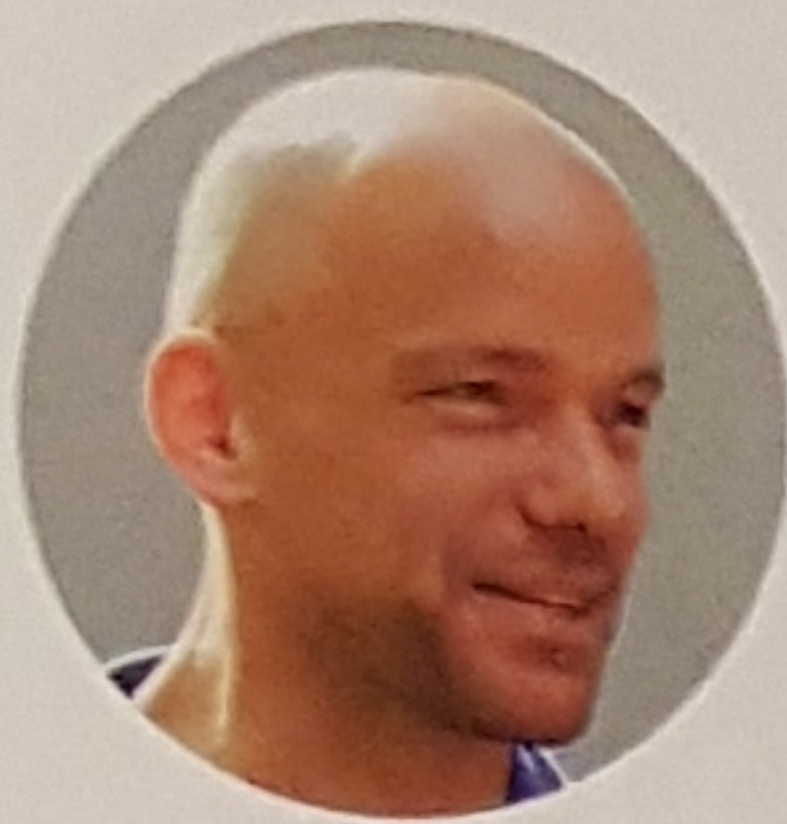


WIKIBASE: DE VOORDELEN VAN WIKIDATA, ZONDER DE NADELEN



Olaf Janssen

Wikimedia-coördinator bij de KB.

Dit artikel is beschikbaar onder de CC-BY-SA 4.0-licentie

Wikidata is de vrije en gratis database die gestructureerde, linked open data van Wikipedia en overige Wikimedia-projecten centraal opslaat. Wikidata kan gelezen en bewerkt worden door zowel mensen als computers. Iedereen kan zonder beperkingen (want: CC0-licentie) gebruikmaken van de beschikbare gegevens en verbanden, én kan er bovendien zelf aan bijdragen. De database bestaat inmiddels zeven jaar. In die tijd heeft een actieve community er ruim 62 miljoen personen, plaatsen, gebeurtenissen, organisaties, organismen en andere 'dingen' (entiteiten) in beschreven. Door de sterke koppeling met Wikipedia – sinds jaren in de top-10 van meest geraadpleegde websites – zijn de gegevens in Wikidata bovendien uiterst zichtbaar. Het uiteindelijke doel is om de hele wereld op een gestructureerde manier te beschrijven, zodat computers ermee overweg kunnen. Vanwege de vrije licentie en de veelgebruikte protocollen (SPARQL, JSON, RDF, XML en, Mediawiki API) wordt de data uit Wikidata ook veelvuldig buiten de Wiki-context hergebruikt, in Google en Amazons Alexa bijvoorbeeld.

Weinig controle

Hoe geweldig Wikidata ook is, voor kennis- en erfgoedinstellingen kan juist die openheid, schrijfbaarheid en onderwerpsbreedte een nadeel vormen. Dit soort instituten heeft vele decennia geïnvesteerd in het zeer zorgvuldig beheren van gespecialiseerde, diepgaande catalogi en thesauri. Hun angst om deze datasets via Wikidata open te maken is dan ook begrijpelijk. Er is te weinig controle, die overijverige community kan alles zomaar aanpassen en de minutieuze beschrijving van dat heel speci-

De openheid, schrijfbaarheid en onderwerpsbreedte van Wikidata kunnen voor kennis- en erfgoedinstellingen een nadeel zijn. Wie liever meer controle over zijn datasets houdt, zet een eigen Wikibase op.



Beeld: pixabay / geralt

Plaque uit de collectie van het Kunstmuseum Den Haag

fieke collectiestuk past niet goed genoeg in het brede en daardoor soms oppervlakkige datamodel van Wikidata.

Wikibase

Einde oefening voor *LOD-hungry* en *Wiki-aware* instellingen? Nee, gelukkig kun je de

software waar Wikidata op draait – dus Wikidata minus de data – ook op een eigen server installeren en er dan je eigen datamodel en data instoppen. Dan heb je een eigen Wikibase opgezet. De Wikibase-software is open source en heeft alle voordelen van Wikidata, minus bovengenoemde nadelen. Over een eigen Wikibase-instantie heb je een stuk meer controle: je kunt alleen geautoriseerde gebruikers het recht geven om bij te dragen en je kunt er je eigen diepgaande, specialistische data(model) in kwijt. Tegelijkertijd kan iedereen – mensen en software – jouw Wikibase gewoon raadplegen, net als bij Wikidata. En functionaliteit die voor Wikidata ontwikkeld wordt, komt vaak ook terug in Wikibase. SPARQL, Mediawiki API, datavisualisaties, kant-en-klare code *snippets* voor Python en andere talen, alles out of the box.

Ontwikkelingen

De nationale bibliotheken van Duitsland (DNB) en Frankrijk (BnF) doen inmiddels serieuze experimenten met Wikibase om hun thesauri naar de LOD-cloud te krijgen. In Nederland doet het Kunstmuseum Den Haag als een van de eerste instellingen een pilot om de website delftwaardewerk.nl een eigen Wikibase te geven. Om partijen uit de academische en erfgoedwereld kennis te laten maken met de voor- en nadelen van Wikibase en Wikidata, is de KB van plan begin 2020 een studiedag rondom deze onderwerpen te organiseren.

Verder lezen

Inside the Alexa-Friendly World of Wikidata: tinyurl.com/y3l5hqwd

Wikibase: wikiba.se

Wikibase BnF: www.transition-bibliographique.fr/fne/french-national-entities-file/

Wikibase DNB: blog.factgrid.de/archives/1527