



OpenRefine

Werkmaterialen introductieworkshop OpenRefine, KB, 4 juli 2023

(Olaf Janssen, KB, 27 juni 2023)

In deze workshop gaan we aan de slag met de winnaars van de Halewijnprijs

- Website: <https://www.halewijnprijs.nl/> (tabblad Winnaars)
- Wikipedia: <https://nl.wikipedia.org/wiki/Halewijnprijs>
- Wikidata: <https://www.wikidata.org/wiki/Q1570893>

De Halewijnprijs is een Nederlandse literatuurprijs

- Voorbeeld van een Wikidata-item over een literatuurprijs incl. de winnaars:
P.C. Hooft-prijs - <https://www.wikidata.org/wiki/Q1379623>
- Voorbeeld Wikidata-item over een winnaar van literatuurprijzen:
Theun de Vries - <https://www.wikidata.org/wiki/Q2143934>

Reconciliatie

- Lijst van beschikbare reconciliatie-services voor OpenRefine : <https://reconciliation-api.github.io/testbench/>
- Wikidata (NL) reconciliatie-service: <https://wikidata.reconci.link/nl/api>
- NTA reconciliatie-service (o.b.v. Termennetwerk NDE): <https://termennetwerk-api.netwerkdigitaalerfgoed.nl/reconcile/http://data.bibliotheken.nl/thesp/sparql>

Extra data van de schrijvers ophalen uit Wikidata

- Afbeelding - <https://www.wikidata.org/wiki/Property:P18>
- Geboorteplaats - <https://www.wikidata.org/wiki/Property:P19>
- Geboortedatum - <https://www.wikidata.org/wiki/Property:P569>
- Beroep - <https://www.wikidata.org/wiki/Property:P106>

URLs op Wikimedia Commons

- Naam afbeelding = "Eva Meijer 2018.jpg"
- Bijbehorende URL op Commons = [https://commons.wikimedia.org/wiki/File:Eva Meijer 2018.jpg](https://commons.wikimedia.org/wiki/File:Eva_Meijer_2018.jpg)
- GREL: "https://commons.wikimedia.org/wiki/File:" + value

Geboorteplaatsen schrijvers – o.b.v. Wikipedia en andere bronnen

- Peter Drehmanns = Roermond
bron: https://nl.wikipedia.org/wiki/Peter_Drehmanns
- Fred Papenhove = Den Haag
bron: <https://www.singeluitgeverijen.nl/auteur/fred-papenhove>
- Stijn van der Loo = Eindhoven
bron: <https://www.brabantcultureel.nl/2019/10/13/stijn-van-der-loo-stopt-dementie-scheiden-en-een-midlifecrisis-in-een-verhaal>
- Adriaan Jaeggi = Wassenaar
bron: https://nl.wikipedia.org/wiki/Adriaan_Jaeggi
- Laurens Spoor = Den Haag
bron: https://theaterencyclopedie.nl/wiki/Laurens_Spoor
- Gijs IJlander = Alkmaar
bron: https://nl.wikipedia.org/wiki/Gijs_IJlander
- André Janssens = Sint-Amandsberg
bron: https://www.dbnl.org/tekst/_ons003198701_01/_ons003198701_01_0063.php
- Jan Huyskens = Horn
bron: <http://streektaalzang.nl/strk/limb/limbihun.htm>

Voorbeeld NTA URIs over Hermsen, Joke J. (1961-)

- HTML: <http://data.bibliotheken.nl/doc/thes/p075253429>
- JSON: <http://data.bibliotheken.nl/doc/thes/p075253429.json>
- GREL: value + ".json"

Extractie van externe IDs (ISNI, VIAF, Wikidata) uit JSON-response NTA

- GREL: value.parseJson()["@graph"][1]["sameAs"].join("-")

Toevoegen van alle winnaars aan Wikidata-item over Halewijnprijs, voor opbouw OR-schema

- Winnaar: <https://www.wikidata.org/wiki/Property:P1346>
- Qualifier = tijdstip / point in time (jaartal):
<https://www.wikidata.org/wiki/Property:P585>

T.b.v. de bronvermelding

- reference URL / URL voor bron (P854) = <https://www.halewijnprijs.nl>
- retrieved / geraadpleegd op (P813) = 4 juli 2023