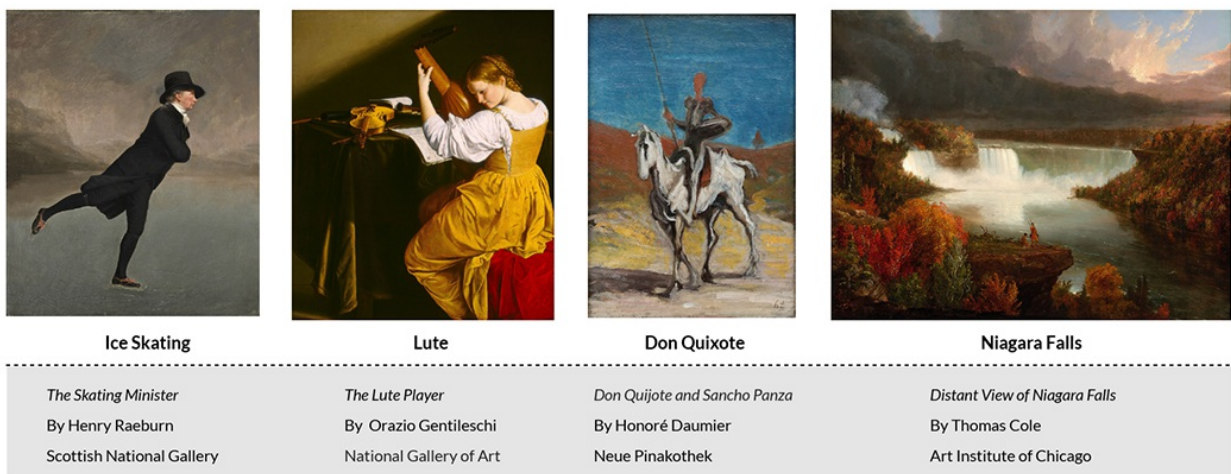


Figure 9: Examples of painting images used on Wikipedia articles

The same painting can be used in various articles to illustrate multiple topics belonging to different categories. For example, the image of *The Starry Night* by Van Gogh appears on 16 articles, some are related to the artist, the artwork, or are included in generic art articles such as “art movement,” “history of painting,” or “night in paintings.” This iconic painting also constitutes a visual element in the following articles: psychosis, moon, moonlight, Venus, Saint-Rémy-de-Provence, and culture of the Netherlands. The image caption highlights the relevant information for each article (Figure 10).

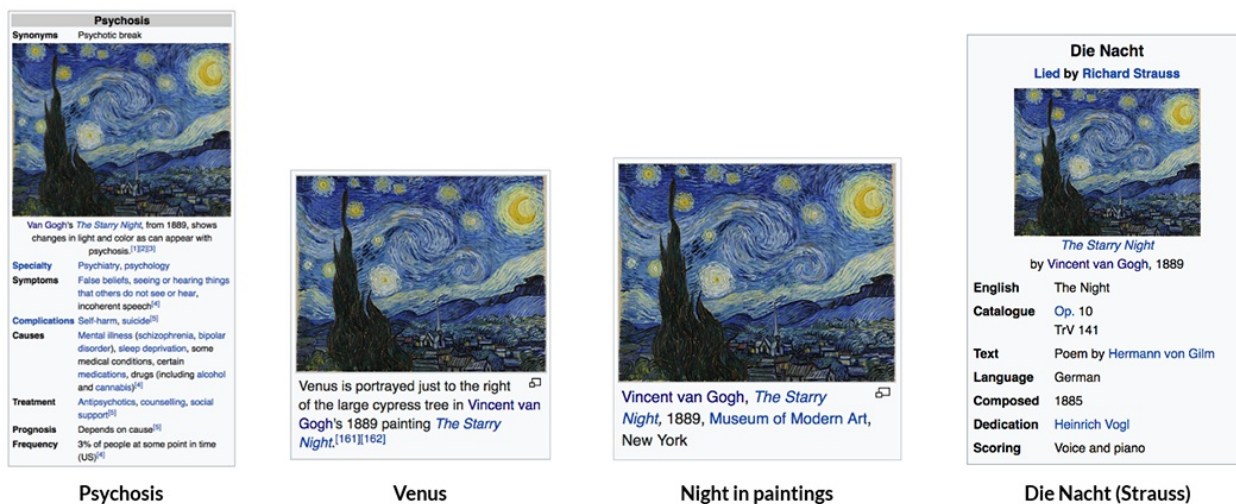


Figure 10: *The Starry Night* image appears on the infobox of the article about psychosis and as an illustration of several other articles.

The key effect of the inclusion of painting images on articles that are not directly related to art is that museums are getting visibility in seemingly unrelated contexts by people searching for information, reaching new audiences.

Unequal Representation of Paintings from Countries and Museums in the World

The map in Figure 11 presents the volume of views of paintings from museums used in Wikipedia articles by country, which denotes a clear presence of paintings from museums in Western countries. (See Table 5 for the top ten countries of the dataset.) This analysis may return slightly different results when looking at other Wikipedia languages. Interestingly, countries that have organized Wikipedia-in-residence programs get the majority of views (e.g. U.K., U.S., Spain, Netherlands). Other countries with large population sizes (e.g. Russia) or long histories of working with open data (e.g. France) also receive significant views in the English Wikipedia.

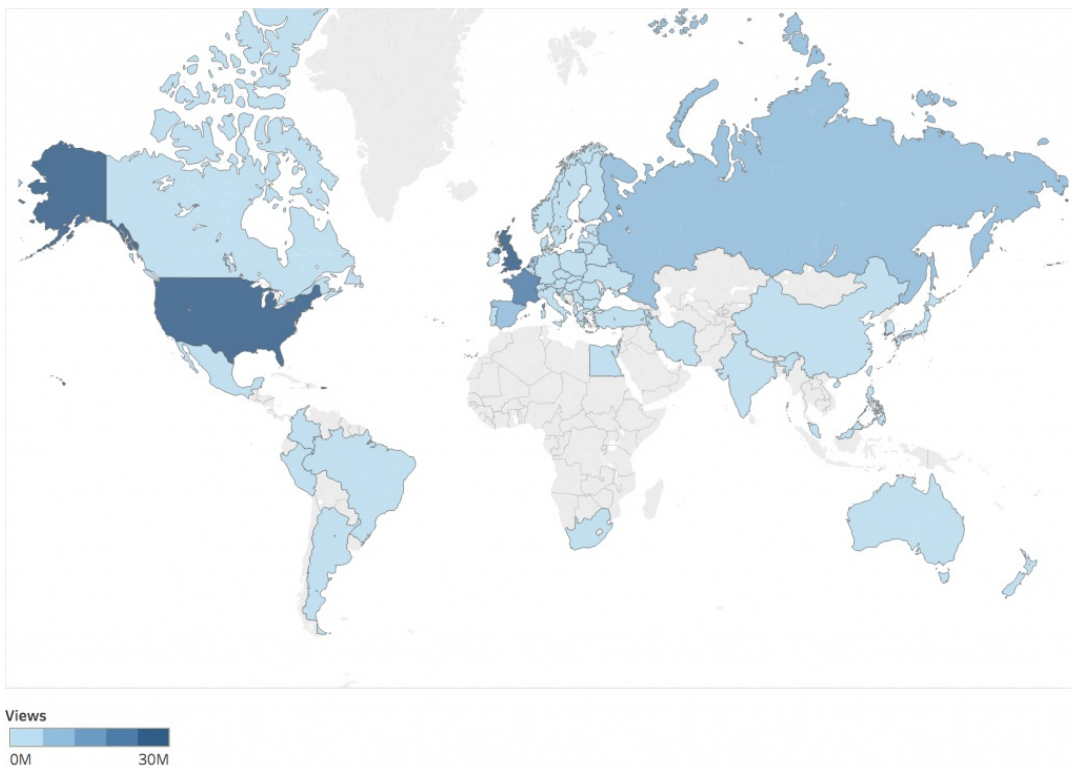


Figure 11: Map of views by country

Country	Views	Number of articles	Number of paintings
United Kingdom	27,866K	2,582	1,307
United States	24,559K	2,229	1,611
France	18,255K	1,599	897
Spain	10,097K	726	433
Netherlands	7,786K	1,026	1,001
Russia	6,528K	520	276
Germany	5,460K	645	412
Italy	4,504K	499	282
Poland	2,960K	244	133
Austria	2,585K	323	303

Table 5: Top 10 countries ordered by volume of article views (monthly)

This leads to the next breakdown of the data—the analysis by museum collection. Larger well-known museums are at the top of the list (Table 6). Some museums listed have open-access policies such as The Metropolitan Museum of Art or the Rijksmuseum, but that is not the case for all the institutions represented on the list. These numbers only scratch the surface of the representation of painting images by country and museum collection. It is very likely affected by the country language and the popularity of some of the artworks in these museums. Further research points to a more comparative and detailed analysis by language and individual collections.

Museum	Views	Articles	Paintings
--------	-------	----------	-----------

Louvre Museum	10,846,301	933	495
National Portrait Gallery (UK)	8,184,376	766	373
National Gallery	8,048,944	728	340
Museo del Prado	7,182,202	481	238
Metropolitan Museum of Art	6,316,237	517	296
National Gallery of Art	4,342,791	411	224
Musée d'Orsay	4,275,493	303	203
Hermitage Museum	3,855,089	327	199
Rijksmuseum	3,614,756	522	437
Tate	3,481,855	336	163

Table 6: Top 10 museums ordered by volume of article views (monthly)

Impact on Copyright Limitations and Open Licenses

A striking observation was revealed when analyzing the presence of an image of Wikidata “painting” items. Legal protection of images has a direct impact on the availability and use of images to illustrate Wikipedia articles as well as any other global use, including the knowledge panel or Siri searches. We identified all the “paintings” and arranged them by year of creation, noting a particular increase in items belonging to the Renaissance, Modern, and Contemporary periods, as well as a surprising absence of images for items made after the 1930s—even though an item page identifies such an item. Copyright restrictions vary by country but generally protect art objects 70 years after the death of the artist, explaining the lack of images of contemporary art—which has been referred to as the copyright hole (Boyle, 2009). The peak at 1650 may reflect collections dating from the “second half of the 17th Century,” being dated as 1650.

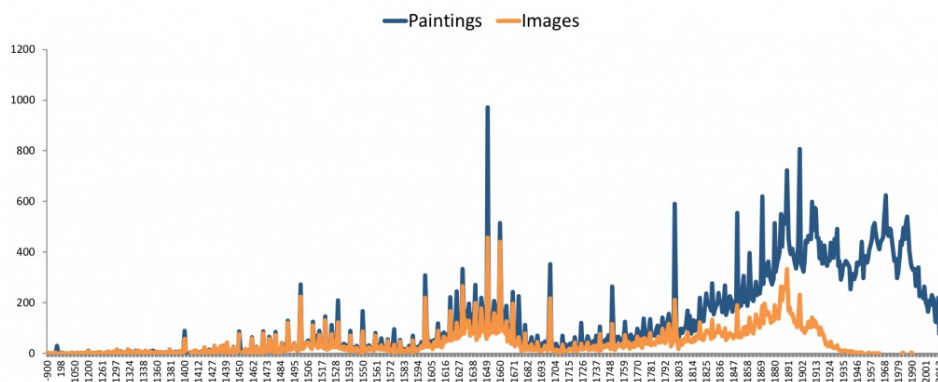


Figure 12: Wikidata “painting” pages and images by year

Tools Available to Measure the Impact of Open Access

This research aims to capture the usage of painting images on Wikipedia and, therefore, one of the first steps in the process was to review the analytics tools to gather the data. There are different tools available to measure the impact of adding all these images and data to the different Wiki projects. The most used in the sector have been Glamorous, Treeviews and BaGlama developed by Magnus Manske. Glamorous (<https://tools.wmflabs.org/glamtools/glamorous.php>) shows the number of files used in each of the Wikimedia projects at the time of running the query for a specific category. The details section provides a list of the top thousand images, each listing the pages where the image has been used. Similarly, Treeviews (<https://tools.wmflabs.org/glamtools/treeviews/>) shows monthly page views for Wikipedia category trees. The

BaGlama2 tool (<https://tools.wmflabs.org/glamtools/baglamma2/>) shows the number of article views, with breakdown by Wikipedia language and by article, per “category,” which is the name of the institution or project. “Categories” have to be added by the developer in order to be tracked. The first institutions were added in 2010 and currently include 80 categories that track museums (Figure 13). The advantage of using these tools is that they provide harmonized information, useful for benchmarking with other cultural organizations and to demonstrate the impact of joining Wikimedia projects. The drawbacks include the dependency on categorization so that images uploaded need to be clearly identified and categorized in order to be tracked and the simplified article “views,” ignoring the positioning of the image within the article or the interactions with the image, such as number of downloads or clicks to see the image details.

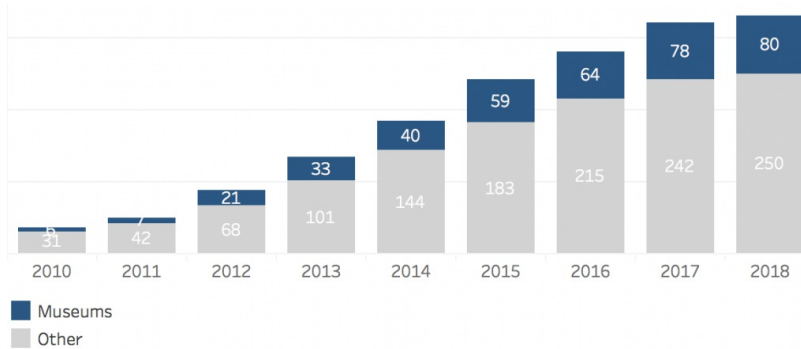


Figure 13: Number of GLAM on BaGLAMA tool (cumulative since year added)

While these tools created by volunteers are extremely valuable, there is a need for robust tools that integrate results from all Wikimedia projects in an automated and easy-to-use form. The GLAM sector could benefit from having a dashboard, or reporting interface, where to integrate all metrics. There have been a few attempts to capture requirements and develop tools to improve the reporting of key metrics. Wikimedia Switzerland developed a dashboard for their GLAM community to display the category network, user contribution, usage of files, and number of views (GLAM stat tool). The goal is to display the usage of images in any Wikimedia project and search engines.

Museums and Wikimedia: Past, Present, and Future

Since the early 2010s, collaborations between the GLAM sector and the Wikimedia Foundation have been explored. (For a detailed review of the GLAM-Wiki collaborations, see Stinson et al. 2018 and Lih, 2018.) We briefly review four types of collaborations that influence the data results presented in this paper: 1.) Wikipedians-in-residence, 2.) edit-a-thons, 3.) batch uploads to Wikimedia Commons 4.) data entry on Wikidata.

Wikipedians-in-Residence

One frequent collaboration in the museum sector is the establishment of a Wikipedian-in-residence program. This program, in a variety of contract formats (voluntary or paid, temporary or permanent, part-time or full-time), brings a Wikipedian editor to the museum to coordinate and promote initiatives to improve the presence of museum’s resources on Wikipedia. Some of the activities undertaken by Wikipedians-in-residence can include training sessions or edit-a-thons.

In 2010, the first Wikipedian-in-residence program took place at the British Museum, where during five weeks, new articles about objects were added by Wikipedia editors in collaboration with curators. A challenge to the Hoxne Hoard page was proposed bringing the article from the stub category to a featured article, and Wikireaders were given this page to be used in the school program (Wyatt, 2010). The same year, another Wikipedian in residence started at the Children’s Museum of Indianapolis. During this period, new articles were created in edit-a-thons; pages got translated

into other languages, and a QR code project was implemented in the galleries allowing visitors to find more information on Wikipedia about a set of objects on display. About 200 object images were uploaded to Wikimedia Commons to be used by editors on Wikipedia (Byrd Phillips, 2011).

At the time of writing this paper, the articles where these images have been added have received 128M views. Another museum that hosted a Wikipedian-in-residence was the Museu Picasso in Barcelona. The goals of this residency were to create basic content for some highlights of the collection and improve the article in Catalan about the series of paintings, *Las Meninas*, which became a “quality article.” For the museum, their expertise combined with the reach of Wikipedia will provide, as a result, a better service for the user (Rodà, 2011). Other museums that have hosted a Wikipedian-in-residence in this first wave of Wikipedian- in-residence include Derby Museum and Art Gallery, Museum of Modern Art, and Israel Museum. Initial published results of some of these collaborations in 2010-11 included over 194K images uploaded, 76 events, and 2,000 articles improved in over 50 languages (GLAM-Wiki US, 2012). Other museums and cultural organizations followed the lead and hosted a Wikipedian-in-residence, reaching 165 residences, of which 44 took place in museums (Wikimedia Outreach: Wikipedian in Residence).

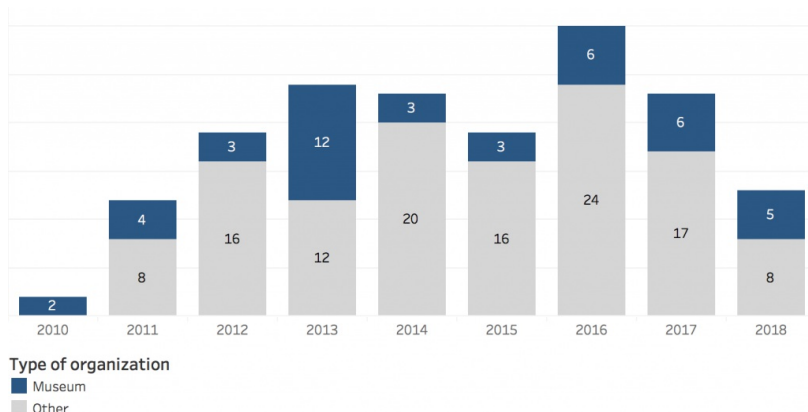


Figure 14: Number of Wikipedians-in-residence represented by starting date (as of December 2018)

Country	Number (museum numbers in brackets)
United Kingdom	32 (9)
United States	25 (5)
Spain	14 (8)
Serbia	13 (6)
Italy	13 (4)
Netherlands	7 (1)
Macedonia	5
Australia	5 (1)
Sweden	4 (1)
Germany	4 (1)
Brazil	4 (1)
Tunisia	3
Switzerland	3
Mexico	3 (2)
Israel	3 (1)
France	3
Austria	3
South Africa	2
Russia	2
India	2
Denmark	2 (1)
Chile	2
Poland	1 (1)
Philippines	1
Nigeria	1
New Zealand	1 (1)
Finland	1
Egypt	1
Czech Republic	1

Table 7: Number of Wikipedians in residence by country. (Museum figures in brackets.)

Edit-a-thons

Edit-a-thons are organized events hosted at a museum location, online, or a combination of both, where Wikipedia editors, either beginners or experts, are invited to create or add information to articles based on the collections, generally about a specific topic, artist, event, or anniversary. Success from this type of meetup has been reported in terms of reach, diversity of content added and collaboration with local communities. For example, during an edit-a-thon organized by Europeana about WWI ended up with 20 images added to 62 articles, which got almost 2 million views in 6 months. (Europeana, 2013). Museums around the globe have gotten involved in the *Art+Feminist* initiative which aims at “improving coverage of cis and transgender women, non-binary folks, feminism, and the arts on Wikipedia” (<http://www.artandfeminism.org>). The year 2018 closed with the participation of 3,816 editors that worked on 22K articles in total. These new or edited articles have received 79M views (*Art+Feminist* dashboard, 2018). Edit-a-thons have become popular in the sector, some breaking records in attendance and engagement duration, but some argue such community involvement could become part of the museum’s regular events organized internally (Snyder, 2018).

Bulk Uploads of Images to Wikimedia Commons

Wikimedia Commons is a wiki platform that hosts over 50 million media items under open licenses. Wikipedia editors use this media repository to add content to the articles. Therefore, having object images on Wikimedia Commons may result in the usage of these on Wikipedia articles. In order to facilitate the batch upload process of images, the “GLAMwiki Toolset” was built based on a collaboration between Wikimedia in several countries and Europeana. Besides its limitations, data mapping, and technical requirements, the tool has been used by various museums, including the Rijksmuseum, the Nordiska Museet, and The Metropolitan Museum of Art (Knipel, 2017). The reach of these images on Wikipedia can be tracked with the tool BaGLAMa2, and it has been used internally to provide evidence of the impact of this work. The tool estimates the number of images from the GLAM-Wiki partnerships uploaded to Wikimedia Commons is 2.5M images, which in fact is probably higher due to the limitations of capturing all the images and institutions currently participating (Stinson, 2018).

Data Imports to Wikidata

Wikidata is a repository of currently 54 million data items and growing that serves to structure data for other Wikimedia projects. Wikidata is one of the focuses of the new strategic direction of the Wikimedia Foundation and currently the fastest growing project (Wikimedia Strategy, 2017). This site, which is edited by both humans and bots, received 568.62M page views in the past 12 months (Wikimedia Statistics, 2018). The clear advantages of adding museum data to Wikidata include tapping into the active community that will continue enhancing the records, automated translations into multiple languages, integration with other Wikimedia projects, linkage to other data records, and access to the search query and visualization tools available. Data entry can be done manually or automatically; the latter is only possible if the metadata is under a CCO license. Museums are already collaborating to publish data on this repository. For example, the *Sum of all paintings* project (https://www.wikidata.org/wiki/Wikidata:WikiProject_sum_of_all_paintings) focuses on creating items for each collection painting and adding all the relevant metadata. Tutorials, case studies, and best practices to support the sector are available at Wikiproject Culture Heritage and Wikiproject Museums.

Future in the GLAM-Wiki

Collaborations between museums and the Wikimedia projects will increasingly focus on structured and linked open data. There is a major focus on Wikimedia's agenda to work with the sector to use Wikidata in order to build a central repository of heritage data (Stinson, Fauconnier, and Wyatt, 2018). Moreover, the *Structure data on Wikimedia Commons* project (2017-19) is taking place "to convert the free media files on Wikimedia Commons to a structured and machine-readable format, so that they become easier to view, search, edit, organize and re-use." This initiative will provide specific features for the GLAM sector, including structured, machine-readable and multilingual metadata (linked open data), structured copyright and attribution information, and rich APIs (Wikimedia Commons: Structured data|GLAM).

Opening up museum collection data has encountered internal resistance in museums as a result of fear about the loss of authority, quality, and control of the images and information, as well as arguments about the reduced income that comes from rights and reproduction permissions activities. It also brings some important challenges with regards to the technology and logistics needed to support the infrastructure and the processes for cleaning the collection data. Besides these risks and challenges, there are strong arguments made to open up museum collections, including a potential loss of funding opportunities, the museum's relevance and branding, and most importantly, how this open practice could contribute to the achievement of the museum's mission (Kapsalis, 2016, Kelly, 2013). Moreover, restricting the distribution and usage of images is practically impossible due to people's expectations and behaviors online (Sanderhoff, 2013).

Therefore, different museum and wiki professionals have been encouraging the museum sector to embrace the collaborative element of the Internet to provide access to their collections and increase the collective knowledge (Byrd Phillips, 2013; Fouseki & Vacharopoulou, 2013). This strong advocacy to open up has been evidenced in several museums, including the Rijksmuseum, SMK museum, Te Papa, and The Met, who have published results in the year after their launch of these policies. Open access brought a positive impact on the brand reputation of the museum, and an increase in funding and sponsor opportunities, which for these museums outweigh the loss of revenue from rights and reproduction work. In terms of access, the number of image downloads increased, and so did the views of these images on Wikipedia, including views in various languages. (Kingston & Edgar, 2015; Schmidt, 2017; Pekel, 2014; Tallon, 2018; Villaespesa, 2018; Navarrete and Borowiecki, 2016). This paper provides further insight into the reach in volume and diversity of audiences that come as a result of uploading painting images to the wiki tools.

An Action Plan to Open Up

The results of this study highlights important implications for the museum sector to open up their collections. From a strategic perspective, structuring and opening data should be of paramount importance if museums want to be findable and relevant in the future. The Wikimedia projects allow the positioning of museum data in the global network of information available to Google, Siri, and any other potential user. Automated data linkage, data translation, and data enrichment is possible once the object has been entered in Wikidata, which can be useful for museums interested in expanding their data. Examples of museums "taking data back" from the Wikimedia projects include Tate and MoMA who present Wikipedia artists' biographies on the museum's website, and in the case of MoMA, also the Wikidata ID (Foe, 2016). There is also a range of third-party tools created by the community that museums could customize to browse and dig into their collections, integrating those on the museum platforms (online and onsite), for example, the Collection Explorer (<http://glam-discovery.bodleian.ox.ac.uk/>) that pulls art data from Wikidata.

The key to enabling all this is integrating Wikimedia projects into the organizational workflow. In this way, museums adopt a linked open data strategy within a linked open data structure that immediately gives visible returns. Figure 15 visualizes an open-knowledge creation and distribution process in which museums audit their data and practices in each of the arrows of the workflow in order to define a strategic plan towards open access. This diagram is a very simplified visual of the practical implications in each of the data pipelines, but it could serve as a guidance tool to review the current status and plan for a better final discovery and usage of the museum's collection.

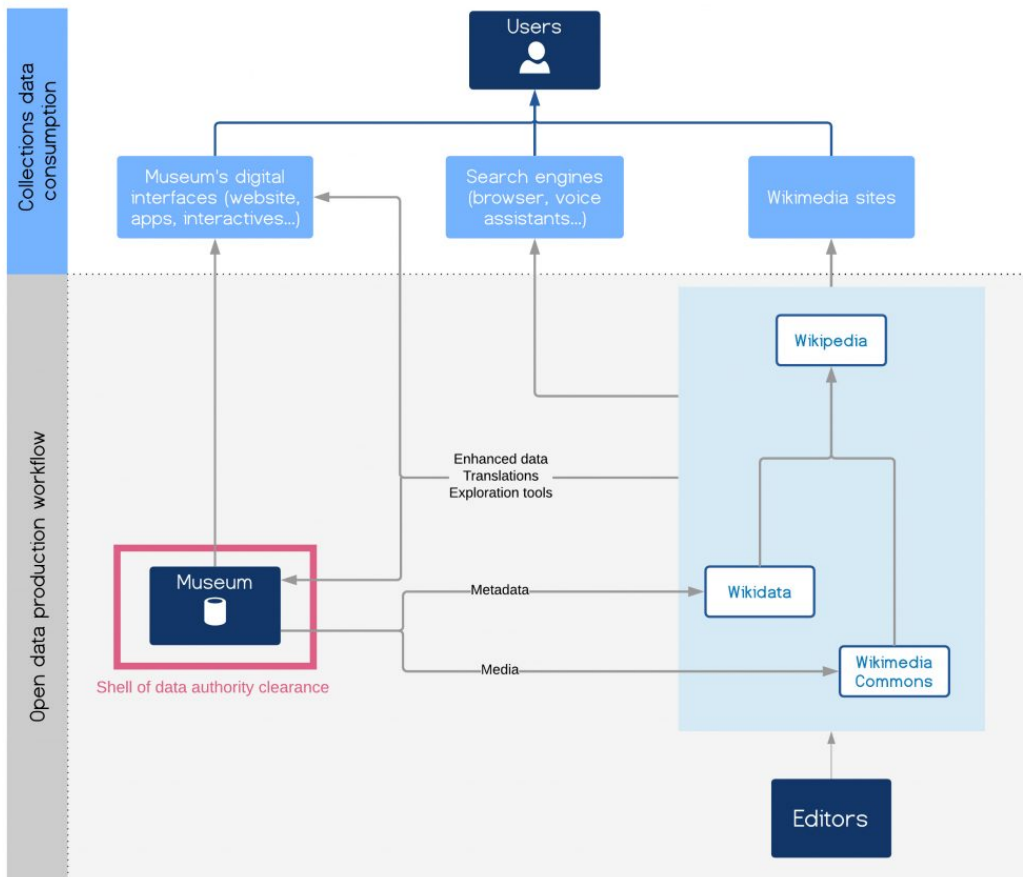


Figure 15: Open knowledge production and delivery workflow

Collaborations reviewed on this paper are normally a short-term add-on activity that, in the majority of cases, is promoted by the digital or technology department. However, this approach is not sustainable due to the speed of change of the online landscape and daily museum activities, such as new acquisitions, new photos with better quality, and other changes happening on an ongoing basis in the collection information management system that do not get reflected in the Wikimedia tools. Therefore, from a tactical perspective, this collaborative work needs to be embedded in museum processes going beyond edit-a-thons or Wikipedians-in-residence schemes. For example, when a collection object record is added or edited, there should be a step to update this data on the wiki tools, and whenever possible, this needs to be automated. This brings us to the need for better tools and dissemination of skills that are required to support this work for museums of all types and sizes. In the next step of the work chain, when it comes to measuring the impact of this work, there is a clear need of an evaluation framework across the sector and user-friendly, robust tools to track and report activities across all Wikipedia projects.

Conclusion

The Web is our global library of information that depends on open linked data to improve query results. Museums are a vast source of authoritative, quality, structured data that can enrich and be enriched by joining the Wikimedia projects information network. This research looked at the use of painting images to illustrate English Wikipedia articles and showed the substantial reach, particularly regarding the views to non-art related articles. Besides page views, the analysis of article categories where images have been used shows the value of museum collections as information sources. Our analysis highlighted the need for a revised legal framework to support the dissemination of quality information as a substantial portion of contemporary art is absent from the Web. So far, paintings located in Western museums are overrepresented, as well as well-known paintings, painters, and museum collections. Most importantly, our results show that museums that tap into the Wikimedia projects, and Wikidata in particular, also tap into the network of open linked data used by people (e.g. Wikipedia readers) and machines (e.g. Google and Siri).

The list of museums collaborating with Wikipedia in various forms continues to increase, but limited evidence is available on the overall impact of these efforts. The research presented in this paper sheds light on the impact of partnering with Wikimedia projects, including increased views, increased data links, data translation into multiple languages, magnified findability, and a great potential still waiting to be measured. Insights gained from this research may be of assistance to museums considering the adoption of a linked open data strategy, including an adjusted workflow to integrate the Wikimedia projects and, subsequently, capture the impact of those platforms.

This research is limited in scope, as we only captured “paintings” labeled in Wikidata used in the English Wikipedia. We, nevertheless, feel confident our results will serve as a stepping stone for future research involving other language editions of the encyclopedia and a more encompassing distribution of collections beyond “paintings.” Another study of interest could undertake the user journey of readers when reading articles, including museum collections and the relation to search engine results on a browser or while using voice assistants.

References

- Art+Feminism Dashboard (2018). Consulted December 12, 2018. Available at: https://outreachdashboard.wmflabs.org/campaigns/artfeminism_2018/programs
- Berners Lee, T. and Cailliau R. (1992). World-Wide Web. Geneva: CERN. Available at: <https://cds.cern.ch/record/245440/files/p69.pdf>.
- Boyle, J. (2009). “A copyright black hole swallows our culture.” *Financial Times*. 7 September 2009, section Comment, pp. 7. Consulted January 3, 2019. Available at: <https://www.ft.com/content/6811a9d4-9b0f-11de-a3a1-00144feabdc0>
- Byrd Phillips, L. (2011) “What’s that Wikipedian-in-Residence been up to?” *Children’s Museum Indianapolis’ Blog*. Consulted January 12, 2019. Available at: <https://www.childrensmuseum.org/blog/whats-that-wikipedian-in-residence-been-up-to>
- Clough, P., Hill, T, Lestari Paramita, M., and Goodale, P. (2017). “Europeana: What Users Search For and Why.” *Research and Advanced Technology for Digital Libraries. Theory and Practice of Digital Libraries (TPDL 2017)*, 18-09-2017 – 21-09-2017, Thessaloniki, Greece. Lecture Notes in Computer Science. Springer Cham, 207-219. Available at: <http://paramita.staff.shef.ac.uk/papers/cloughetal-TPDL2017.pdf>
- English Wikipedia Siteview Analysis. Consulted December 15, 2018. Available at: <https://tools.wmflabs.org/siteviews/?platform=all-access&source=pageviews&agent=user&start=2017-01&end=2017-12&sites=en.wikipedia.org>
- Estermann, B. (2013). “Swiss heritage Institutions in the Internet Era.” Results of a pilot survey on open data and crowdsourcing. Bern University of Applied Sciences.
- Europeana (2013). “Case study: Wikipedia edit-a-thons.” Consulted December 1, 2018. Available at: <https://pro.europeana.eu/post/case-study-wikipedia-edit-a-thons>
- Fouseki, K. and Vacharopoulou, K. (2013). “Digital Museum Collections and Social Media: Ethical Considerations of Ownership and Use.” *Journal of Conservation and Museum Studies*, 11(1), 5. doi:10.5334/jcms.1021209.
- GLAM Stat tool (Cassandra) version 1.2. Consulted December 17, 2018. Available at: <http://stats.wikimedia.swiss/>
- GLAM-Wiki US (2012). “Museums Collaborating with Wikipedia.” Available

at: https://upload.wikimedia.org/wikipedia/commons/8/88/GLAM_One-Pager.pdf

Kapsalis, E. (2016). "The Impact of Open Access on Galleries, Libraries, Museums and Archives." Smithsonian Emerging Leaders Development Program. Available

at: http://siarchives.si.edu/sites/default/files/pdfs/2016_03_10_OpenCollections_Public.pdf

Kelly, K. (2013). "Images of Works of Art in Museum Collections: The Experience of Open Access. A Study of 11 Museums." Council on Library and Information Resources. Available at: <https://www.clir.org/wp-content/uploads/sites/6/pub157.pdf>

Kingston, A. and Edgar, P. (2015). "A review of a year of open access images at Te Papa." *MWA2015: Museums and the Web Asia 2015*. Consulted January 14, 2019. Available at: <https://mwa2015.museumsandtheweb.com/paper/a-review-of-a-year-of-open-access-images-at-te-papa/>

Knipel, R. (2017). "The Metropolitan Museum of Art: 375,000 windows on art history, and that's just the beginning." *Wikimedia Blog*. Consulted December 22, 2018. Available at: <https://blog.wikimedia.org/2017/07/25/met-open-access-initiative/>

Lih, A. (2018). "What Are Galleries, Libraries, Archives, and Museums (GLAM) to the Wikimedia Community?." Proffitt, M. (ed.) *Leveraging Wikipedia: connecting communities of knowledge*. ALA Editions, pp. 7–16.

Marty, P. F. (2008). "Museum websites and museum visitors: digital museum resources and their use." *Museum Management and Curatorship*, 23:1, 81-99. Available at: [10.1080/09647770701865410](https://doi.org/10.1080/09647770701865410)

Marty, P. F. and Burton Jones, K. (2008). *Museum Informatics. People, Information and Technology in Museums*. New York: Routledge.

Marty, P. F. (2007). "Museum Websites and Museum Visitors: Before and After the Museum Visit." *Museum Management and Curatorship*, 22:4, 337-360. Available at: [10.1080/09647770701757708](https://doi.org/10.1080/09647770701757708)

Navarrete, T. and Borowiecki, K. J. (2016). "Changes in cultural consumption: ethnographic collections in Wikipedia." *Cultural Trends*, Vol.25(4), p.233-248.

Pekel, J. (2014). "Democratising the Rijksmuseum." *Europeana*. Available at: https://pro.europeana.eu/files/Europeana_Professional/Publications/Democratising%20the%20Rijksmuseum.pdf

Phillips, L. B. (2013). "The Temple and the Bazaar: Wikipedia as a Platform for Open Authority in Museums." *Curator: The Museum Journal*, 56(2), 219–235. Available at: doi: [10.1111/cura.12021](https://doi.org/10.1111/cura.12021).

Rodà, C. (2011). "What is a Wikipedian-in-residence doing in the Museu Picasso?." *Museum Picasso Blog*. Consulted January 12, 2019. Available at: <http://www.blogmuseupicassobcn.org/2011/07/what-is-a-wikipedian-in-residence-doing-in-the-museu-picasso/?lang=en>

Romeo, F. (2016). "Bringing [art] knowledge to everyone who seeks it." *Digital @ MoMA–Medium*. Consulted January 4, 2019. Available at: <https://medium.com/digital-moma/bringing-art-knowledge-to-everyone-who-seeks-it-899ec257a55c>

Sanderhoff, M. (2013). "Open Images. Risk or opportunity for art collections in the digital age?" *Nordisk Museologi*, 0(2), 131. Available at: doi: [10.5617/NM.3083](https://doi.org/10.5617/NM.3083).

Schmidt, A. (2018). "MKG Collection Online: The potential of open museum collections." *Hamburger Journal Für Kulturanthropologie (HJK)*, (7), 25-39. Available at <https://journals.sub.uni-hamburg.de/hjk/article/view/1191>

Snider, A. (2018). "Edit-a-thons and beyond." Proffitt, M. (ed.) *Leveraging Wikipedia: connecting communities of knowledge*. ALA Editions, 105-118.

Stinson, A. (2018). "Let's sum all GLAM results." GLAM TLV Conference 2018. Consulted December 27, 2018. Available at: https://meta.wikimedia.org/wiki/GLAMTLV2018/Submissions/Let%27s_sum_all_GLAM_results

Stinson, A. D., Fauconnier, S., and Wyatt, L. (2018). "Stepping Beyond Libraries: The Changing Orientation in Global GLAM-Wiki". *Italian Journal of Library, Archives and Information Science*, 9(3), 16–34. Available at: <https://doi.org/10.4403/jlis.it-12480> <https://www.jlis.it/article/view/12480/11333>

Tallon, L. (2018). "Creating Access beyond metmuseum.org: The Met Collection on Wikipedia." *The Metropolitan Museum of Art Blog*. Consulted January 2, 2019. Available at: <https://www.metmuseum.org/blogs/now-at-the-met/2018/open-access-at-the-met-year-one>

Villaespesa, E. (2018). "Expanding Our Collection's Global Reach on the Spanish Wikipedia." *The Metropolitan Museum of Art Blog*. Consulted January 2, 2019. Available at: <https://www.metmuseum.org/blogs/collection-insights/2018/open-access-images-spanish-wikipedia>

Wikimedia Commons: GLAMwiki Toolset. Consulted December 22, 2018. Available at https://commons.wikimedia.org/wiki/Commons:GLAMwiki_Toolset

Wikimedia Commons: Structured data|GLAM. Consulted December 10, 2018. Available at https://commons.wikimedia.org/wiki/Commons:Structured_data/GLAM

Wikimedia Meta-Wiki: Strategy—Wikimedia movement—2017—Sources—Considering 2030: Future of reference and open knowledge. Consulted December 27, 2018. Available at https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2017/Sources/Considering_2030:_Future_of_reference_and_open_knowledge

Wikimedia Outreach: Wikipedian-inResidence. Consulted December 30, 2018. Available at https://outreach.wikimedia.org/wiki/Wikipedian_in_Residence

Wikimedia Statistics: Wikidata total page views. Consulted December 29, 2018. Available at <https://stats.wikimedia.org/v2/#/wikidata.org/reading/total-page-views/normal|bar|All|~total>

Wikipedia: Wikipedian in residence. Consulted December 20, 2018. Available at https://en.wikipedia.org/wiki/Wikipedian_in_residence

World Wide Web Consortium (W3C): Mission. Consulted January 14, 2019. Available at <https://www.w3.org/Consortium/mission>

Wyatt, L. (2010) End of my residency, *Witty's Blog*. Consulted January 12, 2019. Available at: <https://wittylama.com/2010/07/03/end-of-my-residency/>

Villaespesa, Elena and Navarrete, Trilce. "Museum Collections on Wikipedia: Opening Up to Open Data Initiatives." MW19: MW 2019. Published January 14, 2019. Consulted September 7, 2019. <https://mw19.mwconf.org/paper/museum-collections-on-wikipedia-opening-up-to-open-data-initiatives/>

C O M M U N I T Y C O N V E N



P L A T I N U M S P O N S O R



G O L D S P O N S O R



S I L V E R S P O N S O R



B R O N Z E S P O N S O R



P A R T N E R S



Subscribe to get our free MuseLetter and conference updates!

Subscribe

Presented by Museums and the Web, LLC

703 Dale Drive, Silver Spring, MD, 20910, USA

info@mwconf.org

Founded by Archives and Museum Informatics

www.archimuse.com

