Structured Data on Wikimedia Commons
Structuring and linking the world's heritage (and more!)

Sandra Fauconnier – Wikimedia Foundation
BAAC Conference 2018, Tallinn
sfaconnier@wikimedia.org / @sanseveria
Not just Wikipedia

WIKIQUOTE  WIKIMEDIA COMMONS  WIKIDATA  WIKIVERSITY  WIKISOURCE  WIKTIONARY

WIKIMEDIA LABS  WIKINEWS  WIKIBOOKS  WIKISPECIES  WIKIMEDIA INCUBATOR  WIKIVOYAGE
Wikis - Editable by anyone, collaborative
Freely licensed - can be re-used by anyone
Worldwide
Multilingual
Volunteer-driven
No ads - entirely funded by donations
Imagine a world in which every single human being can freely share in the sum of all knowledge. That's our commitment.

https://meta.wikimedia.org/wiki/Vision
A worldwide movement

Independent Wikimedia affiliates (chapters and regional user groups), August 2015

Lodewijk Gelauff, CC0
Hundreds of partnerships around the world since the mid-2000s

https://outreach.wikimedia.org/wiki/GLAM
Edit-a-thons
¡Ay, qué tiempos, señor Don Simón!

¡Ay, qué tiempos, señor don Simón! es una película mexicana de comedia de 1941. Dirigida por Julio Bracho y protagonizada por Joaquín Pardavé, Arturo de Córdova y Mapy Cortés, retrata y satiriza la vida de México a principios del siglo XX durante la época del Porfiriato.

<table>
<thead>
<tr>
<th>Índice [ocultar]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Sinopsis</td>
</tr>
<tr>
<td>2 Reparto</td>
</tr>
<tr>
<td>3 Contexto</td>
</tr>
<tr>
<td>4 Recepción</td>
</tr>
<tr>
<td>5 Referencias</td>
</tr>
</tbody>
</table>

**Sinopsis** [editar]

Durante la época del porfiriato, a principios del siglo XX, Inés (Mapy Cortés) es una atractiva joven que recientemente envió y que tiene por novio a un capitán del ejército llamado Miguel (Arturo de Córdova) quien la engaña con la bailarina tiple Coco Achondo. A fin de descubrir el engaño, Inés asiste en compañía de su amiga Beatriz (Anita Blanx) a un teatro exclusivo para caballeros en el que Coco se presentaba y en donde suele ser vista con Miguel. El lugar también es frecuentado por don Simón (Joaquín Pardavé), quien es presidente de la Liga de las Buenas Costumbres. Al momento de que Inés y Beatriz entran al lugar son vistas por las hermanas Méndez (Dolores Camarillo y Consuelo Guerrero de Luna), un par de solteronas que también pertenecen a la liga que preside don Simón.
Audiovisual media in Wikipedia articles?

Voice recording (short excerpt) from BBC's archives

https://en.wikipedia.org/wiki/Wangari_Maathai
First steps
The Netherlands Institute for Sound and Vision donates 392 videos to Commons

We expanded the collection videos on Wikimedia Commons with historical newsreel footage. Hereunder you can see some examples in this category all 392 videos can be found. This donation involves footage and newsreels from WWII and the Dutch state mines (coal mines) and footage of daily life. If you would like to help in matching videos to articles relevant to the subject, you can use this tool and it will take you directly to an overview of all videos that have not been used yet on the Dutch Wikipedia.

Media uploads to Wikimedia Commons: videos illustrating Wikipedia articles

Historical newsreels from the Netherlands Institute for Sound and Vision

https://lv.wikipedia.org/wiki/1._sievie%C5%A1u_%C5%A1aha_olimpi%C4%81de
https://commons.wikimedia.org/wiki/Category:Media_from_Open_Beelden
Marcel Breuer


Breuer tuli tunnetuksi polkupyörän ohjaustangosta ideaansa saaneasta putkinrunkoisesta huonekaluista, joista tuli yksi modernismin symboleista 1920-luvulla. Hän pääsi itseläiseltä viitattavasta vuotavan hampaiden arkkitehtien Le Corbusierin ja Ludwig Mies van der Rothen suosioon ja alko Gropiuksen avulla saada ensimmäisiä sisustusSuunnittelutilauksia.


Category details for Media from Open Beelden

98 months have a data point, with 367,000,512 page views in total. Click on individual time points in the graph to see monthly data.

Page views in 2018-9

Total monthly page views: 5,095,815. Download this table.

<table>
<thead>
<tr>
<th>Site</th>
<th>Pages Views</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dutch Wikipedia</td>
<td>1,961,210,063</td>
</tr>
<tr>
<td>English Wikipedia</td>
<td>3,411,965,757</td>
</tr>
<tr>
<td>German Wikipedia</td>
<td>165,307,959</td>
</tr>
<tr>
<td>Spanish Wikipedia</td>
<td>56,211,157</td>
</tr>
<tr>
<td>French Wikipedia</td>
<td>164,167,700</td>
</tr>
<tr>
<td>Indonesian Wikipedia</td>
<td>29,56,147</td>
</tr>
<tr>
<td>Russian Wikipedia</td>
<td>52,47,311</td>
</tr>
<tr>
<td>Italian Wikipedia</td>
<td>84,35,976</td>
</tr>
<tr>
<td>Polish Wikipedia</td>
<td>52,30,448</td>
</tr>
<tr>
<td>Portuguese Wikipedia</td>
<td>30,26,451</td>
</tr>
<tr>
<td>Japanese Wikipedia</td>
<td>32,25,964</td>
</tr>
</tbody>
</table>

http://tools.wmflabs.org/glamtools/baglama2/#gid=81&month=201809
Loic Tallon: "Creating Access beyond metmuseum.org: The Met Collection on Wikipedia"
Yoshua Bengio

Yoshua Bengio (born 1964 in France) is a Canadian computer scientist, most noted for his work on artificial neural networks and deep learning.[1][2]

Bengio received his Bachelor of Engineering (electrical engineering), Master of Science (computer science) and PhD (computer science) from McGill University.[3] He was a post-doctoral fellow at MIT (under Michael I. Jordan) and AT&T Bell Labs.[4] Bengio has been a faculty member at the Université de Montréal since 1993, heads the MILA (Montreal Institute for Learning Algorithms) and is co-director of the Learning in Machines & Brains project of the Canadian Institute for Advanced Research.[5]

Along with Geoffrey Hinton and Yann LeCun, Bengio is considered one of the three people most responsible for the advancement of deep learning during the 1990s and 2000s.[6] Whereas the other two went to work for Google and Facebook respectively, Bengio has stayed in academia. Among the computer scientists with the largest h-index, Bengio is the one with the most recent citations per day.[7][8]

In October 2016, Bengio co-founded Element AI, a Montreal-based business incubator that seeks to transform artificial intelligence (AI) research into real-world business applications.[9] In May 2017, Bengio announced that he was joining Montreal-based legal tech startup Botler AI, as a strategy adviser.[9]

Awards

In 2017, Bengio was named an Officer of the Order of Canada.[10] The same year, he was nominated Fellow of the Royal Society of Canada[11] and received the Marie-Victorin Quebec Prize.[12]

References


Mind of the Universe open licensing by a public broadcaster VPRO in partnership with the Netherlands Institute for Sound and Vision

https://en.wikipedia.org/wiki/Yoshua_Bengio
This Is How Open Publication of Broadcasting Content Can Make a Difference
By Charlery Smeets and Jesse de Vos June 13, 2018

Last year, the Dutch public broadcaster VPRO produced a ten-part documentary series titled *The Mind of the Universe*, in which leading scientists from various fields were interviewed about their work. Aware of their role as a public institution, and looking for bigger impact, VPRO decided to distribute this series under an open license, allowing other people to re-use the materials for educational purposes. To cater to the needs of these ‘secondary’ users, VPRO, in collaboration with The Netherlands Institute for Sound and Vision, launched a platform “Open Source Science TV”. On the platform the materials can be searched and downloaded, allowing anyone to create their own new productions with the materials.

We have the pleasure of talking a bit more about the project with Jesse de Vos, Researcher Interactive Media at The Netherlands Institute for Sound and Vision. Jesse was closely involved with the project and wrote down the experiences of the project in the paper: “Publication of Public Broadcasting Content in the Commons”.

— Could you tell us about your role in the project?
JVa After the publication of the materials on the platform “Open Source Science TV”, I uploaded the videos to Wikimedia Commons, the media database of Wikipedia. We tried to encourage the reuse of those videos in articles on Wikipedia, which proved to be quite a challenge for various reasons. After the project was finished I interviewed various participants and wrote a practitioner’s statement, reflecting on our experiences in publishing this content under an open license.

— As mentioned in the paper, in 2010, Peter B. Kaufman published the white paper “Video for Wikipedia and the Open Web”, about the open publication of video. Almost a decade later, however, publishing openly is still far from common practice. Could you tell

http://blog.euscreen.eu/2018/06/this-is-how-open-publication-of-broadcasting-content-can-make-a-difference/
Releasing your own productions under Creative Commons licenses

https://commons.wikimedia.org/wiki/Category:Media_from_Netherlands_Media_Art_Institute
Most recent approaches and challenges
Bridging data across the web

GLAM collections are not islands
Linking Open Data cloud diagram
https://www.lod-cloud.net
Linked Open Data in Wikimedia projects
• Centralizes interwiki links
• Centralizes data in infoboxes
• Offers an interface for ‘rich queries’

Structures the ‘sum of all human knowledge’

• Multilingual
• Referenced
• All data is CC0 – freely reusable by everyone
Wikidata as authority hub

https://tools.wmflabs.org/reasonator/?q=Q236656&lang=en
https://query.wikidata.org

http://tinyurl.com/y8gssvpj
Histropedia timeline - Swedish women who are actresses AND film directors
How do GLAMs(*) benefit from Wikidata?

(*) and major web platforms!

http://tinyurl.com/zqjz6q2
1 - have data on wikidata
2 - link to Wikidata
3 - use data from Wikidata
4 - work together through Wikidata
1 - Data publication on Wikidata places collections in their worldwide context
National Library of Wales uploaded data about its Welsh Landscape Collection on Wikidata and enriched it there.
Enrichment via Wikidata:
Welsh Landscape Collection
– depicted locations now geotagged (Query)
Multilinguality via Wikidata: Things depicted in the collections of National Library of Wales, shown in French (Query)
2 - Linking to Wikidata in collection databases makes metadata linked (and more robust)
Yle: keywords connected with Wikidata items

- Helps to link the different 'data silos' of YLE
- Allows tagging with niche topics that are not in other vocabularies

http://wikimedia.fi/2016/04/15/yle-3-wikidata/ & presentation by Pia Virtanen at EBU MDN workshop 2016
Get your vocabularies in Wikidata...

so Europeana and others can get them

Europeana offers a trusted online repository of cultural heritage objects. In order to achieve its mission it is crucial for the metadata described in Europeana to be semantically rich and multilingual.

Enriching the metadata with Linked Open Vocabularies has allowed us to improve our metadata. We perform automatic metadata enrichment using external value vocabularies and datasets such as GeoNames and DBpedia and exploit the semantic relations and translations offered by those vocabularies.

When selecting which vocabularies we use for our automatic enrichment, we seek to apply some criteria, especially for minimising our semantic commitment and abstracting from the

https://pro.europeana.eu/page/get-your-vocabularies-in-wikidata
Musical instruments on Wikidata linked to the MIMO vocabulary

http://tinyurl.com/yao9xmwq
3 - Data pulled from Wikidata can be used to enrich one's own collection website

Rush, (Performer [iii]). Lee, Geddy, (Instrumentalist [iii]). Lifeson, Alex, (Instrumentalist [iii]). Peart, Neil, (Instrumentalist [iii]).

Rush was a Canadian rock band comprising Geddy Lee (bass, vocals, keyboards), Alex Lifeson (guitars), and Neil Peart (drums, percussion, lyrics). Formed in 1968, the band went through several configurations until arriving at its longest and most popular line-up when Peart replaced original drummer John Rutsey in July 1974, two weeks before the group's first United States tour. W

Available copies

• 1 of 1 copy available at Conifer. (Show)
• 1 of 1 copy available at Laurentian University.

Current holds

0 current holds with 1 total copy.
Wikidata: a platform for your library’s linked open data

Seized with the desire to improve the visibility of Canadian music in the world, a ragtag band of librarians led by Stacy Allison-Cassan set out to host Wikipedia edit-a-thons in the style of Art+Feminism, but with a focus on addressing Canadian music instead. Along the way, they recognized that Wikidata offered a low-barrier, high-reward method of making that data not only visible but reusable as linked open data, and consequently incorporated Wikidata into their edit-a-thons. This is their story.

by Stacy Allison-Cassan†, Dan Scott‡

† Both authors contributed equally to this work.

Introduction

Creating and using linked open data (LOD) in library and GLAM (galleries, libraries, archives, museums) projects has historically been associated with a high level of institutional requirements. Erik et al. (2015 [3]) asserted the fact that LAM institutions are still having to select triplestores, SPARQL engines, indexing platforms, and other services means that there is still a relatively high bar for institutions to cross in taking up LOD projects. The requirement to select, host, and administer all of these systems establishes technical and resource barriers that can prevent organizations and individuals from participating in LOD entirely (Goddard and Byrne, 2010 [5]). Creating and publishing LOD has traditionally required technical skills to transform relational data, to support content-negotiation and alternate serializations, and to understand vocabularies and ontologies typically documented in RDF or OWL.

Wikidata, launched in 2012 by the Wikimedia Foundation as the machine readable store for all Wikimedia Foundation projects, is a freely available hosted platform that anyone—including libraries—can use to create, publish, and use LOD. Powered by Blazegraph, the platform offers a triplestore and high-availability SPARQL endpoint that (as of April 2018) has served roughly 3 million queries per day over the past year (Wikimedia Foundation, n.d. [11]): a full text search engine; and is administered by the Wikimedia Foundation. Its vocabulary is published and editable alongside other items in the platform using the same relatively user-friendly interfaces. In effect, Wikidata has responded to the barriers identified by Erik et al. and Goddard and Byrne by providing a ready-made platform for any person or organization that wants to create, publish, and use LOD, including libraries. In their 2018 IFLA discussion paper, Goddard and Byrne noted: Wikidata has a...
Angelica Kauffmann
Swiss-Austrian painter (1741-1807)

Maria Anna Angelika Kauffmann, usually known in English as Angelica Kauffman, was a Swiss Neoclassical painter who had a successful career in London and Rome. Remembered primarily as a history painter, Kauffmann was a skilled portraitist, landscape and decoration painter. She was one of the two female founding members of the Royal Academy in London in 1768.

- Birth Date: 30 Oct 1741
- Place of Birth: Chur
- Death Date: 5 Nov 1807 (age 66)
- Place of Death: Rome
- Spouse: Antonio Zucchi (1781-1795)
- Parents: Joseph Johann Kauffmann
- Genre: Portrait painting, History painting
4 - Work together through Wikidata as shared knowledge base
Shared knowledge infrastructure for digital preservation
http://wikidp.org/
Federated Linked Open Data with Wikibase

[Diagram showing relationships between Wikibase, Wikidata, Registries, and other data sources like Enslaved, GeneWiki network, and OSM.]
Catalog of internet artworks preserved by Rhizome.org using Wikibase.
Many faces of Wikibase: Rhizome’s archive of born-digital art and digital preservation

By Sandra Feuerstein
6 September 2018

Simple Net Art Diagram

The art happens here

MTAA ca. 1997

Rhizome, an arts organisation in New York City, was one of the early adopters of Wikibase, having been using it since 2015 for its archive of born-digital art and digital preservation activities. Sandra Feuerstein interviewed Dragan Espenschied (Rhizome’s preservation director), Lyndsey Alubaid (software curator), and Lizana Rosenova (external Ph.D. researcher) to ask them why and how they use Wikibase.

Sandra: On your website, you describe Rhizome as an organization that deals with born-digital art and culture. What does this mean?

Dragan: Rhizome was founded in 1996, as an online initiative by net artists. Internet-based art or net art is an art form that uses new media, and specifically the Internet, as its medium. So, we deal with art pieces that have been made especially for the Internet, not digitized ‘classical’ artworks. In the mid-1990s, this art form was emerging and quite new. Rhizome and many other organizations around the world have helped to make it more established. Today, Rhizome is an affiliate of the New Museum in New York City.

Image

WikibaseNYC conference explores the frontier of linked open data infrastructure

By Alex Stesin, Jake Orlowich and Jens Dihlig
24 October 2018

When you think about the work of art historians or genetics researchers, installing database software is not the first thing that comes to mind. For over 15 years, Wikibase, Wikimedia’s knowledge base—also available as free and open software, and can be used for external databases and linked open data projects. To highlight how organizations use Wikibase, the Wikimedia Foundation and Wikimedia Germany (Deutschland) are publishing a series of blog posts called the ‘Many faces of Wikibase’.

WikiBaseNYC conference explores the frontier of linked open data infrastructure

By Alex Stesin, Jake Orlowich and Jens Dihlig
24 October 2018

When you think about the work of art historians or genetics researchers, installing database software is not the first thing that comes to mind. For over 15 years, Wikibase, Wikimedia’s knowledge base—also available as free and open software, and can be used for external databases and linked open data projects. To highlight how organizations use Wikibase, the Wikimedia Foundation and Wikimedia Germany (Deutschland) are publishing a series of blog posts called the ‘Many faces of Wikibase’.

WikiBaseNYC conference explores the frontier of linked open data infrastructure

By Alex Stesin, Jake Orlowich and Jens Dihlig
24 October 2018

When you think about the work of art historians or genetics researchers, installing database software is not the first thing that comes to mind. For over 15 years, Wikibase, Wikimedia’s knowledge base—also available as free and open software, and can be used for external databases and linked open data projects. To highlight how organizations use Wikibase, the Wikimedia Foundation and Wikimedia Germany (Deutschland) are publishing a series of blog posts called the ‘Many faces of Wikibase’.

WikiBaseNYC conference explores the frontier of linked open data infrastructure

By Alex Stesin, Jake Orlowich and Jens Dihlig
24 October 2018

When you think about the work of art historians or genetics researchers, installing database software is not the first thing that comes to mind. For over 15 years, Wikibase, Wikimedia’s knowledge base—also available as free and open software, and can be used for external databases and linked open data projects. To highlight how organizations use Wikibase, the Wikimedia Foundation and Wikimedia Germany (Deutschland) are publishing a series of blog posts called the ‘Many faces of Wikibase’.

WikiBaseNYC conference explores the frontier of linked open data infrastructure

By Alex Stesin, Jake Orlowich and Jens Dihlig
24 October 2018

When you think about the work of art historians or genetics researchers, installing database software is not the first thing that comes to mind. For over 15 years, Wikibase, Wikimedia’s knowledge base—also available as free and open software, and can be used for external databases and linked open data projects. To highlight how organizations use Wikibase, the Wikimedia Foundation and Wikimedia Germany (Deutschland) are publishing a series of blog posts called the ‘Many faces of Wikibase’.

WikiBaseNYC conference explores the frontier of linked open data infrastructure

By Alex Stesin, Jake Orlowich and Jens Dihlig
24 October 2018

When you think about the work of art historians or genetics researchers, installing database software is not the first thing that comes to mind. For over 15 years, Wikibase, Wikimedia’s knowledge base—also available as free and open software, and can be used for external databases and linked open data projects. To highlight how organizations use Wikibase, the Wikimedia Foundation and Wikimedia Germany (Deutschland) are publishing a series of blog posts called the ‘Many faces of Wikibase’.

WikiBaseNYC conference explores the frontier of linked open data infrastructure

By Alex Stesin, Jake Orlowich and Jens Dihlig
24 October 2018

When you think about the work of art historians or genetics researchers, installing database software is not the first thing that comes to mind. For over 15 years, Wikibase, Wikimedia’s knowledge base—also available as free and open software, and can be used for external databases and linked open data projects. To highlight how organizations use Wikibase, the Wikimedia Foundation and Wikimedia Germany (Deutschland) are publishing a series of blog posts called the ‘Many faces of Wikibase’.

WikiBaseNYC conference explores the frontier of linked open data infrastructure

By Alex Stesin, Jake Orlowich and Jens Dihlig
24 October 2018

When you think about the work of art historians or genetics researchers, installing database software is not the first thing that comes to mind. For over 15 years, Wikibase, Wikimedia’s knowledge base—also available as free and open software, and can be used for external databases and linked open data projects. To highlight how organizations use Wikibase, the Wikimedia Foundation and Wikimedia Germany (Deutschland) are publishing a series of blog posts called the ‘Many faces of Wikibase’.
Structured Data on Wikimedia Commons 2017-19

https://commons.wikimedia.org/wiki/Commons:Structured_data
Strings, not things :(

October: Ten Days That Shook The World, a 1928 film by Grigori Aleksandrov and Sergei Eisenstein about the 1917 October Revolution.
Public Domain.
Video uploaded from YouTube to Commons by Racconish.
Structured Data on Commons
2017-2019

adding metadata on Commons in a *structured* & *machine-readable* format making Commons files easier to view, search, edit, organize and re-use, in many *languages*
Commons:Structured data

From Wikimedia Commons, the free media repository

The project Structured Data on Wikimedia Commons (2017–19) converts the free files on Wikimedia Commons to a structured and machine-readable format, so that they become easier to view, search, edit, organize and re-use. To achieve that, the Commons backend is migrated to Wikibase, the same technology as used for Wikidata.

What is this?

Wikimedia Commons holds a lot of (meta)data about the media files it hosts. Structuring this data more and making it machine-readable has many benefits: it makes it easier to view, search, edit, curate, use and re-use the files on Commons.

The Structured Data on Commons project is an effort to make this happen. In early 2017, this project was funded by the Alfred P. Sloan Foundation. For more information about the grant, see the documentation here.


Latest updates

https://commons.wikimedia.org/wiki/Commons:Structured_data
GLAM pilot projects

→ Diverse and representative
→ Support and documentation
files on Commons as Linked Open Data
multilinguality
refined APIs
  ○ data synchronization and metadata round-tripping!
  ○ more refined impact statistics
Commons:International Image Interoperability Framework

From Wikimedia Commons, the free media repository

This page in a nutshell: This page collects information on how Wikimedia Commons and other Wikimedia projects might make use of the International Image Interoperability Framework.

Contents

1. What is IIIF?
2. IIIF and Wikimedia projects
   2.1 How to stay up to date with IIIF and Wikimedia discussions
   2.2 IIIF tools and software for Wikimedia projects
      2.2.1 Wikimedia Commons
      2.2.2 Wikidata
3. People interested in IIIF on Wikimedia projects

What is IIIF? [edit]

The International Image Interoperability Framework (IIIF, pronounced as 'triple-eye-ef') is a standard for sharing images (and other media files) across the web, which makes image and media repositories interoperable with each other. IIIF is widely supported by cultural institutions around the world.

To get an idea – with examples – of what types of applications become possible with IIIF, see this blog post (June 2018) from the Wellcome Collection.

General IIIF links:
- Official website
- IIIF Showcased – software and websites that implement IIIF
- IIIF for Museums - Introductory slide-deck, November 2014
- Awesome IIIF - big list of IIIF resources

https://commons.wikimedia.org/wiki/Commons:International_Image_Interoperability_Framework
Thank you!

Questions?