

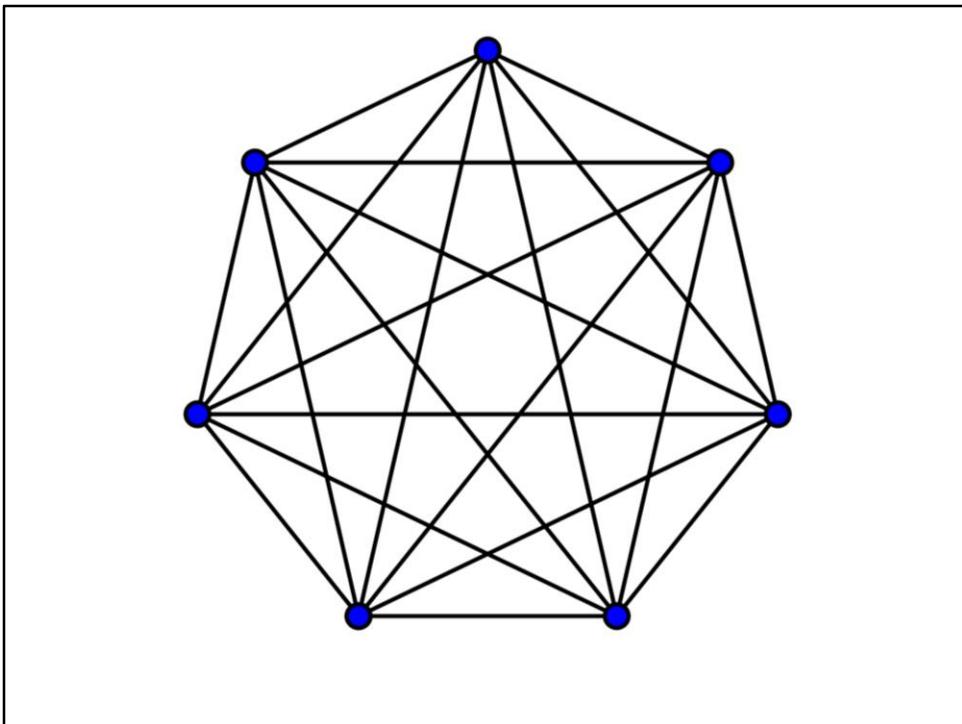
Wikidata introductie

Wat kun je ermee? Hoe voeg je zelf data toe?

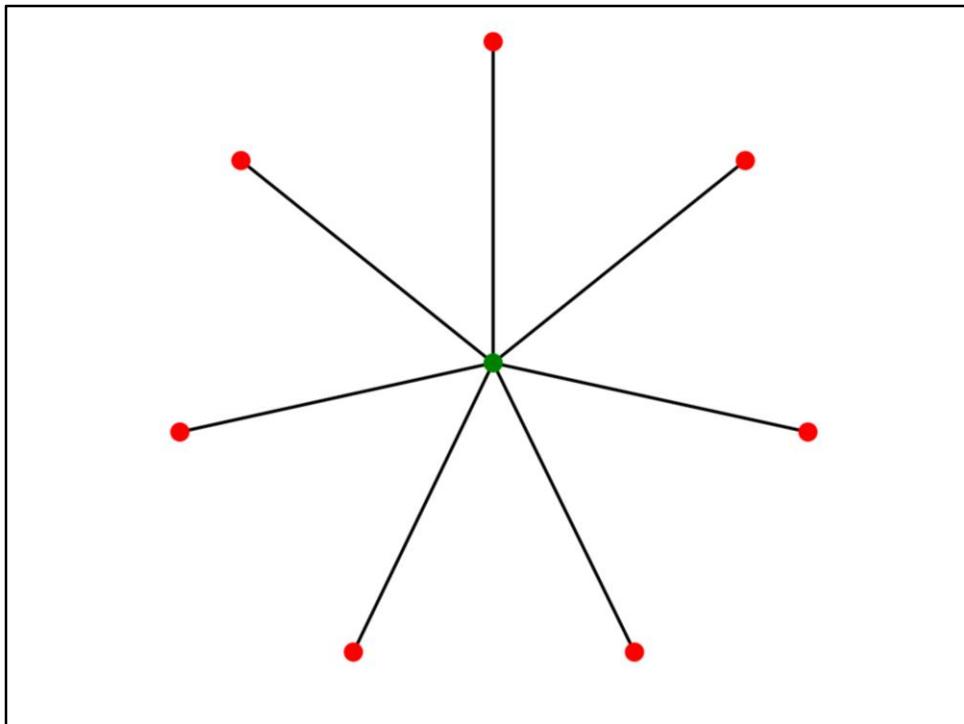
Maarten Dammers

@mdammers

- A general introduction to Wikidata, voor Vigin 2018
- Introduce myself. Nlwp, Commons, Wikidata, GLAMwiki, bots, Wiki Loves Monuments, uploads,



- Interlanguage (interwiki) links used to be local
- Lot's of bots running around doing updates.
- Didn't scale
- https://commons.wikimedia.org/wiki/File:Complete_graph_K7.svg



- Proposal to put this in a central place
- Start of Wikidata

universe (Q1)

totality of planets, stars, galaxies, intergalactic space, and all matter and energy
cosmos | The Universe

▼ In more languages

Language	Label	Description	Also known as
English	universe	totality of planets, stars, galaxies, interg...	cosmos The Universe
German	Universum	Gesamtheit der Energie, des Raumes u...	Weltall All Kosmos Weltraum
Spanish	universo	totalidad del espacio-tiempo, la materia ...	cosmos
French	Univers	ensemble des planètes, des étoiles, de...	Cosmos
Dutch	heelal	alle materie en energie binnen het geh...	universum kosmos cosmos
Portuguese	universo	tudo o que existe fisicamente, a totalida...	Universo cosmos

- It started with Q1, the universe somewhere back in October 2012
- This is what you call an item

 WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikisource store

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes
Upload file
Special pages
Permanent link
[Page information](#)
[Wikidata item](#)
[Cite this page](#)

Print/export
Create a book
Download as PDF
Printable version

In other projects

Article [Talk](#) Read Edit Edit source View history TW Search Wikipedia

This is the [current revision](#) of this page, as edited by [Drow](#) ([talk](#) | [contribs](#)) at 12:14, 26 March 2018 ([Astronomical concepts](#)). The present address (URL) is [https://en.wikipedia.org/w/index.php?title=Universe&oldid=832491446](#)

(diff) ← Previous revision | Latest revision (diff) | Newer revision → (diff)

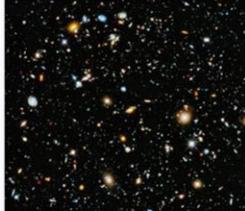
For other uses, see [Universe \(disambiguation\)](#)

The **Universe** is all of [space](#) and [time](#) and their contents, including [planets](#), [stars](#), [galaxies](#), and all other forms of [matter](#) and [energy](#). While the spatial size of the entire Universe is still unknown,¹⁵ it is possible to measure the [observable universe](#).

The earliest scientific models of the Universe were developed by [ancient Greek](#) and [Indian philosophers](#) and were [geocentric](#), placing [Earth](#) at the centre of the Universe.¹⁵²¹⁵³ Over the centuries, more precise astronomical observations led [Nicolaus Copernicus](#) to develop the [heliocentric model](#) with the Sun at the center of the [Solar System](#). In developing the [law of universal gravitation](#), [Sir Isaac Newton](#) built upon Copernicus's work as well as observations by [Tycho Brahe](#) and [Johannes Kepler's laws of planetary motion](#).

Further observational improvements led to the realization that our Solar System is located in the [Milky Way](#) galaxy, which is one of many galaxies in the Universe. It is assumed that galaxies are distributed uniformly and the same in all directions, meaning that the Universe has neither an edge nor a center. Discoveries in the early 20th century have suggested that the Universe had a beginning and that it is expanding¹⁵⁴ at an increasing rate.¹⁵⁵ Roughly eighty percent of mass in the Universe appears to exist in an unknown form called [dark matter](#) which cannot be directly observed.¹⁵⁶

The [Big Bang](#) theory is the prevailing [cosmological](#) description of the development of the Universe. Under this theory, space and time emerged together 13.799 ± 0.021 billion years ago¹⁵⁷ with a fixed amount of energy and matter that has become less dense as the Universe has expanded. After the initial expansion, the Universe cooled, allowing the first [subatomic particles](#) to form and then simple [atoms](#). Giant clouds later merged through gravity to form galaxies, stars, and everything else seen today. It is possible to see objects that are now further away than 13.799 billion [light-years](#) because


The [Hubble Ultra-Deep Field](#) image shows some of the most remote [galaxies](#) visible with present technology, each consisting of billions of stars. (Apparent image area about 1/79 that of a full moon)¹⁵⁸

Age (Within Lambda-CDM model)	13.799 ± 0.021 billion years ¹⁵⁷
Diameter	At least 156–554 billion light-years (47.8–170 billion parsecs) ¹⁵²¹⁵³ (maybe infinite ¹⁵⁹)

- <https://en.wikipedia.org/w/index.php?title=Universe&oldid=832491446>



- Wikidata was born
- That was almost 5 years ago
- First edit October 2012
- Relatively young project (Wikipedia started in 2001)
- https://commons.wikimedia.org/wiki/File:Logo_5th_Wikidata_Birthday.png

Johannes Vermeer (Q41264)

[edit]

Dutch 17th century painter

Jan Vermeer | Johannes Vermeer van Delft

▼ In more languages

Language	Label	Description	Also known as
English	Johannes Vermeer	Dutch 17th century painter	Jan Vermeer Johannes Vermeer van Delft
German	Jan Vermeer	Holländischer Maler	Johannes Vermeer Joannis van der Meer Jan Vermeer van Delft Joannis ver Meer
Spanish	Johannes Vermeer	No description defined	Johannes van der Meer Jan Vermeer Vermeer de Delft Vermer Jan Vermeer van Delft Johannes Vermeer van Delft Vermeer
French	Johannes Vermeer	peintre hollandais	Vermer Jan Vermeer van Delft Jan Vermeer Vermeer de Delft Vermeer
Dutch	Johannes Vermeer	Nederlands kunstschilder	

- The different parts of an item
- Label – No disambiguation
- Description – Short description
- Alias – Other names
- Every Wikipedia articles has an item

Wikipedia (82 entries) [edit]	
als	Jan Vermeer
ar	يوهانس فرمير
az	Yan Vermeer
be_x_old	Ян Вэрмээр
be	Ян Вермер
bg	Йоханес Вермеер
bn	জোহান্স ভার্মিয়ার
br	Johannes Vermeer
bs	Johannes Vermeer
ca	Johannes Vermeer de Delft
co	Jan Vermeer
cs	Jan Vermeer
cy	Johannes Vermeer
da	Johannes Vermeer ☈
de	Jan Vermeer ☈
el	Ιωάνες Βερμέερ
en	Johannes Vermeer
eo	Johannes Vermeer
es	Johannes Vermeer
et	Johannes Vermeer
eu	Jan Vermeer
fa	یوهانس فرمیر

- Links to Wikipedia
- Also to other sites
- Badges

place of birth (P19)

[edit]

the most specific known (e.g. city instead of country, or hospital instead of city)

birthplace | born in | POB | birth place | location born | born at | birth location | location of birth

▼ In more languages

Language	Label	Description	Also known as
English	place of birth	the most specific known (e.g. city instead of c...)	birthplace born in POB birth place location born born at birth location location of birth
German	Geburtsort	Ort, in dem ein Mensch geboren wurde	geboren in
Spanish	lugar de nacimiento	Sitio donde nació una persona. Usar el sitio m...	
French	lieu de naissance	lieu où la personne est née, le plus spécifique...	né à naissance née à
Dutch	geboorteplaats	de meest specifieke bekende geboorteplaats ...	
Portuguese	local de nascimento	local de nascimento da pessoa (o mais espec...)	lugar de nascimento nasceu em nascido em

- Properties, do be used on items
- Same structure with labels, descriptions, aliases
- Just like items all multilingual



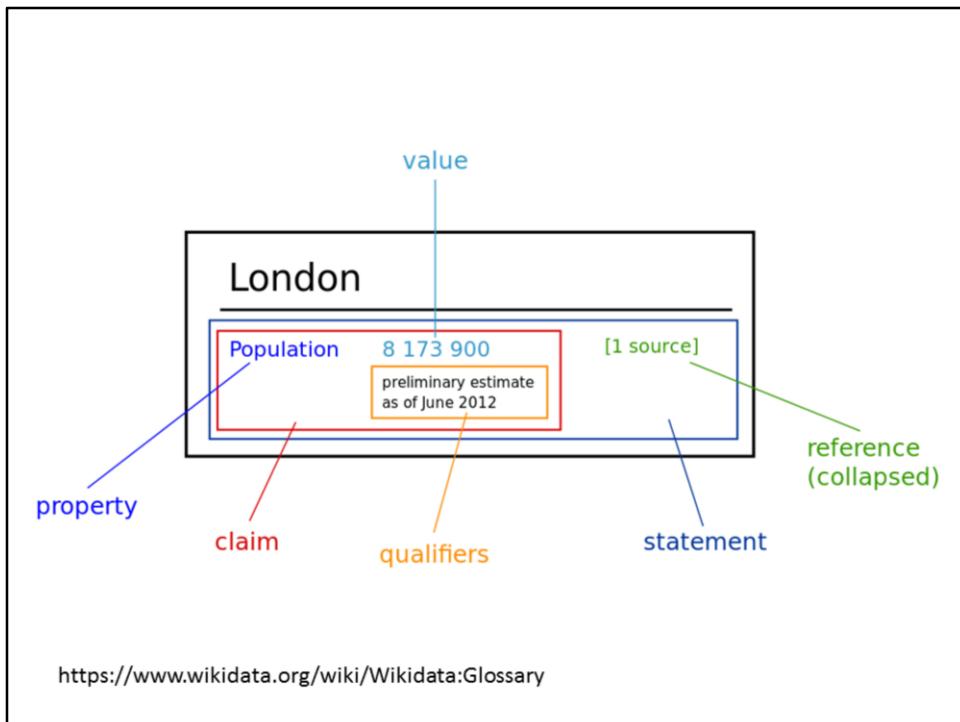
- A claim is a property -> right hand (can be item of some other type like string, image, coordinate)

inventory number	<input type="text"/> SK-A-2344	[edit]
collection	Rijksmuseum	
▼ 1 reference		[edit]
reference URL	https://www.rijksmuseum.nl /en/collection/SK-A-2344 ⓘ	
retrieved	9 April 2015	
		[add reference]
		[add]

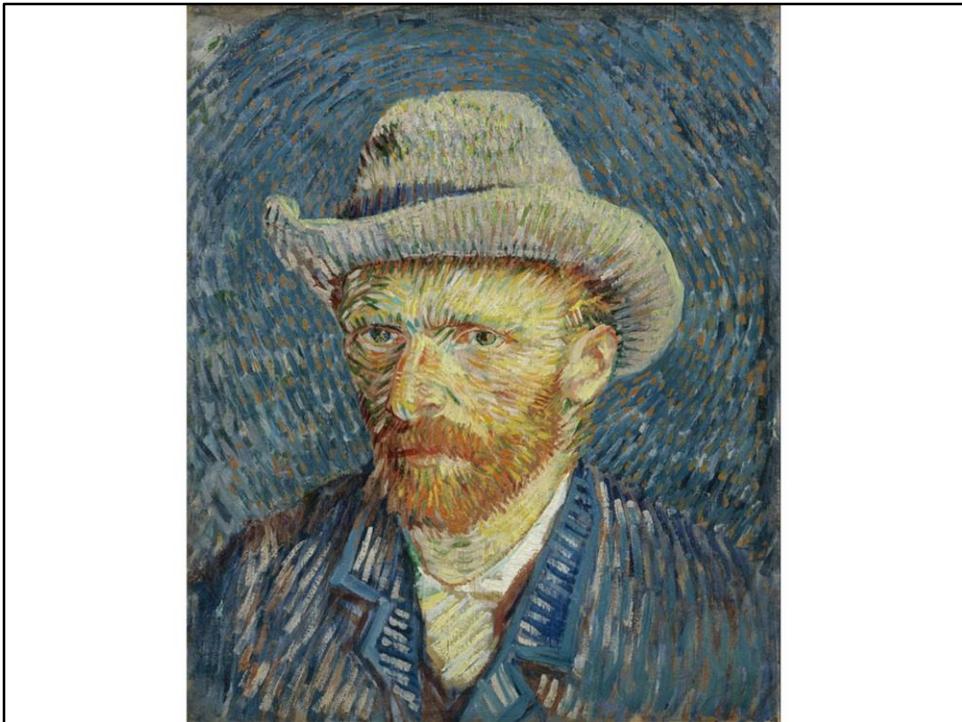
- Claim with source added and here we have a qualifier
- This is called a statement
- References, we like references *a lot*: There is no truth, only sources



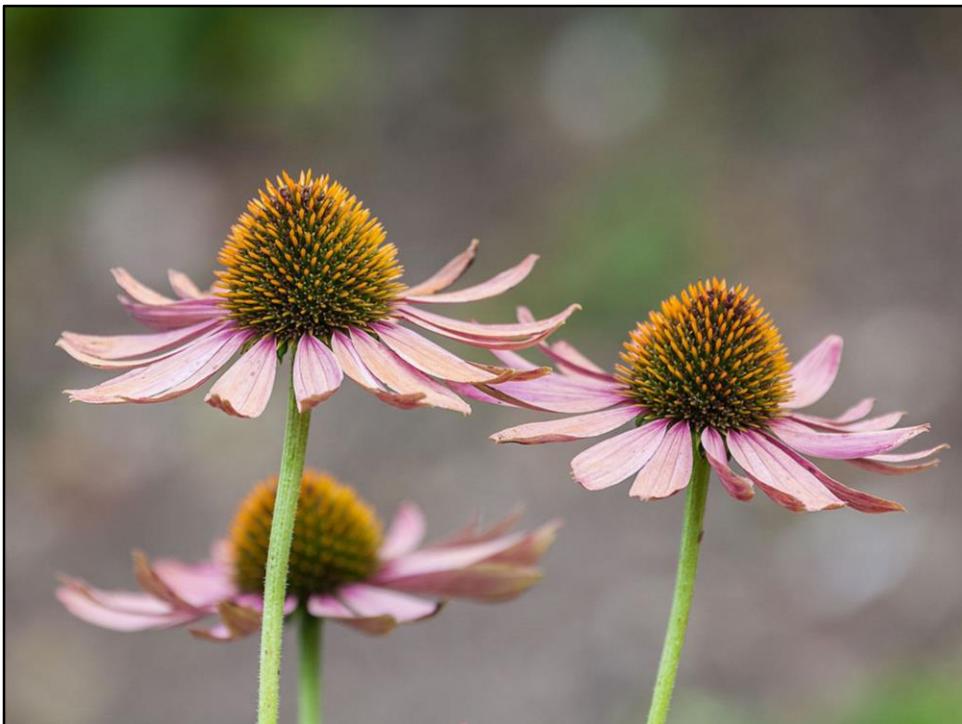
- Milkmaid!
- Previous statement is on the item about this painting
- https://commons.wikimedia.org/wiki/File:Johannes_Vermeer_-_Het_melkmeisje_-_Google_Art_Project.jpg (?)



- Here you see the different parts
- Difference between claim and statement
- https://commons.wikimedia.org/wiki/File:Wikidata_statement.svg



- Total about 37M items
- Millions of people (3M), many artists
- Deze personen linken ook weer naar andere databases
- https://commons.wikimedia.org/wiki/File:Vincent_van_Gogh_-_Self-portrait_with_grey_felt_hat_-_Google_Art_Project.jpg



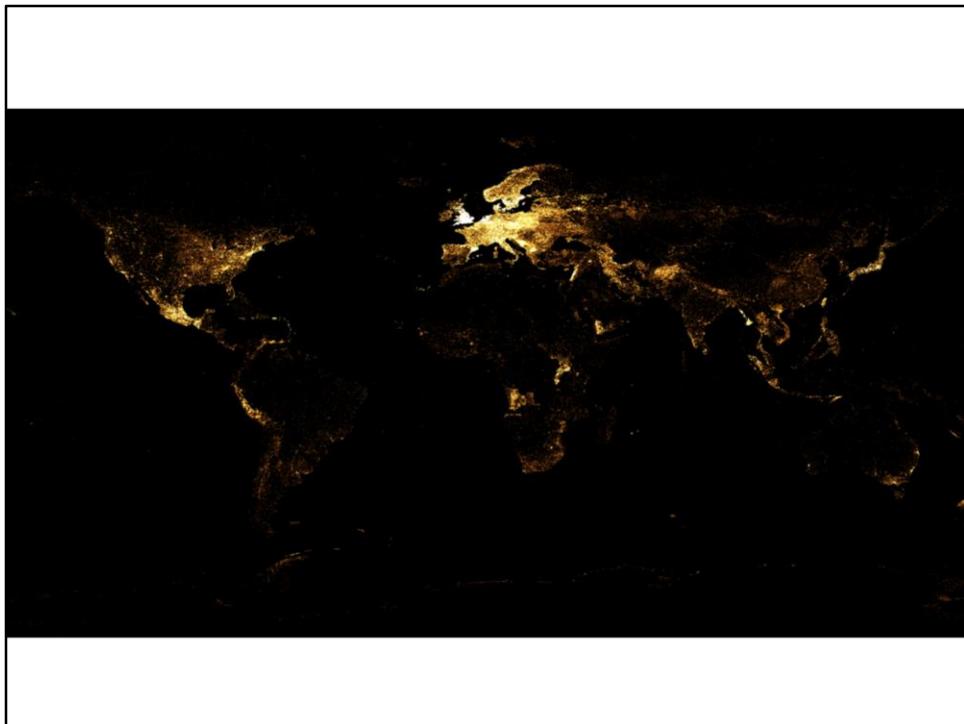
- About 2M items about taxons
- Flowers (Echinacea purpurea)
- https://commons.wikimedia.org/wiki/File:Echinacea_purpurea,_Jard%C3%ADn_Bot%C3%A1nico,_M%C3%BAnich,_Alemania,_2013-09-08,_DD_01.jpg



- About 2 M locations
- From the biggest cities (Hong Kong from Victoria peak)
- https://commons.wikimedia.org/wiki/File:Hong_Kong_Night_Skyline.jpg



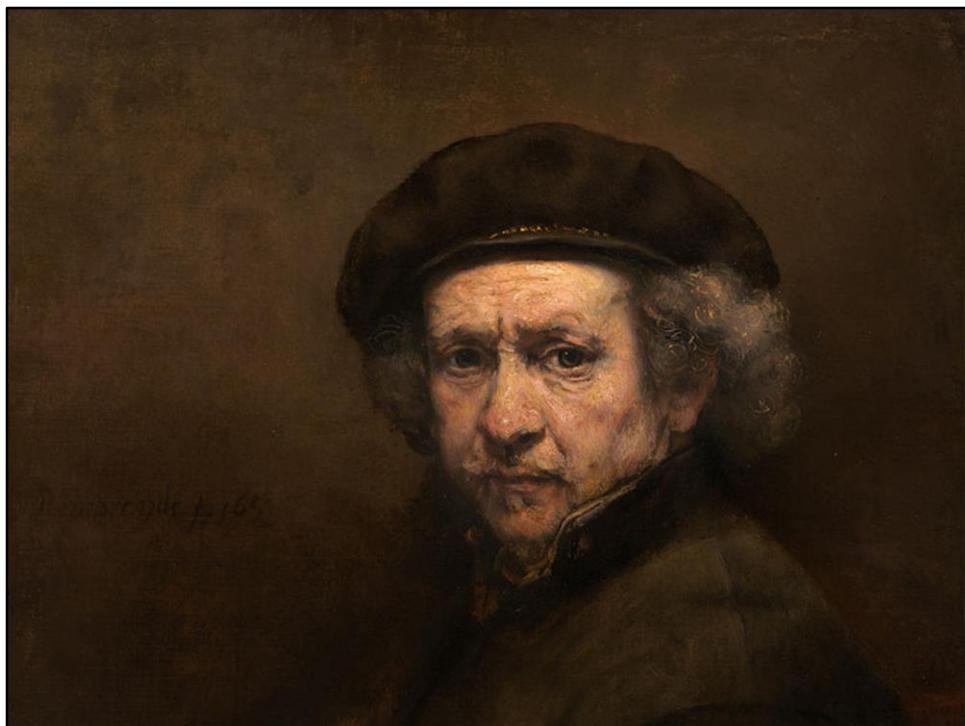
- To the tiniest towns Bodie California (ghost town)
- Most of these things come from Wikipedia. Now also slowly working on items without articles
- Took the picture myself, might be somewhere on Commons too



- Kaart van Wikidata items
- Elke item met coordinaat is een stipje



- Nederland zo goed te zien door de straten
- Daar kan je mooi de verbanden zien



- Het leggen van verbanden.
- Meeste straten vernoemd naar een schilder: 188 naar Rembrandt (en nog bezig)
- https://commons.wikimedia.org/wiki/File:Rembrandt_van_Rijn_-_Self-Portrait_-_Google_Art_Project.jpg



- For Wiki Loves Monuments I build the monuments database
- It contains 1,3 million records
- Create an entry for every historic building and kill it!
- Netherlands done, working on other countries
- https://commons.wikimedia.org/wiki/File:LUSITANA_WLM_2011_d.svg



- Historic buildings. More about that later
- Ongeveer 1,5 miljoen
- https://commons.wikimedia.org/wiki/File:Sint-Hippolytuskerk,_Hervormde_kerk.jpg

- K (2008) PowerPoint files citations. *Information Science and Technology* 19: 1–10.
- M (2008) Assessing the impact of disciplinary information analysis of online syllabuses. *Journal of the American Association Science and Technology* 59: 2060–2069.
- E, Cabezas-Clavijo Á (2012) Google scholar metrics: An assessment scientific journals. *El Profesional De La Información* 37: 1–10.
- wall M (2007) Google scholar citations and Google Web/URL multi-discipline exploratory analysis. *Journal of the American Information Science and Technology* 58: 1055–1065.
- K (2007) Impact of data sources on citation counts and rankings of Web of science vs Scopus and Google scholar. *Journal of the Society for Information Science and Technology* 58: 2105–2125.
- Thelwall M (2009) Google book search: Citation analysis for social media in the humanities. *Journal of the American Society for Information Science and Technology* 60: 1537–1549.
- Harnad S, Carr L (2006) Earlier web usage statistics as predictors of citation impact. *Journal of the American Society for Information Science and Technology* 57: 1060–1072.
- F (2005) Statistical relationships between downloads and citations at the individual documents within a single journal. *Journal of the American Society for Information Science & Technology*, 56: 1088–1097.
- L (2002) Research dissemination and impact: Evidence from web site loads. *Journal of Finance* 57: 485–499.
- J (2013) Altmetrics. In: Cronin B, Sugimoto C, editors. *Bibliometrics and Beyond: Metrics-Based Evaluation of Scholarly Research*. Cambridge: MIT Press, in press.
- J, Shema H, Thelwall M (2013) Bibliographic References in Web 2.0. In: Cronin B, Sugimoto C, editors. *Bibliometrics and Beyond: Metrics-Based Evaluation of Scholarly Research*. Cambridge: MIT Press, in press.
- J, Groth P, Taraborrelli D (2012) The Altmetrics Collection. *PLoS ONE* 7: e48753.
- Taraborrelli D (2008) Soft peer review: social software and distributed scientific evaluation. In: Proceedings of the 8th International Conference on the Design of Systems. pp. 99–110.
- A, Kats M, Fang X, et al. (2012) Tweeting the social impact based evaluation. In: Proceedings of Kidney Week 2011. *PLoS ONE* 7: e48753.
- Nielsen F (2004) <http://firstmonday.org/humanities/1997/1872>. Accessed 21 January, 2013.
- Yan K-K, Gerstein M (2011) The spread of scientific Web usage statistics in PLoS Article-Level Metrics. *PLoS ONE* 6: e20000.
- Bar-Ilan J (2012a) JASIST@mendeley. Presented at the ACM Web Science Conference Workshop on Altmetrics. Evanston, IL. Available: <http://altmetrics.org/altmetrics12/bar-ilan>. Accessed 21 January 2013.
- Bar-Ilan J (2012b) JASIST 2001–2010. *Bulletin of the American Society for Information Science and Technology* 38: 24–28.
- Bar-Ilan J, Haustein S, Peters I, Priem J, Shema H, et al. (2012) Beyond citations: Scholars' visibility on the social Web. In: Proceedings of the International Conference on Science and Technology Indicators. Montréal, Canada. pp. 98–109.
- Li X, Thelwall M (2012) F1000, Mendeley and traditional bibliometrics indicators. In: Proceedings of the 17th International Conference on Science and Technology Indicators. Montréal, Canada. pp. 451–551.
- Li X, Thelwall M, Giustini D (2012) Validating online reference managers' scholarly impact measurement. *Scientometrics* 91: 461–471.
- Priem J, Piwowar HA, Hemminger BM (2012) Altmetrics in the wild: social media to explore scholarly impact. [arXiv.org](http://arxiv.org/abs/1203.4745). Available: <http://arxiv.org/abs/1203.4745>. Accessed 2013 January 21.
- Schlogl C, Stock WG (2004) Impact and relevance of LIS j scientometric analysis of international and German-language LIS Citation analysis versus reader survey. *Journal of the American Information Science and Technology* 55: 1155–1168.
- Rowlands I, Nicholas D (2007) The missing link: Journal usage in Proceedings 59: 222–228.
- Kurtz M, Ballen J (2010) Usage bibliometrics. *Annual Review of Science and Technology* 44: 1–64.
- NLM (2013) MEDLINE Fact sheet. Available: <http://www.nlm.nih.gov/factsheets/medline.html>. Accessed 2013 March 20.
- Wan JK, Hua PH, Rousseau R, Sun XK (2010) The immediacy index (II): experiences using a Chinese Scientometrics 82: 555–566.
- Walman L, Costas R (2013) F1000 recommendations as a situation: A comparison with citations. [arXiv.org](http://arxiv.org/abs/1304.4775). Available: <http://arxiv.org/abs/1304.4775>. Accessed 2013 April 4.

- https://www.wikidata.org/wiki/Wikidata:Database_reports/Popular_items
- We verzuipen bijna in de 14M wetenschappelijke artikelen
- <https://meta.wikimedia.org/wiki/WikiCite>
- There is no truth, only sources
- Thesaurus specifiek gebied, net zoals Wikipedia is Wikidata alles omvattend



- 280.000 paintings
- https://commons.wikimedia.org/wiki/File:Jan_Asselijn_-_De_doorbraak_van_de_Sint_Anthonisdijk_in_Amsterdam.jpg



The Art & Architecture Thesaurus (AAT)®

Catherine wheel or rose window?

AAT is a structured vocabulary, including terms, descriptions, and other information for generic concepts related to art, architecture, conservation, archaeology, and other cultural heritage.



The Cultural Objects Name Authority (CONA)®

Mona Lisa or La Gioconda? CONA includes titles, attributions, depicted subjects, and other information about works of art and architecture as derived from scholarly literature, museum collections, special collections, archives, libraries, and indexing projects. CONA is linked to the AAT, TGN, and ULAN.



The Getty Thesaurus of Geographic Names (TGN)®

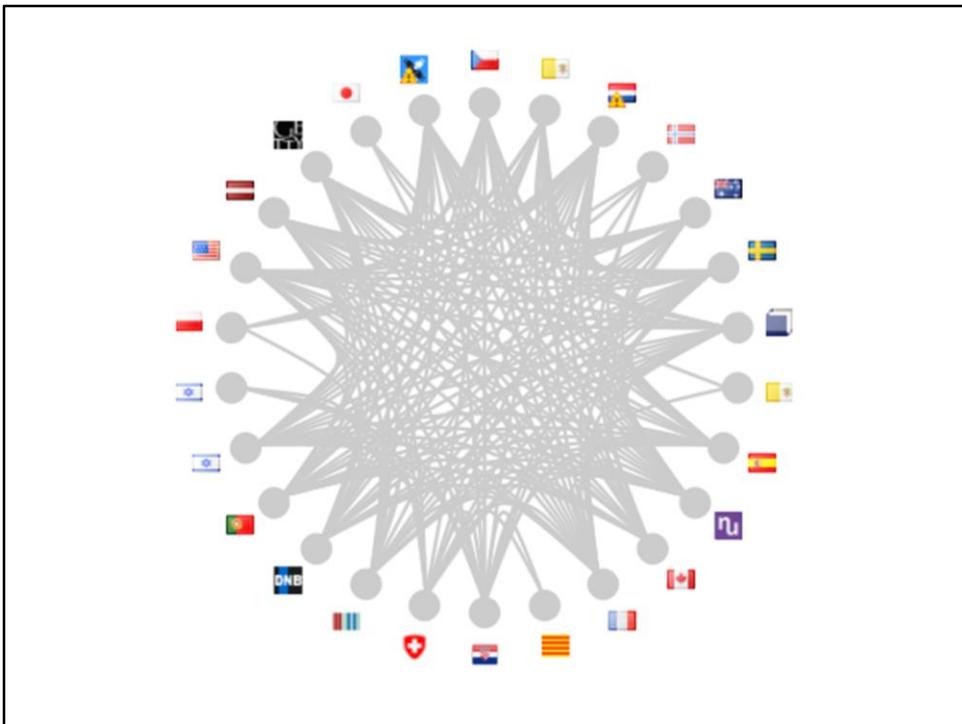
London or Londinium? TGN is a structured vocabulary, including names, descriptions, and other information for current and historical administrative places and physical features important to art and architecture.



The Union List of Artist Names (ULAN)®

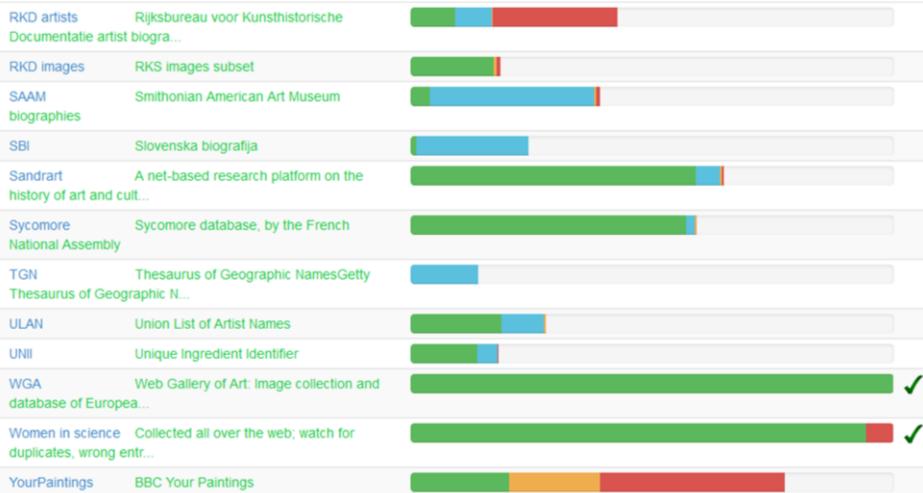
Titian or Tiziano Vecellio? ULAN is a structured vocabulary, including names, biographies, related people, and other information about artists, architects, firms, studios, museums, patrons, sitters, and other people and groups involved in the creation and study of art and architecture.

- Getty does authority control, we do to!
- We link to many. AAT is concepts, CONA is objects, TGN is locations, ULAN is people
- Screenshot of <http://www.getty.edu/research/tools/vocabularies/>



- Viaf! Already a history on Wikipedia
- Over 1 million links now
- GND over 0,5 M links now
- Viaf is switching from Wikipedia to Wikidata. YOU SHOULD DO THAT TOO!
- We also have the connected items (like NTA, 120.000) and we run bots
- https://viaf.org/viaf/64013650/#Rembrandt_Harmenszoon_van_Rijn,_1606-1669

<https://tools.wmflabs.org/mix-n-match/>



- Mix'n'Match is used to make initial matches
- Bot also do imports.
- Still a lot of work to do
- Meer dan duizend verschillende identifiers nu!
- Hier kan je zelf helpen
- <https://tools.wmflabs.org/mix-n-match/>

<https://tools.wmflabs.org/wikidata-game/>

 **Random game**
A random game, from the list below, is selected each time.
Makes for more variety.

 **Merge items**
Some topics have duplicate items on Wikidata. Two items with the same title or alias will be suggested to you.
Merge identical topics, tag items as different, or skip an item pair if you are not sure. Please be very careful!

 **Person**
Many items about people on Wikidata have no "instance of" property.
Decide if one of them is a [human](#), or skip the item if you are not sure.

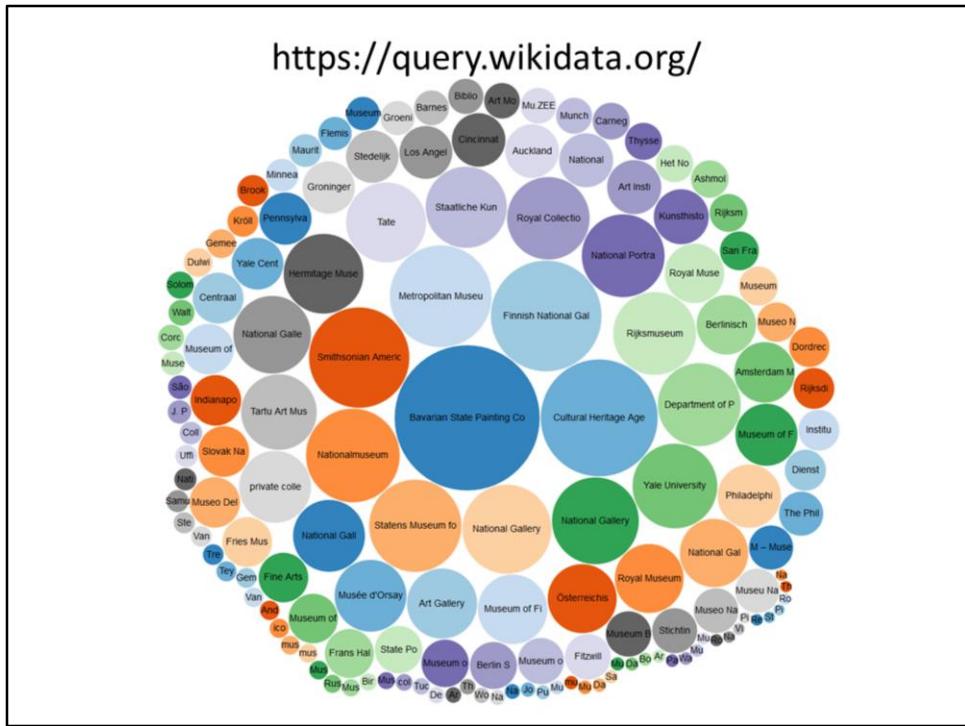
 **Gender**
Many items about people on Wikidata have no [gender](#) property set.
Add a sex/gender statement "male" or "female", or skip the item if you are not sure.

 **Occupation**
These people have no [occupation](#), but their articles suggest some.

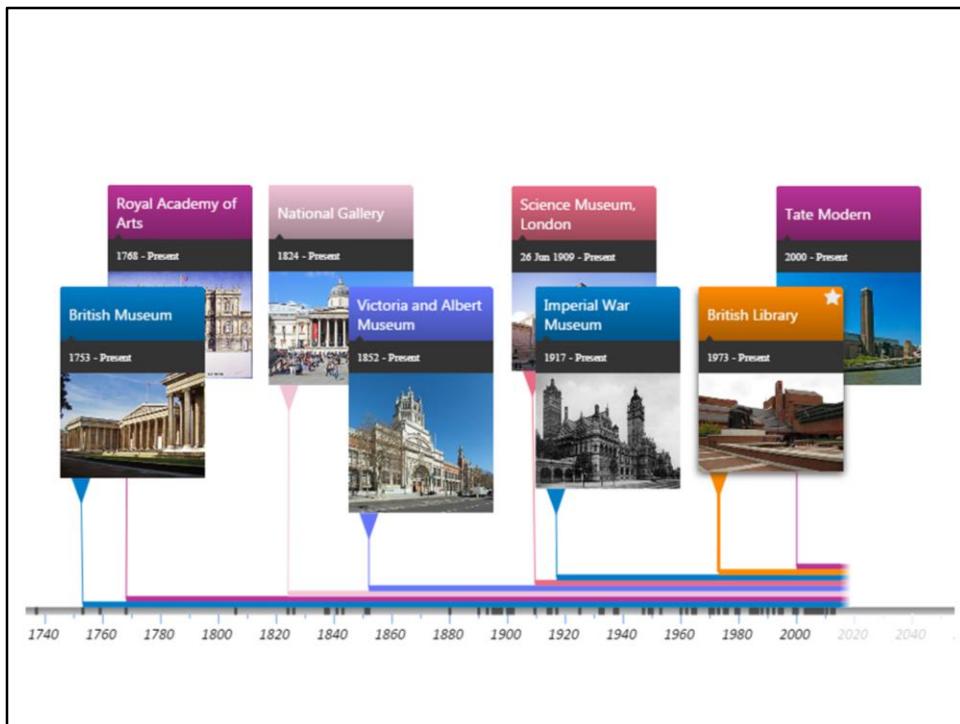
 **Alma mater**
These people have no [alma mater](#), but their articles suggest some.

 **Country of citizenship**
These people have a birth place, but no country of citizenship (P27).

- Wikidata games zijn een andere goede manier om te helpen



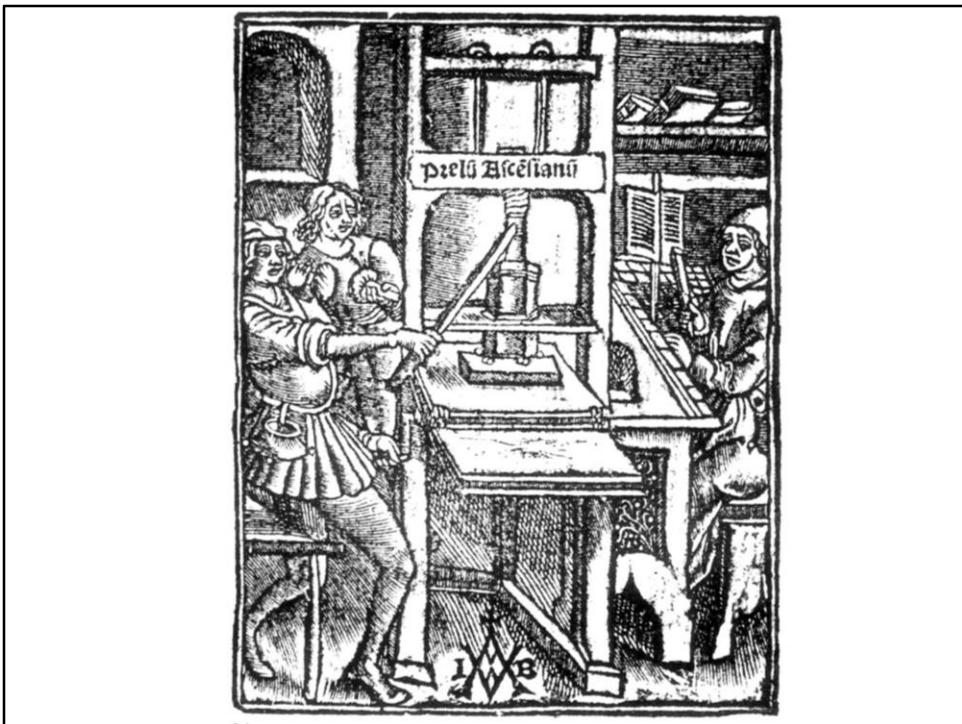
- SPARQL
- Query taaltje om door de hele database te gaan
- Voorbeeld: De grootste collecties schilderijen
- Misschien volgende keer een SPARQL workshop?



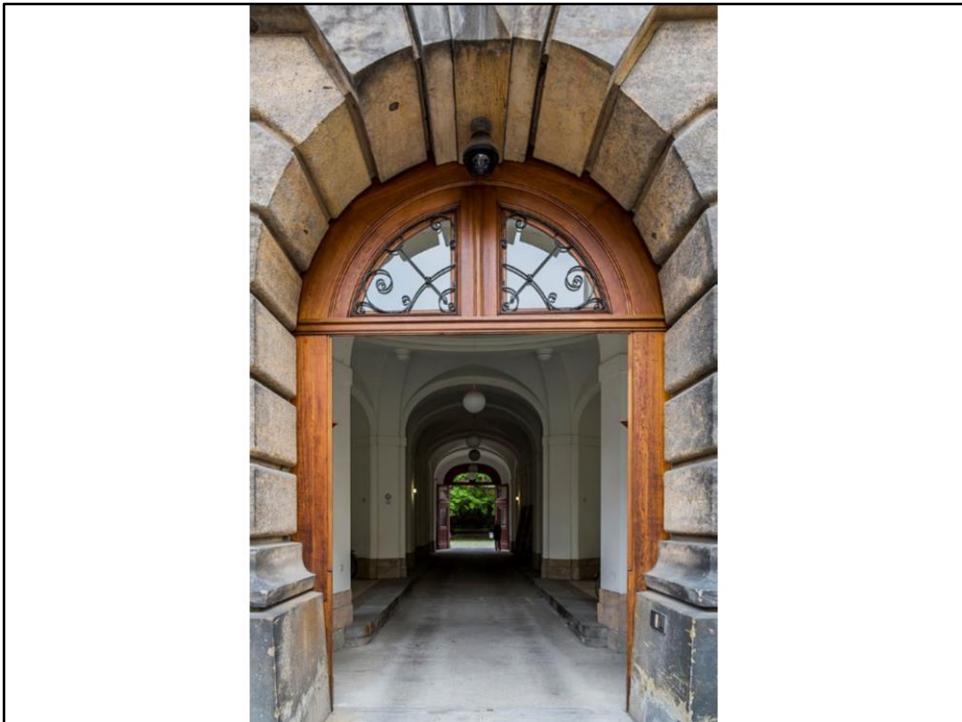
- Histropedia is een van de hergebruikers
- Siri kan je ook gebruiken
- Eurowings gebruikt het ook
- Ask Platypus noemen
- https://commons.wikimedia.org/wiki/File:Histropedia_timeline_of_GLAM_institutions_in_London.jpg



- Notability for example is very much in movement
- Is it useful data to add or is it just dumping trash?
- For the Wikipedians: Inclusionists vs deletionists
- Data has to be used and maintainable, we don't want to become a data landfill
- https://commons.wikimedia.org/wiki/File:LOCAL_RESIDENTS_DUMP_TRASH_IN_CLYDE_KEPLEY%27S_SANITARY_LAND_FILL_-_NARA_-_542549.jpg



- Publish your data online
- Own website or platform like Europeana
- [https://commons.wikimedia.org/wiki/File:Jodocus_Badius_Ascensius_Druckermarke_Paris_1519_\(Isny\).jpg](https://commons.wikimedia.org/wiki/File:Jodocus_Badius_Ascensius_Druckermarke_Paris_1519_(Isny).jpg)



- Open it under a free license to encourage re-use
- Have an open access policy
- https://commons.wikimedia.org/wiki/File:Dresden_Germany_Kunstakademie-01.jpg



- ST Jerome in the Desert
- Wiki people like to work in peace and quiet from their homes
- They need to be able to find the data online. If the data isn't online, it looks like it doesn't exist
- https://commons.wikimedia.org/wiki/File:Giovanni_Girolamo_Savoldo_-_Saint_Jerome_-_Google_Art_Project.jpg



- Any questions?
- Photo van Sebastiaan