

# Setting up your own Wikibase reconciliation service (e.g. for OpenRefine)



OpenRefine

PLACE OF REVOLUTION  
AND RECONCILIATION

**Olaf Janssen / Hanno Lans**

[olaf.janssen@kb.nl](mailto:olaf.janssen@kb.nl) / [hanno@datascape.nl](mailto:hanno@datascape.nl)

[@ookgezellig](#) / [@hannolans](#)

Q66439268 / Q61996680


[slideshare.net/OlafJanssenNL](https://slideshare.net/OlafJanssenNL)

KB } nationale  
bibliotheek

4th meeting Dutch Wikibase Knowledge Group

Wednesday 23 November 2022

# Reconciliation of person names against Wikidata using OpenRefine

 **OpenRefine** clipboard [Permalink](#)

Facet / Filter [Undo / Redo](#) 127 / 127

44 rows [Schema](#) [Issues](#) [Preview](#)

Show as: **rows** records Show: 5 10 **25** 50 100

▼ All ▼ Jaar ▼ PrijsWinnaar

☆	🗨	1.	2013	Eveline Hagenbeek
☆	🗨	2.	2000	Paul Lochtenberg
☆	🗨	3.	1998	Luc van Brussel
☆	🗨	4.	1994	Bob vanden Eynde
☆	🗨	5.	1984	Herman Lutgerink
☆	🗨	6.	1982	Ruud van Gessel
☆	🗨	7.	1982	Jan Koonings
☆	🗨	8.	1980	Michel Desmyter
☆	🗨	9.	1978	Toos Staalman
☆	🗨	10.	1978	Paul Vanbossuyt
☆	🗨	11.	1978	Herman van Winsen
☆	🗨	12.	1976	Frits Enk
☆	🗨	13.	1976	W.J. Jansen
☆	🗨	14.	1976	Nel Bakker
☆	🗨	15.	1976	Yonne Dillen
☆	🗨	16.	1976	Gaston Gheuens
☆	🗨	17.	1976	H.A. Voet
☆	🗨	18.	1975	Johan Breedveld
☆	🗨	19.	1975	Rudolf Bruyn
☆	🗨	20.	1975	Toos Staalman
☆	🗨	21.	1974	R.B. Barkman
☆	🗨	22.	1974	Michel Desmyter
☆	🗨	23.	1974	Per Justesen
☆	🗨	24.	1973	Wim Bischot
☆	🗨	25.	1971	Achilles Gautier

**Using facets and filters**

Use facets and filters to select subsets of your data to act on. Choose facet and filter methods from the menus at the top of each data column.

Not sure how to get started?  
[Watch these screencasts](#)

## Reconciliation

For these person names (*strings*) we want to look up the corresponding Wikidata Q-items (*things*)

# Reconciliation of person names against Wikidata using OpenRefine

44 rows   Schema   Issues   Preview

Show as: **rows** records   Show: 5 10 **25** 50 100 500 1000 rows

Facet / Filter

Using facets and filters of your data to analyze and filter methods at the top of each column.

Not sure how to get started? Watch these screencasts.

All   Jaar   **Prijswinnaar**

- Facet ▶
- Text filter
- Edit cells ▶
- Edit column ▶
- Transpose ▶
- Sort...
- View ▶
- Reconcile ▶**
  - Start reconciling...**
  - Facets ▶
  - Actions ▶
  - Copy reconciliation data...
  - Use values as identifiers...
  - Add entity identifiers column...

☆	🗨	1.	2013	
☆	🗨	2.	2000	
☆	🗨	3.	1998	
☆	🗨	4.	1994	
☆	🗨	5.	1984	
☆	🗨	6.	1982	
☆	🗨	7.	1982	
☆	🗨	8.	1980	
☆	🗨	9.	1978	
☆	🗨	10.	1978	
☆	🗨	11.	1978	Herman van Winsen
☆	🗨	12.	1976	Frits Enk
☆	🗨	13.	1976	W.J. Jansen
☆	🗨	14.	1976	Nel Bakker
☆	🗨	15.	1976	Yonne Dillen
☆	🗨	16.	1976	Gaston Gheuens
☆	🗨	17.	1976	H.A. Voet
☆	🗨	18.	1975	Johan Breedveld
☆	🗨	19.	1975	Rudolf Bruyn
☆	🗨	20.	1975	Toos Staalman
☆	🗨	21.	1974	R.B. Barkman
☆	🗨	22.	1974	Michel Desmyter
☆	🗨	23.	1974	Per Justesen
☆	🗨	24.	1973	Wim Bischof
☆	🗨	25.	1971	Achilles Gautier

# Reconciliation of person names against Wikidata using OpenRefine

44 rows

Schema

Issues

Preview

## Reconcile column "PrijsWinnaar"

### Services

Nederlandse Thesaurus van  
Auteursnamen



RKDartists



VIAF



Wikidata reconci.link (en)



Wikidata reconci.link (nl)



Wikimedia Commons (en)



Pick a service or extension on left

Add standard service...

Discover services...

Start reconciling...

Cancel

# Reconciliation of person names against Wikidata using OpenRefine

OpenRefine 44 rows Schema Issues Preview

## Reconcile column "PrijsWinnaar"

[Access service API](#)

Reconcile each cell to an entity of one of these types:

- mens  
Q5
- wetenschappelijk artikel  
Q13442814
- gilde van Brussel  
Q111525228
- sculptuur  
Q860861
- Sint-Lucasgilde  
Q1110992

Also use relevant details from other columns:

**Column Include? As property**

Jaar

Reconcile against type:

Reconcile against no particular type

Auto-match candidates with high confidence

Maximum number of candidates to return

# Reconciliation of person names against Wikidata using OpenRefine

OpenRefine

44 rows Schema Issues Preview

## Reconcile column "PrijsWinnaar"

Access service API

Reconcile each cell to an entity of one of these types: mens

Also use relevant details from other columns: Column Include? As property

**Reconcile cells in column PrijsWinnaar to type Q5**  
57% complete Cancel

44 rows Schema Issues Preview

Show as: rows records Show: 5 10 25 50 100 500 1000 rows

All	Jaar	PrijsWinnaar	
	1.	2013	Eveline Hagenbeek
	2.	2000	Paul Lochtenberg
	3.	1998	Luc van Brussel

Reconcile against type:

Reconcile against no particular type

Auto-match candidates with high confidence

Maximum number of candidates to return

Add standard service... Discover services... Start reconciling... Cancel


# Reconciliation of person names against Wikidata using OpenRefine

44 rows   Schema   Issues   Preview

Show as: **rows** records   Show: 5 10 **25** 50 100 500 1000 rows

**All**   **Jaar**   **Prijswinnaar**

		Jaar	Prijswinnaar
☆	🗨	1. 2013	<a href="#">Eveline Hagenbeek</a> <small>Choose new match</small>
☆	🗨	2. 2000	<a href="#">Paul Lochtenberg</a> <small>Choose new match</small>
☆	🗨	3. 1998	<a href="#">Luc van Brussel</a> <small>Choose new match</small>
☆	🗨	4. 1994	<a href="#">Bob vanden Eynde</a> <small>Choose new match</small>
☆	🗨	5. 1984	<a href="#">Herman Lutgerink</a> <small>Choose new match</small>
☆	🗨	6. 1982	<a href="#">Ruud van Gessel</a> <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Ruud van G <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Ruud van G <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Create new item <small>Search for match</small>
☆	🗨	7. 1982	<a href="#">Jan Koonings</a> <small>Choose new match</small>
☆	🗨	8. 1980	<a href="#">Michel Desmyter</a> <small>Choose new match</small>
☆	🗨	9. 1978	<a href="#">Toos Staalman</a> <small>Choose new match</small>
☆	🗨	10. 1978	<a href="#">Paul Vanbossuyt</a> <small>Choose new match</small>
☆	🗨	11. 1978	<a href="#">Herman van Winsen</a> <small>Choose new match</small>
☆	🗨	12. 1976	<a href="#">Frits Enk</a> <small>Choose new match</small>
☆	🗨	13. 1976	<a href="#">W.J. Jansen</a> <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> <a href="#">W.J. Jansen</a> (100)



Bob vanden Eynde (Q111605502)  
winnaar ANV-Visser Neerlandia-prijs voor Drama

Add standard service...

Discover services...

Start reconciling...

Cancel

# Reconciliation service for Wikidata

## Wikidata reconciliation for OpenRefine

This web service can be used to align datasets to Wikidata items in OpenRefine.

Use the following URL in OpenRefine: <https://wikidata.reconci.link/en/api>.

Replacing "en" by another language code will display items and properties in your language, when they are available.

This interface works with OpenRefine from 2.6 rc2 onwards. It is not compatible with Google Refine.

- [Documentation of the protocol](#)
- [Documentation of this application for developers and Wikibase admins](#)
- [Source](#)
- [Bug tracker](#)
- [Service status](#)

## Tutorials for OpenRefine users

- [Official documentation in English](#)
- [Official documentation in Japanese](#) by 山本 泰智
- [Tutorial by Antoine Courtin](#) (in French)
- [Comparison of reconciliation services for OpenRefine](#) by Aaron Tay (in English)
- [Using OpenRefine to Reconcile Name Entities](#) by Karen H. (in English)
- [Reconciling in OpenRefine - eine Einführung von histHub](#) (in German)

## Videos

- [Reconciliation in OpenRefine](#), screencast by Owen Stephens (in English, 24:06)
- [OpenRefine demo at WikidataCon 2017](#), live demo by Arthur P. Smith (in English, 19:52)
- [Exemple de réconciliation et d'enrichissement avec Wikidata](#) (silent screencast, 3:27) by Ettore Rizza
- [Wikidata editing with OpenRefine](#), screencast by Antonin Delpeuch (in English, 3 videos of 5 minutes each)

## Documentation

Contents:

- [Reconciling via unique identifiers](#)
- [Property paths](#)
- [Special properties](#)
- [Subfields](#)

### Reconciling via unique identifiers

If a unique identifier is supplied in an additional property, then reconciliation candidates will first be searched for using the unique identifier.

If you only have unique identifiers and no names for the entities you want to reconcile, you can therefore supply a fake column of names.

### Property paths

<https://wikidata.reconci.link>



# Reconciliation service for Wikidata



## TABLE OF CONTENTS

- 1. Introduction**
  - 1.1 Data Matching on the Web
  - 1.2 History of the Reconciliation API
  - 1.3 External Resources
  - 1.4 Versions
    - 1.4.1 0.1
    - 1.4.2 0.2
    - 1.4.3 This Draft
  - 1.5 Conformance
- 2. Core Concepts**
  - 2.1 Entities
  - 2.2 Types
  - 2.3 Properties
  - 2.4 Property Values
  - 2.5 Identifier and Schema Spaces
- 3. Service Definition**
  - 3.1 Service Manifest
  - 3.2 Overview of Possible Routes
  - 3.3 HTTP(S) Access
  - 3.4 Cross-Origin Access
  - 3.5 Error Handling and Rate-limiting
- 4. Reconciliation Queries**
  - 4.1 Structure of a Reconciliation Query
  - 4.2 Reconciliation Query Responses
  - 4.3 Sending Reconciliation Queries to a Service

## Reconciliation Service API

A protocol for data matching on the Web



Draft Community Group Report 10 November 2022

### Latest editor's draft:

<https://reconciliation-api.github.io/specs/latest/>

### Editors:

[Antonin Delpuch](#) (University of Oxford)  
[Adrian Pohl](#) (Hochschulbibliothekszentrum NRW)  
[Fabian Steeg](#) (Hochschulbibliothekszentrum NRW)  
[Thad Guidry Sr.](#)   
[Osma Suominen](#) (National Library of Finland)

Add Yourself Here!

### Participate:

[GitHub reconciliation-api/specs](#)

[File a bug](#)

[Commit history](#)

[Pull requests](#)

Copyright © 2022 the Contributors to the Reconciliation Service API Specification, published by the [Entity Reconciliation Community Group](#) under the [W3C Community Contributor License Agreement \(CLA\)](#). A human-readable [summary](#) is available.

## Reconciliation Service API

### Abstract

This document describes the reconciliation service API, a protocol edited by the [W3C Entity Reconciliation Community Group](#). It is intended as a comprehensive and definitive specification of this API in its given state. Various aspects of this API need to be improved, as hinted by notes throughout this document.

<https://reconciliation-api.github.io/specs/latest/>

# Reconciliation service for Wikidata

Search docs

## CONTENTS:

- Installing the reconciliation service
- Architecture overview
- Scoring mechanism
- Testing infrastructure
- Documenting

## Welcome to the documentation of the OpenRefine-Wikibase reconciliation interface

This software offers a reconciliation interface for a Wikibase instance, following the [specifications of the reconciliation API](#).

This manual is intended for developers and Wikibase administrators. Reconciliation users should instead refer to the main page of the corresponding reconciliation endpoint, where user documentation can be found. For Wikidata, this is available at <https://wikidata.reconci.link/>.

### Contents:

- [Installing the reconciliation service](#)
  - [Requirements](#)
  - [Configuration](#)
  - [Installing with Docker](#)
  - [Installing manually](#)
  - [Deploying in production](#)
  - [Tips about Redis configuration](#)
- [Architecture overview](#)
  - [Reconciliation](#)
  - [Auto-complete \(suggest\) services](#)
  - [Preview](#)
  - [Data extension](#)

<https://openrefine-wikibase.readthedocs.io/en/latest/>

# Reconciliation service for Wikidata


```
JSON  Onbewerkte gegevens  Headers
Opslaan  Kopiëren  Alles samenvouwen  Alles uitvouwen  JSON filteren
▼ defaultTypes:
  ▼ 0:
    id:      "Q35120"
    name:    "entity"
  ▼ extend:
    ▼ property_settings:
      ▼ 0:
        default: 0
        help_text: "Maximum number of values to return per row (0 for no limit)"
        label: "Limit"
        name: "limit"
        type: "number"
      ▼ 1:
        ▼ choices:
          ▼ 0:
            name: "Any rank"
            value: "any"
          ▼ 1:
            name: "Only the best rank"
            value: "best"
          ▼ 2:
            name: "Preferred and normal ranks"
            value: "no_deprecated"
        default: "best"
        help_text: "Filter statements by rank"
        label: "Ranks"
        name: "rank"
        type: "select"
      ▼ 2:
        ▼ choices:
          ▼ 0:
            name: "Any statement"
            value: "any"
          ▼ 1:
            name: "At least one reference"
            value: "referenced"
          ▼ 2:
            name: "At least one non-wiki reference"
            value: "no_wiki"
        default: "any"
        help_text: "Filter statements by their references"
```







## Service endpoints (as needed by OpenRefine)

- <https://wikidata.reconci.link/en/api>
- <https://wikidata.reconci.link/nl/api>
- <https://wikidata.reconci.link/fr/api>
- Etc..

# Reconciliation of filenames against Wikimedia Commons

Reconcile column "PrijsWinnaar"

Services 

- Nederlandse Thesaurus van Auteursnamen 
- RKDartists 
- VIAF 
- Wikidata reconci.link (en) 
- Wikidata reconci.link (nl) 
- Wikimedia Commons (en)** 

Pick a service or extension on left

Add standard service... Discover services... Start reconciling... Cancel

# Reconciliation of filenames against Wikimedia Commons

Reconcile column "PrijsWinnaar"

**OpenRefine** Nederland Japan 7voor demo.xlsx [Permalink](#)

Facet / Filter Undo / Redo 70 / 70

1048 records Schema Issues 8 Preview

Show as: rows **records** Show: 5 10 25 50 100 500 1000 records

All	TitleNL	TitleEN	dcx:recordIdentifier	CommonsTitle	fullimageURL	
☆	1.	Kaswaars	Cassowary	KONB11:3-5	File:Cassowary - KONB11:3-5.jpg new <a href="#">Choose new match</a>	<a href="http://resolver.kb.nl/resolve?urn=urn:gvn:KONB11:3-5&amp;role=image&amp;size=large">http://resolver.kb.nl/resolve?urn=urn:gvn:KONB11:3-5&amp;role=image&amp;size=large</a>

Using facets and filters

Use facets and filters to select subsets of your data to act on. Choose facet and filter methods from the menus at the top of each data column.

Not sure how to get started?  
[Watch these screencasts](#)

String

Add standard service...

Discover services...

Start reconciling...

Cancel

# Reconciliation of filenames against Wikimedia Commons

1048 records   Schema   Issues **8**   Preview

Show as: **rows** **records**   Show: 5 **10** 25 50 100 500 1000 records

<input type="checkbox"/> All	<input type="checkbox"/> TitleNL	<input type="checkbox"/> TitleEN	<input type="checkbox"/> dcx:recordIdentifier	<input type="checkbox"/> CommonsTitle	<input type="checkbox"/> fullimageURL
	1.	Kaswaars	Cassowary	KONB11:3-5	<a href="http://resolver.kb.nl/resolve?urn=urn:grole=image&amp;size=large">http://resolver.kb.nl/resolve?urn=urn:grole=image&amp;size=large</a>

Facet

Text filter

Edit cells

Edit column

Transpose

Sort...

View

Reconcile

- Start reconciling...
- Facets
- Actions
- Copy reconciliation data...
- Use values as identifiers...
- Add entity identifiers column...

# Reconciliation of filenames against Wikimedia Commons

## Reconcile column "PrijsWinnaar"

### Services

Nederlandse Thesaurus van  
Auteursnamen

RKDartists

VIAF

Wikidata reconci.link (en)

Wikidata reconci.link (nl)

Wikimedia Commons (en)

Pick a service or extension on left

Add standard service...

Discover services...

Start reconciling...

Cancel

# Reconciliation of filenames against Wikimedia Commons


1048 records   Schema   Issues 8   Preview

## Reconcile column "PrijsWinnaar"

Services

1048 records   Schema   Issues   Preview

Show as: rows **records**   Show: 5 10 25 50 100 500 1000 records

<input type="checkbox"/> All	<input type="checkbox"/> TitleNL	<input type="checkbox"/> TitleEN	<input type="checkbox"/> dcx:recordIdentifier	<input type="checkbox"/> CommonsTitle	<input type="checkbox"/> fullimageURL
<input type="checkbox"/> ☆ <input type="checkbox"/> 1.	Kaswaars	Cassowary	KONB11:3-5	<a href="#">File:Cassowary - KONB11-3-5.jpg</a> <a href="#">Choose new match</a>	 <b>File:Cassowary - KONB11-3-5.jpg</b> 750 x 1063; 249.35 KB

Thing (file)

Add standard service...   Discover services...   Start reconciling...   Cancel



# Reconciliation service for Wikimedia Commons

## Wikimedia Commons reconciliation for OpenRefine

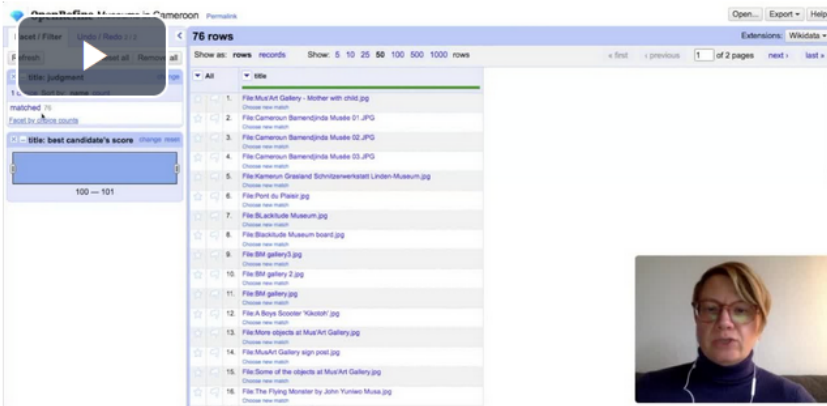
This web service aligns ('reconciles') file names/paths of Wikimedia Commons files with their corresponding **MediaInfo entity URIs** (M-ids). Via data extraction, the Wikitext, multilingual file captions and other **structured data** of these files can then be retrieved for further processing in OpenRefine.

*This interface works with OpenRefine from 2.6 rc2 onwards. It is not compatible with Google Refine.*

Use the following URL in OpenRefine: <https://commonsreconcile.toolforge.org/en/api>

Replacing **en** with another language code will display items and properties in that language, if they are available.

## How do I use this service in OpenRefine?



The screenshot shows the OpenRefine interface with a reconciliation table. The table has columns for 'All' and 'Title'. The first few rows are:

All	Title
1	File:MusArt Gallery - Mother with child.jpg
2	File:Commons:Barnemidde Musée 01.JPG
3	File:Commons:Barnemidde Musée 02.JPG
4	File:Commons:Barnemidde Musée 03.JPG
5	File:Commons:Gastand Schindlerwerkstatt Linden-Museum.jpg
6	File:Point du Platin.jpg
7	File:Blacklode Museum.jpg
8	File:Blacklode Museum board.jpg
9	File:BM gallery3.jpg
10	File:BM gallery 2.jpg
11	File:BM gallery.jpg
12	File:A Boys Soccer "Kicker".jpg
13	File:More objects at MusArt Gallery.jpg
14	File:MusArt Gallery sign post.jpg
15	File:Some of the objects at MusArt Gallery.jpg
16	File:The Flying Monster by John Yarnes Muck.jpg

Demo video in English, 3'25"

## Reconciling

You can find [general documentation](#) about reconciling in OpenRefine's manual.

```
Onbewerkte gegevens Headers
Kopiëren Alles samenvouwen Alles uitvouwen JSON filteren
send:
propose_properties:
  service_path: "/properties"
  service_url: "https://commonsreconcile.toolforge.org/en/api"
  identifierSpace: "https://commons.wikimedia.org/entity/"
  namespace: "Wikimedia Commons (en)"
  preview:
  height: 60
  url: "https://commonsreconcile.toolforge.org/en/api/preview?id={{id}}"
  width: 500
  namespace: "http://www.wikidata.org/prop/direct/"
  suggest:
  entity:
  service_path: "/suggest"
  service_url: "https://commonsreconcile.toolforge.org/en/api"
  property:
  service_path: "/suggest/properties"
  service_url: "https://commonsreconcile.toolforge.org/en/api"
  versions:
  0: "0.2"
  preview:
  url: "https://commons.wikimedia.org/entity/{{id}}"
```

<https://commonsreconcile.toolforge.org>

<https://commonsreconcile.toolforge.org/en/api>

# Reconciliation service for Rhizome Artbase

## Rhizome Artbase reconciliation for OpenRefine

This web service can be used to align datasets to Rhizome Artbase items in

Use the following URL in OpenRefine: <https://reconciliation.artbase.rhizome.org/en/api/>

Replacing "en" by another language code will display items and properties in your language, when they are available

This interface works with OpenRefine from 2.6 rc2 onwards. It is not compatible with Google Refine.

- [Documentation of the API](#)
- [Source](#)
- [Bug tracker](#)
- [Service status](#)

## Tutorials for OpenRefine users

- [Official documentation in English](#)
- [Official documentation in Japanese](#) by 山本 泰智
- [Tutorial by Antoine Courtin](#) (in French)
- [Comparison of reconciliation services for OpenRefine](#) by Aaron Tay (in English)
- [Using OpenRefine to Reconcile Name Entities](#) by Karen H. (in English)
- [Reconciling in OpenRefine - eine Einführung von histHub](#) (in German)

## Videos

- [Reconciliation in OpenRefine](#), screencast by Owen Stephens (in English, 24:06)
- [OpenRefine demo at WikidataCon 2017](#), live demo by Arthur P. Smith (in English, 19:52)
- [Exemple de réconciliation et d'enrichissement avec Wikidata](#) (silent screencast, 3:27) by Ettore Rizza
- [Wikidata editing with OpenRefine](#), screencast by Antonin Delpeuch (in English, 3 videos of 5 minutes each)

## Acknowledgements

- The original reconciliation service was written by Magnus Manske
- The service uses [autodescribe](#) by Magnus Manske
- This work was supported by a [bounty on OpenRefine](#)

```
JSON  Onbewerkte gegevens  Headers
Opslaan  Kopiëren  Alles samenvouwen  Alles uitvouwen  JSON filteren
defaultTypes:  []
extend:
  property_settings:
    0:
      default:  0
      help_text:  "Maximum number of values to return per row (0 for no limit)"
      label:  "Limit"
      name:  "limit"
      type:  "number"
    1:
      choices:
        0:
          name:  "Any rank"
          value:  "any"
        1:
          name:  "Only the best rank"
          value:  "best"
        2:
          name:  "Preferred and normal ranks"
          value:  "no_deprecated"
      default:  "best"
      help_text:  "Filter statements by rank"
      label:  "Ranks"
      name:  "rank"
      type:  "select"
    2:
      choices:
        0:
          name:  "Any statement"
          value:  "any"
        1:
          name:  "At least one reference"
          value:  "referenced"
        2:
          name:  "At least one non-wiki reference"
          value:  "no_wiki"
      default:  "any"
      help_text:  "Filter statements by their references"
      label:  "References"
```

<https://reconciliation.artbase.rhizome.org/>

<https://reconciliation.artbase.rhizome.org/en/api>

# Reconciliation service for EU Knowledge Graph

## EU Knowledge Graph reconciliation for OpenRefine

This web service can be used to align datasets to EU Knowledge Graph items

Use the following URL in OpenRefine: <https://openrefine-reconciliation.linkedopendata.eu/en/api>.

Replacing "en" by another language code will display items and properties in your language, when they are available

This interface works with OpenRefine from 2.6 rc2 onwards. It is not compatible with Google Refine.

- [Documentation of the protocol](#)
- [Documentation of this application for developers and Wikibase admins](#)
- [Source](#)
- [Bug tracker](#)
- [Service status](#)

## Tutorials for OpenRefine users

- [Official documentation in English](#)
- [Official documentation in Japanese](#) by 山本 泰智
- [Tutorial by Antoine Courtin](#) (in French)
- [Comparison of reconciliation services for OpenRefine](#) by Aaron Tay (in English)
- [Using OpenRefine to Reconcile Name Entities](#) by Karen H. (in English)
- [Reconciling in OpenRefine - eine Einführung von histHub](#) (in German)

## Videos

- [Reconciliation in OpenRefine](#), screencast by Owen Stephens (in English, 24:06)
- [OpenRefine demo at WikidataCon 2017](#), live demo by Arthur P. Smith (in English, 19:52)
- [Exemple de réconciliation et d'enrichissement avec Wikidata](#) (silent screencast, 3:27) by Ettore Rizza
- [Wikidata editing with OpenRefine](#), screencast by Antonin Delpeuch (in English, 3 videos of 5 minutes each)

## Documentation

Contents:

























- [Reconciling via unique identifiers](#)
- [Property paths](#)
- [Special properties](#)
- [Subfields](#)


```
JSON  Onbewerkte gegevens  Headers
Opslaan  Kopiëren  Alles samenvouwen  Alles uitvouwen  JSON filteren
▼ defaultTypes:
  ▼ 0:
    id: "Q236932"
    name: "entity"
  ▼ extend:
    ▼ property_settings:
      ▼ 0:
        default: 0
        ▼ help_text: "Maximum number of values to return per row (0 for no limit)"
        label: "Limit"
        name: "limit"
        type: "number"
      ▼ 1:
        ▼ choices:
          ▼ 0:
            name: "Any rank"
            value: "any"
          ▼ 1:
            name: "Only the best rank"
            value: "best"
          ▼ 2:
            name: "Preferred and normal ranks"
            value: "no_deprecated"
        default: "best"
        help_text: "Filter statements by rank"
        label: "Ranks"
        name: "rank"
        type: "select"
      ▼ 2:
        ▼ choices:
          ▼ 0:
            name: "Any statement"
            value: "any"
          ▼ 1:
            name: "At least one reference"
```

<https://openrefine-reconciliation.linkedopendata.eu>

<https://openrefine-reconciliation.linkedopendata.eu/en/api>

# Reconciliation services overview


Wikimedia Commons (docs, source)		 <a href="https://commonsreconcile.toolforge.org/en/api">https://commonsreconcile.toolforge.org/en/api</a>
OpenSanctions (docs)		 <a href="https://api.opensanctions.org/reconcile/default">https://api.opensanctions.org/reconcile/default</a>
GeoNames (source)		 <a href="https://api.exldevnetwork.net/geonames-openrefine/reconcile">https://api.exldevnetwork.net/geonames-openrefine/reconcile</a>
Bionomia ID (docs)		 <a href="https://api.bionomia.net/reconcile">https://api.bionomia.net/reconcile</a>
Artsdata.ca (docs, source)		 <a href="https://api.artsdata.ca/recon">https://api.artsdata.ca/recon</a>
Organized Crime and Corruption Reporting Project (source)		 <a href="https://aleph.occrp.org/api/freebase/reconcile">https://aleph.occrp.org/api/freebase/reconcile</a>
Bahamas Leaks (docs, source)		 <a href="https://aleph.occrp.org/api/2/collections/276/reconcile">https://aleph.occrp.org/api/2/collections/276/reconcile</a>
ORCID (docs, source)		 <a href="http://refine.codefork.com/reconcile/orcid">http://refine.codefork.com/reconcile/orcid</a>
Open Library (docs, source)		 <a href="http://refine.codefork.com/reconcile/openlibrary">http://refine.codefork.com/reconcile/openlibrary</a>
Funder Registry (docs, source)		 <a href="http://recon.labs.crossref.org/reconcile">http://recon.labs.crossref.org/reconcile</a>
OpenUp common names (docs)		 <a href="http://openup.nhm-wien.ac.at/commonNames/">http://openup.nhm-wien.ac.at/commonNames/</a>
kerameikos.org (docs, source)		 <a href="http://kerameikos.org/apis/reconcile">http://kerameikos.org/apis/reconcile</a>

 Add a service

 Refresh table

Muziekweb: genres and styles (docs)



 <https://termennetwerk-api.netwerkdomein.nl/Muziekweb/sparql/Muziekweb%23mw-genresstijlen>

<https://reconciliation-api.github.io/testbench/#/>

# Making a recon service for your own Wikibase and use it in OpenRefine in 3 easy steps



# 1) Install the recon service for your Wikibase

This repository has been archived by the owner. It is now read-only.

wetneb / openrefine-wikibase Public archive Watch 21 Fork 23 Star 97

[Code](#) [Issues 25](#) [Pull requests](#) [Actions](#) [Security](#) [Insights](#)

master 8 branches 2 tags Go to file Code

wetneb Point to new repository ✓ 4589b69 11 days ago 287 commits

.github/workflows	Change CI to use pytest. Closes #103. (#104)	2 years ago
docs	Further fixes to install instructions	3 months ago
static	Use local version of Bootstrap CSS	2 years ago
templates	Add link to Wikidata SiteLinks ID list (#144)	3 months ago
tests	Return descriptions in recon responses, announce supported versio...	14 months ago
wdreconcile	Fix property proposals for data extension. Closes #145	2 months ago
.coveragerc	Fix continuous integration (#102)	2 years ago
.gitignore	Change CI to use pytest. Closes #103. (#104)	2 years ago
Dockerfile	Fix docker config	2 years ago
LICENSE	README and LICENSE	6 years ago
README.md	Point to new repository	11 days ago
app.py	Migrate to aioredis 2.0.0 and pin dependency versions	13 months ago
config_docker.py	Update config_docker.py	2 years ago
config_wikidata.py	Ignore autodesc failures in previews	2 years ago
conftest.py	Migrate to aioredis 2.0.0 and pin dependency versions	13 months ago
docker-compose.yml	Changed docker-compose file volumes so	
requirements.txt	set jinja2==3.0.3 (#138)	

About

This repository has migrated to:  
[gitlab.com/nfdi4culture/ta1-data-enric...](https://gitlab.com/nfdi4culture/ta1-data-enric...)

[openrefine](#) [reconciliation](#) [wikidata](#)  
[wikibase](#) [reconciliation-service](#)  
[reconciliation-interface](#)

Readme  
View license  
97 stars  
21 watching  
23 forks

Releases  
2 tags

Packages  
No packages published

Contributors 9

<https://github.com/wetneb/openrefine-wikibase>

# 1) Install the recon service for your Wikibase

## CONTENTS:

- Installing the reconciliation service
- Architecture overview
- Scoring mechanism
- Testing infrastructure
- Documenting

## Welcome to the documentation of the OpenRefine-Wikibase reconciliation interface

This software offers a reconciliation interface for a Wikibase instance, following the [specifications of the reconciliation API](#).

This manual is intended for developers and Wikibase administrators. Reconciliation users should instead refer to the main page of the corresponding reconciliation endpoint, where user documentation can be found. For Wikidata, this is available at <https://wikidata.reconci.link/>.

### Contents:

- [Installing the reconciliation service](#)
  - [Requirements](#)
  - [Configuration](#)
  - [Installing with Docker](#)
  - [Installing manually](#)
  - [Deploying in production](#)
  - [Tips about Redis configuration](#)
- [Architecture overview](#)
  - [Reconciliation](#)
  - [Auto-complete \(suggest\) services](#)
  - [Preview](#)
  - [Data extension](#)

<https://openrefine-wikibase.readthedocs.io/en/latest/>

# 1) Install the recon service for your Wikibase

## Installing the reconciliation service

### Requirements

The Wikibase instance should have:

- An associated SPARQL query service;
- Some special properties and items to represent its type system, by analogy to the one in place in Wikidata with [instance of \(P31\)](#) and [subclass of \(P279\)](#), with a root type such as [entity \(Q35120\)](#);

In addition it is also recommended that the Wikibase instance uses the [CirrusSearch extension](#) (ElasticSearch-based search engine).

### Configuration

The configuration of the service is done in a Python file *config.py*. A sample configuration file is provided for Wikidata, *config\_wikidata.py*.

### Installing with Docker

You can run this service with docker. First, clone the repository and go to its root directory:

```
git clone https://github.com/wetneb/openrefine-wikibase
cd openrefine-wikibase
```

Then, copy the sample *config\_docker.py* to *config.py* and modify the copy to point to the Wikibase instance of your choice.

Finally, start the service:

```
docker-compose up
```

On Windows you will need to accept the Windows Firewall the service runs.

<https://openrefine-wikibase.readthedocs.io/en/latest/install.html>



# 1) Install the recon service for your Wikibase

## Installing the reconciliation service

### Requirements

The Wikibase instance should have:

- An associated SPARQL query service;
- Some special properties and items to represent its type system, by analogy to the one in place in Wikidata with [instance of \(P31\)](#) and [subclass of \(P279\)](#), with a root type such as [entity \(Q35120\)](#);

In addition it is also recommended that the Wikibase instance uses the [CirrusSearch extension](#) (ElasticSearch-based search engine).

### Configuration

The configuration of the service is done in a Python file *config.py*. A sample configuration file is provided for Wikidata, *config\_wikidata.py*.

### Installing with Docker

You can run this service with docker. First, clone the repository and go to its root directory:

```
git clone https://github.com/wetneb/openrefine-wikibase
cd openrefine-wikibase
```

Then, copy the sample *config\_docker.py* to *config.py* and modify the copy to point to the Wikibase instance of your choice.

Finally, start the service:

```
docker-compose up
```

On Windows you will need to accept the Windows Firewall the service runs.

<https://openrefine-wikibase.readthedocs.io/en/latest/install.html>

160 lines (123 sloc) | 6 KB

Raw Blame

*config.py* for

<https://kbttestwikibase.wikibase.cloud>

```
1
2 """
3 This file defines a few constants which configure
4 which Wikibase instance and which property/item ids
5 should be used
6 """
7
8 # Endpoint of the Mediawiki API of the Wikibase instance
9 mediawiki_api_endpoint = 'https://kbttestwikibase.wikibase.cloud/w/api.php'
10
11 # SPARQL endpoint
12 wikibase_sparql_endpoint = 'https://kbttestwikibase.wikibase.cloud/query/'
13 # Name of the Wikibase instance
14 wikibase_name = 'KBTestWikibase.cloud'
15
16 # URL of the main page of the Wikibase instance
17 wikibase_main_page = 'https://kbttestwikibase.wikibase.cloud/wiki/Main_Page'
18
19 # Wikibase namespace ID, used to search for items
20 # For Wikidata this is 0, but most by default Wikibase uses 120, which is the default
21 # CHANGE THIS TO 120 if you are adapting this configuration file to another Wikibase
22 wikibase_namespace_id = 120
23 # https://kbttestwikibase.wikibase.cloud/wiki/Special:AllPages?from=&to=&namespace=120
24
25 # Namespace prefix of Wikibase items (including colon, e.g. 'Item:')
26 wikibase_namespace_prefix = 'Item:'
27
28 # User agent to connect to the Wikidata APIs
29 user_agent = 'OpenRefine-Wikidata reconciliation interface'
30
31 # Regexes and group ids to extracts Qids and Pids from URLs
32 import re
33 q_re = re.compile(r'(<?https?://kbttestwikibase.wikibase.cloud/(entity|wiki)/)?(Q[0-9]+)'
34 q_re_group_id = 3
35 p_re = re.compile(r'(<?https?://kbttestwikibase.wikibase.cloud/(entity|wiki/Property:))'
36 p_re_group_id = 3
37
38 # Identifier space and schema space exposed to OpenRefine.
39 # This should match the IRI prefixes used in RDF serialization. For instance: https://k
40 # Note that you should be careful about using http or https there,
41 # because any variation will break comparisons at various places.
42 identifier_space = 'https://kbttestwikibase.wikibase.cloud/entity/'
43 schema_space = 'https://kbttestwikibase.wikibase.cloud/prop/direct/'
44
45
46 # Pattern used to form the URL of
47 # This is only used for viewing so it is fine to use any protocol (therefore, preferably HTTPS if supported)
48 qid_url_pattern = 'https://kbttestwikibase.wikibase.cloud/wiki/Item:{{id}}'
```

```
58 # Service name exposed at various places,
59 # mainly in the list of reconciliation services of users
60 service_name = 'KBTestWikibase.cloud recon service'
61
62 # URL (without the trailing slash) where this server runs
63 this_host = 'http://localhost:8000'
64
65 # The default limit on the number of results returned by us
66 default_num_results = 25
67
68 # The maximum number of
69 wd_api_max_search_result
70
71 # The matching score abo
72 validation_threshold = 95
73
74 # Redis client used for caching at various places
75 redis_uri = 'redis://redis:6379/0?encoding=utf-8'
76
77 # Redis prefix to use in front of all keys
78 redis_key_prefix = 'openrefine_wikidata:'
```

Recon service will run at this URL

<https://github.com/KBNLresearch/OpenRefine-Wikibase/blob/main/kb-test-wikibase-cloud-config.py>

# 1) Install the recon service for your Wikibase

## Installing the reconciliation service

### Requirements

The Wikibase instance should have:

- An associated SPARQL query service;
- Some special properties and items to represent its type system, by analogy to the one in place in Wikidata with [instance of \(P31\)](#) and [subclass of \(P279\)](#), with a root type such as [entity \(Q35120\)](#);

In addition it is also recommended that the Wikibase instance uses the [CirrusSearch extension](#) (ElasticSearch-based search engine).

### Configuration

The configuration of the service is done in a Python file *config.py*. A sample configuration file is provided for Wikidata, *config\_wikidata.py*.

### Installing with Docker

You can run this service with docker. First, clone the repository and go to its root directory:

```
git clone https://github.com/wetneb/openrefine-wikibase
cd openrefine-wikibase
```

Then, copy the sample *config\_docker.py* to the *config.py* file of your Wikibase instance of your choice.

Finally, start the service:

```
docker-compose up
```

Start the recon  
service

On Windows you will need to accept the Windows Firewall the service runs.

<https://openrefine-wikibase.readthedocs.io/en/latest/install.html>

# Reconciliation service for Kunstmuseum Den Haag

## API kunstmuseum reconciliation for OpenRefine

This web service can be used to align datasets to API kunstmuseum items in OpenRefine

Use the following URL in OpenRefine: <https://reconciliation.kunstmuseum.nl/en/api>.

Replacing "en" by another language code will display items and properties in your language, when they are available.

This interface works with OpenRefine from 2.6 rc2 onwards. It is not compatible with Google Refine.

- [Documentation of the protocol](#)
- [Documentation of this application for developers and Wikibase admins](#)
- [Source](#)
- [Bug tracker](#)
- [Service status](#)

## Tutorials for OpenRefine users

- [Official documentation in English](#)
- [Official documentation in Japanese](#) by 山本 泰智
- [Tutorial by Antoine Courtin](#) (in French)
- [Comparison of reconciliation services for OpenRefine](#) by Aaron Tay (in English)
- [Using OpenRefine to Reconcile Name Entities](#) by Karen H. (in English)
- [Reconciling in OpenRefine - eine Einführung von histHub](#) (in German)

## Videos

- [Reconciliation in OpenRefine](#), screencast by Owen Stephens (in English, 24:06)
- [OpenRefine demo at WikidataCon 2017](#), live demo by Arthur P. Smith (in English, 19:52)
- [Exemple de réconciliation et d'enrichissement avec Wikidata](#) (silent screencast, 3:27) by Ettore Rizza
- [Wikidata editing with OpenRefine](#), screencast by Antonin Delpeuch (in English, 3 videos of 5 minutes each)

## Documentation

Contents:

- [Reconciling via unique identifiers](#)
- [Property paths](#)
- [Special properties](#)
- [Subfields](#)

[Reconciling via unique identifiers](#)

```
JSON  Onbewerkte gegevens  Headers
Opslaan  Kopiëren  Alles samenvouwen  Alles uitvouwen  JSON filteren
▼ defaultTypes:
  ▼ 0:
    id: "Q2667"
    name: "entiteit"
  ▼ extend:
    ▼ property_settings:
      ▼ 0:
        default: 0
        ▼ help_text: "Maximum number of values to return per row (0 for no limit)"
        label: "Limit"
        name: "limit"
        type: "number"
      ▼ 1:
        ▼ choices:
          ▼ 0:
            name: "Any rank"
            value: "any"
          ▼ 1:
            name: "Only the best rank"
            value: "best"
          ▼ 2:
            name: "Preferred and normal ranks"
            value: "no_deprecated"
        default: "best"
        help_text: "Filter statements by rank"
        label: "Ranks"
        name: "rank"
        type: "select"
      ▼ 2:
        ▼ choices:
          ▼ 0:
            name: "Any statement"
            value: "any"
          ▼ 1:
            name: "At least one reference"
            value: "referenced"
          ▼ 2:
            name: "At least one non-wiki reference"
            value: "no_wiki"
        default: "any"
        help_text: "Filter statements by their references"
        label: "References"
```

<https://reconciliation.kunstmuseum.nl>

<https://reconciliation.kunstmuseum.nl/nl/api>



Facet / Filter

Undo / Redo 1 / 1



19 rows

Show as: **rows** records Show: 5 10 25

All  KunstmuseumPersoon

<input type="checkbox"/>	<input type="checkbox"/>	1.	Vincent van Gogh
<input type="checkbox"/>	<input type="checkbox"/>	2.	Adrianus Kocx
<input type="checkbox"/>	<input type="checkbox"/>	3.	Lambertus van Eenhoorn
<input type="checkbox"/>	<input type="checkbox"/>	4.	Susanne Klüver
<input type="checkbox"/>	<input type="checkbox"/>	5.	Marion van Aken-Fehmers
<input type="checkbox"/>	<input type="checkbox"/>	6.	Suzanne Lambooy
<input type="checkbox"/>	<input type="checkbox"/>	7.	Titus Eliëns
<input type="checkbox"/>	<input type="checkbox"/>	8.	Ida Peelen
<input type="checkbox"/>	<input type="checkbox"/>	9.	Eline van den Berg
<input type="checkbox"/>	<input type="checkbox"/>	10.	August I de Sterke
<input type="checkbox"/>	<input type="checkbox"/>	11.	Chenghua
<input type="checkbox"/>	<input type="checkbox"/>	12.	Daoguang
<input type="checkbox"/>	<input type="checkbox"/>	13.	Guangxu
<input type="checkbox"/>	<input type="checkbox"/>	14.	Hongwu
<input type="checkbox"/>	<input type="checkbox"/>	15.	Hongxi
<input type="checkbox"/>	<input type="checkbox"/>	16.	Hongzhi
<input type="checkbox"/>	<input type="checkbox"/>	17.	Jiajing
<input type="checkbox"/>	<input type="checkbox"/>	18.	Jianwen
<input type="checkbox"/>	<input type="checkbox"/>	19.	Jiaqing

### Using facets and filters



Use facets and filters to select subsets of your data to act on. Choose facet and filter methods from the menus at the top of each data column.

Not sure how to get started?

[Watch these screencasts](#)

aders

Alles uitvouden  JSON filteren

number of values to return per row (0 for no limit)"

best rank"

ed and normal ranks"

ecated"

statements by rank"

statement"

one reference"

ced"

one non-wiki reference"

statements by their references"

ca"

### Reconcile column "KunstmuseumPersoon"

#### Services

- Nederlandse Thesaurus van Auteursnamen
- RKDartists
- VIAF
- Wikidata reconcil.link (en)
- Wikidata reconcil.link (nl)
- Wikimedia Commons (en)
- Wikibase Kunstmuseum (nl)

Pick a service or extension on left

<https://reconciliation.kunstmuseum.nl/nl/api>

Add standard service...

Discover services...

Start reconciling...

Cancel

Reconcile column "KunstmuseumPersoon"

Services

Nederlandse Thesaurus van Auteursnamen

RKDartists

VIAF

Wikidata reconcil.link (en)

Wikidata reconcil.link (nl)

Wikimedia Commons (en)

Wikibase Kunstmuseum (nl)

OpenRefine Clipboard [Permalink](#)

Facet / Filter Undo / Redo 6 / 6 < 19 rows

Refresh Reset all Remove all

KunstmuseumPersoon: judgment change invert reset

1 choice Sort by: name count

matched 19 exclude

Facet by choice counts

▼ All ▼ KunstmuseumPersoon

Show as: rows records Show: 5 10 25 50

☆	🗨	1.	Vincent van Gogh <a href="#">Choose new match</a>
☆	🗨	2.	Adrianus Kocx <a href="#">Choose new match</a>
☆	🗨	3.	Lambertus van Eenhoorn <a href="#">Choose new match</a>
☆	🗨	4.	Susanne Klüver <a href="#">Choose new match</a>
☆	🗨	5.	Marion van Aken-Fehmers <a href="#">Choose new match</a>
☆	🗨	6.	Suzanne Lambooy <a href="#">Choose new match</a>
☆	🗨	7.	Titus Eliëns <a href="#">Choose new match</a>
☆	🗨	8.	Ida Peelen <a href="#">Choose new match</a>
☆	🗨	9.	Eline van den Berg <a href="#">Choose new match</a>
☆	🗨	10.	August I de Sterke <a href="#">Choose new match</a>
☆	🗨	11.	Chenghua <a href="#">Choose new match</a>
☆	🗨	12.	Daoguang <a href="#">Choose new match</a>
☆	🗨	13.	Guangxu <a href="#">Choose new match</a>



Add standard service...

Cancel Reconciling...

Cancel

# 2) Connect OpenRefine to your Wikibase instance

The screenshot shows the OpenRefine documentation website. The top navigation bar includes 'OpenRefine Documentation', 'User Manual', and 'Technical Reference'. A left sidebar lists various manual sections, with 'Connecting to Wikibase' highlighted. The main content area shows the breadcrumb 'User Manual > Wikibase, Wikidata and Wikimedia Commons > Connecting to Wikibase', a version tag 'Version: 3.6', and the main title 'Connecting OpenRefine to a Wikibase instance'. The text explains that the page covers connecting to any Wikibase instance, with a note that Wikidata is pre-configured. It is divided into three sections: 'For Wikibase end users', 'For Wikibase administrators', and 'Requirements'. In the 'For Wikibase end users' section, the word 'manifest' is highlighted with a red box. The 'Requirements' section mentions the need for a reconciliation service and a Python wrapper.

[OpenRefine Documentation](#) User Manual Technical Reference

User Manual > Wikibase, Wikidata and Wikimedia Commons > [Connecting to Wikibase](#)

Version: 3.6

## Connecting OpenRefine to a Wikibase instance

This page explains how to connect OpenRefine to any Wikibase instance. If you just want to work with [Wikidata](#), you can ignore this page as Wikidata is configured out of the box in OpenRefine.

### For Wikibase end users

All you need to configure OpenRefine to work with a Wikibase instance is a **manifest** for that instance, which provides some metadata and links required for the integration to work.

We offer some off-the-shelf manifests for some public Wikibase instances in the [wikibase-manifests](#) repository. But the administrators of your Wikibase instance should provide one that is potentially more up to date, so it makes sense to request it to them first.

### For Wikibase administrators

To let your users contribute to your Wikibase instance with OpenRefine, you will need to write a manifest as described above. There is currently no canonical location where this manifest should be hosted - just make sure can be found easily by your users. This section explains the format of the manifest.

### Requirements

To work with OpenRefine, your Wikibase instance needs an associated reconciliation service. For instance you can use a [Python wrapper](#) for this. Also, in addition to Wikibase, the [UniversalLanguageSelector extension](#) should be installed.

<https://docs.openrefine.org>

<https://docs.openrefine.org/manual/wikibase/configuration>



## 2) Connect OpenRefine to your Wikibase instance

The screenshot shows the GitHub repository page for 'OpenRefine/wikibase-manifests'. The repository is public and has 4 watchers, 8 forks, and 13 stars. The main navigation includes Code, Issues, Pull requests (2), Actions, Projects, Security, and Insights. The repository is on the 'master' branch with 3 branches and 0 tags. A list of files is shown, with 'wikimedia-commons-manifest.json' highlighted by a red box. The README file is open, showing the title 'Wikibase manifests' and the text: 'This GitHub repository stores manifests of some public Wikibase instances. By entering a manifest in OpenRefine's Wikibase selection dialog window ( Select Wikibase instance... in the Wikidata/Wikibase extension menu), you can start (batch) editing that particular Wikibase.'

File Name	Description	Commit Date
LICENSE.md	Create LICENSE.md	12 months ago
README.md	Update README.md	8 months ago
biblissima-data-manifest-v1.json	Rename Biblissima Manifest (v1) and add Manifest v2 (#9)	
biblissima-data-manifest.json	Rename Biblissima Manifest (v1) and add Manifest v2 (#9)	
factgrid-manifest.json	Add FactGrid manifest	
openrefine-wikibase-test-manifest.j...	init	2 years ago
rhizome-artbase-manifest.json	Update reconciliation endpoint for Rhizome Artbase	5 months ago
wikibase-manifest-schema-v1.json	Add tag and max_edits_per_minute fields	9 months ago
wikibase-manifest-schema-v2.json	Add new field 'hide_structured_fields_in_mediainfo'	2 months ago
wikidata-manifest.json	include editgroups in the manifest	2 years ago
wikimedia-commons-manifest.json	Revert "Add all Wikidata constraint settings to Commons too, why n...	last month

**Wikibase example manifests**

### Wikibase manifests

This GitHub repository stores manifests of some public Wikibase instances.

By entering a manifest in OpenRefine's Wikibase selection dialog window ( Select Wikibase instance... in the Wikidata/Wikibase extension menu), you can start (batch) editing that particular Wikibase.

<https://github.com/OpenRefine/wikibase-manifests>

2)

789 lines (789 sloc) | 24.4 KB

```
1 {
2   "version": "2.0",
3   "mediawiki": {
4     "name": "Wikimedia Commons",
5     "root": "https://commons.wikimedia.org/wiki/",
6     "main_page": "https://commons.wikimedia.org/wiki/Main_Page",
7     "api": "https://commons.wikimedia.org/w/api.php"
8   },
9   "wikibase": {
10    "site_iri": "https://commons.wikimedia.org/entity/",
11    "maxlag": 5,
12    "properties": {
13      "instance_of": "P31",
14      "subclass_of": "P279"
15    },
16    "constraints": {
17      "property_constraint_pid": "P2302",
18      "exception_to_constraint_pid": "P2303",
19      "constraint_status_pid": "P2316",
20      "mandatory_constraint_qid": "Q21502408",
21      "suggestion_constraint_qid": "Q62026391",
22      "distinct_values_constraint_qid": "Q21502410"
23    }
24  },
25  "oauth": {
26    "registration_page": "https://commons.wikimedia.org/wiki/Special:OAuthConsumerRegistration/propose"
27  },
28  "entity_types": {
29    "item": {
30      "site_iri": "http://www.wikidata.org/entity/",
31      "reconciliation_endpoint": "https://wikidata.reconci.link/${lang}/api",
32      "mediawiki_api": "https://www.wikidata.org/w/api.php"
33    },
34    "property": {
35      "site_iri": "http://www.wikidata.org/entity/",
36      "mediawiki_api": "https://www.wikidata.org/w/api.php"
37    },
38    "mediainfo": {
39      "site_iri": "https://commons.wikimedia.org/entity/",
40      "reconciliation_endpoint": "https://commonsreconcile.toolforge.org/${lang}/api"
```

## Wikimedia Commons manifest

instance

Fork 8 Star 13

About

Manifests v2 are  
3.6 and later

Readme

CC0-1.0 license

13 stars

4 watching

8 forks

Releases

No releases published

Packages

No packages published

Contributors 7



<https://github.com/OpenRefine/wikibase-manifests/blob/master/wikimedia-commons-manifest.json>

2)

789 lines

OpenRefine / OpenRefine Public

Search Watch 478

<> Code Issues 6

## Write a Wikibase manifest

Lu Liu edited this page on Apr 10, 2015

This is a tutorial on how to write a Wikibase manifest.

### Before you start

Before you start writing a manifest for the Wikibase.

If there is no manifest for the Wikibase.

### The format of the manifest

Here is the manifest of Wikibase:

```

{
  "version": "1.0",
  "mediawiki": {
    "name": "Wikidata",
    "root": "https://www.wikidata.org/wiki/Wikidata:Main_Page",
    "main_page": "https://www.wikidata.org/wiki/Wikidata:Main_Page",
    "api": "https://www.wikidata.org/wiki/Wikidata:API",
  },
  "wikibase": {
    "site_iri": "http://www.wikidata.org/entity/",
    "maxlag": 5,
    "properties": {
      "instance_of": "http://www.wikidata.org/entity/Q5887",
      "subclass_of": "http://www.wikidata.org/entity/Q5887",
    },
    "constraints": {
      "property_constraint": "http://www.wikidata.org/entity/Q5887",
      "exception_to_constraint": "http://www.wikidata.org/entity/Q5887",
      "mandatory_constraint": "http://www.wikidata.org/entity/Q5887",
      "suggestion_constraint": "http://www.wikidata.org/entity/Q5887",
    }
  }
}

```

### version

The version should be in the format "1.x". The minor version should be in a backward-compatible manner. The major version should be "1" if the Wikibase is in a beta or alpha state. See [v1.json](#).

### mediawiki

This part contains some basic information of the Wikibase.

#### name

The name of the Wikibase, should be unique for different Wikibase instances.

#### root

The root of the Wikibase. Typically in the form "[https://foo.bar/wiki/](#)". The trailing slash cannot be omitted.

#### main\_page

The main page of the Wikibase. Typically in the form "[https://foo.bar/wiki/Main\\_Page](#)".

#### api

The MediaWiki API endpoint of the Wikibase. Typically in the form "[https://foo.bar/w/api.php](#)".

### wikibase

This part contains configurations of the Wikibase extension.

#### site\_iri

The IRI of the Wikibase, in the form '[http://foo.bar/entity/](#)'. This should match the IRI prefixes used in RDF serialization. Be careful about using "http" or "https", because any variation will break comparisons at various places. The trailing slash cannot be omitted.

#### maxlag

Maxlag is a parameter that controls how aggressive a mass-editing tool should be when uploading edits to a Wikibase instance. See [https://www.mediawiki.org/wiki/Manual:Maxlag\\_parameter](https://www.mediawiki.org/wiki/Manual:Maxlag_parameter) for more details. The value should be adapted according to the actual traffic of the Wikibase.

Write a Wikibase manifest

<https://github.com/OpenRefine/OpenRefine/wiki/Write-a-Wikibase-manifest>

# Manifest.json for Kunstmuseum Den Haag

## Openrefine

Het reconciliation-endpoint kan in Openrefine worden ingevoerd om te koppelen. Om items vervolgens te uploaden is een account en een manifest nodig. Onderstaand is te gebruiken als manifest:

```
{
  "version": "2.0",
  "mediawiki": {
    "name": "Kunstmuseum API",
    "root": "https://api.kunstmuseum.nl/wiki/@",
    "main_page": "https://api.kunstmuseum.nl/wiki/Kunstmuseum_API@",
    "api": "https://api.kunstmuseum.nl/w/api.php@"
  },
  "wikibase": {
    "site_iri": "https://api.kunstmuseum.nl/entity/@",
    "maxlag": 5,
    "max_edits_per_minute": 240,
    "tag": "openrefine-${version}",
    "properties": {
      "instance_of": "P4",
      "subclass_of": "P47"
    }
  },
  "oauth": {
    "registration_page": "https://api.kunstmuseum.nl/wiki/Speciaal:OAuthGebruikerRegistratie/propose@"
  },
  "entity_types": {
    "item": {
      "site_iri": "https://api.kunstmuseum.nl/entity/@",
      "reconciliation_endpoint": "https://reconciliation.kunstmuseum.nl/${lang}/api@",
      "mediawiki_api": "https://api.kunstmuseum.nl/w/api.php@"
    },
    "property": {
      "site_iri": "https://api.kunstmuseum.nl/entity/@",
      "mediawiki_api": "https://api.kunstmuseum.nl/w/api.php@"
    }
  },
  "hide_structured_fields_in_mediainfo": true
}
```

[https://api.kunstmuseum.nl/wiki/Kunstmuseum\\_API#Openrefine](https://api.kunstmuseum.nl/wiki/Kunstmuseum_API#Openrefine)

# Manifest.json for Kunstmuseum Den Haag

## Openrefine

Het reconciliation-endpoint kan in Openrefine worden ingevoerd om te koppelen. Om items vervolgens te uploaden is een account en een manifest nodig. Onderstaand is te gebruiken als manifest.

19 rows

Extensions Wikibase ▾

Show as: **rows** records Show: 5 10 **25** 50 100 500 1000 rows

« first

▾ All ▾ **KunstmuseumPersoon**

☆	🗨	1.	<a href="#">Vincent van Gogh</a> <small>Choose new match</small>
☆	🗨	2.	<a href="#">Adrianus Kocx</a> <small>Choose new match</small>
☆	🗨	3.	<a href="#">Lambertus van Eenhoorn</a> <small>Choose new match</small>
☆	🗨	4.	<a href="#">Susanne Klüver</a> <small>Choose new match</small>
☆	🗨	5.	<a href="#">Marion van Aken-Fehmers</a> <small>Choose new match</small>
☆	🗨	6.	<a href="#">Suzanne Lambooy</a> <small>Choose new match</small>
☆	🗨	7.	<a href="#">Titus Eliëns</a> <small>Choose new match</small>
☆	🗨	8.	<a href="#">Ida Peelen</a> <small>Choose new match</small>
☆	🗨	9.	<a href="#">Eline van den Berg</a> <small>Choose new match</small>
☆	🗨	10.	<a href="#">August I de Sterke</a> <small>Choose new match</small>
☆	🗨	11.	<a href="#">Chenghua</a> <small>Choose new match</small>

Manage Wikibase instances...

Edit Wikibase schema

Manage schemas...

Manage Wikibase account...

Upload edits to Wikibase...

Export to QuickStatements

# Manifest.json for Kunstmuseum Den Haag

## Openrefine

Het reconciliation-endpoint kan in Openrefine worden ingevoerd om te koppelen. Om items vervolgens te uploaden is een account en een manifest nodig. Onderstaand is te gebruiken als manifest.

19 rows

Extensions Wikibase ▾

Show as: rows records Show: 5 10 25 50 100 500 1000 rows

« first

Manage Wikibase instances...

Edit Wikibase schema

▼ All

### Manage Wikibase instances

Click on an item below to select it as the target Wikibase to work against. This will clear any existing schema. After switching to the target Wikibase, you should reconcile your data against it before editing the schema.



Wikidata

<https://www.wikidata.org/wiki/>

active



Wikimedia Commons

<https://commons.wikimedia.org/wiki/>

Delete

Add Wikibase

Discover Wikibase instances...

Cancel

OK

# Manifest.json for Kunstmuseum Den Haag

## Add Wikibase manifest

The manifest should specify a reconciliation service linked to the Wikibase, the reconciliation service will be added to OpenRefine if not added yet.

Enter the Wikibase manifest's URL (recommended, this is helpful for keeping track of the latest manifest):

Or paste the manifest JSON text directly (manifests for some public Wikibases can be found [here](#), you can also write one yourself according to this [tutorial](#)):

```
{
  "version": "2.0",
  "mediawiki": {
    "name": "Kunstmuseum API",
    "root": "https://api.kunstmuseum.nl/wiki/",
    "main_page": "https://api.kunstmuseum.nl/wiki/Kunstmuseum_API",
    "api": "https://api.kunstmuseum.nl/w/api.php"
  },
  "wikibase": {
    "site_iri": "https://api.kunstmuseum.nl/entity/",
    "maxlag": 5,
    "max_edits_per_minute": 240,
    "tag": "openrefine-${version}",
    "properties": {
      "instance_of": "P4",
      "subclass_of": "P47"
    }
  },
  "oauth": {
    "registration_page": "https://api.kunstmuseum.nl/wiki/Special:OAuthGebruikerRegistratie/propose"
  },
  "entity_types": {
    "item": {
      "site_iri": "https://api.kunstmuseum.nl/entity/",
      "reconciliation_endpoint": "https://reconciliation.kunstmuseum.nl/${lang}/api",
      "mediawiki_api": "https://api.kunstmuseum.nl/w/api.php"
    }
  }
}
```

Cancel

Add Wikibase


# Manifest.json for Kunstmuseum Den Haag

Add Wikibase manifest

## Manage Wikibase instances

Click on an item below to select it as the target Wikibase to work against. This will clear any existing schema. After switching to the target Wikibase, you should reconcile your data against it before editing the schema.

**Kunstmuseum API** active  
Kunstmuseum API Logo  
<https://api.kunstmuseum.nl/wiki/>

 Wikidata  
<https://www.wikidata.org/wiki/>

 Wikimedia Commons Delete  
<https://commons.wikimedia.org/wiki/>

Now you can write to the KM Wikibase

Add Wikibase

Discover Wikibase instances...

Cancel

OK

```
"mediawiki_api": "https://api.kunstmuseum.nl/w/api.php"
```

Cancel

Add Wikibase



# Manifest.json for Kunstmuseum Den Haag

Add Wikibase manifest

## Manage Wikibase instances

Click on an item below to select it as the target Wikibase instance.  
After switching to the target Wikibase, you can add a manifest.

(But you'll need an account to be able to write to the KM Wikibase)

### Kunstmuseum API account

Logging in to [Kunstmuseum API](#) lets you upload edits directly from OpenRefine.

You are encouraged to login with [bot passwords](#), see this [wiki page](#) to learn how to get one.

[Kunstmuseum API Logo](#)

Username

Password

Remember me [?](#)

You can also [login with your owner-only consumer](#).

Close

Log in

Add Wikibase

Discover Wikibase instances...

Cancel

OK

```
"mediawiki_api": "https://api.kunstmuseum.nl/w/api.php"
```

Cancel

Add Wikibase



# 3) Configure your Wikibase instance

main ▾ OpenRefine-Wikibase / kb-test-wikibase-cloud-manifest.json

English ▾ OlafJanssen ▾ 15 ▾ 17 Talk Preferences

KB

Special page

## Tags

This page lists the tags that the software may mark an edit with, and their meaning.

### Create a new tag

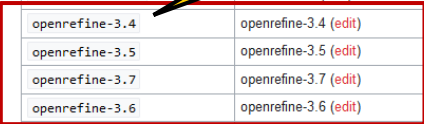
By default, newly created tags will be made available for use by users and bots.

Tag name:

Reason:

Tag name	Appears on	Description	Source	Active?	Tagged changes	Actions
mw-manual-revert	Manual revert (edit)	Edits that manually restore the page to an exact previous state (edit)	Defined by the software	Yes	16 changes	
mw-blank	Blanking (edit)	Edits that blank a page (edit)	Defined by the software	Yes	2 changes	
OAuth CID: 4	Quickstatements [OAuth CID: 4] (edit)	Quickstatements (edit)	Defined by the software	Yes	2 changes	
mw-reverted	Reverted (edit)	Edits that were later reverted by a different edit (edit)	Defined by the software	Yes	1 change	
mw-new-redirect	New redirect (edit)	Edits that create a new redirect or change a page to a redirect (edit)	Defined by the software	Yes	1 change	
openrefine-3.4	openrefine-3.4 (edit)	used for edits made by OpenRefine 3.4. Tag is added directly by the tool without an AbuseFilter. (edit)	Applied manually by users and bots	Yes	0 changes	delete   deactivate
openrefine-3.5	openrefine-3.5 (edit)	used for edits made by OpenRefine 3.5. Tag is added directly by the tool without an AbuseFilter. (edit)	Applied manually by users and bots	Yes	0 changes	delete   deactivate
openrefine-3.7	openrefine-3.7 (edit)	used for edits made by OpenRefine 3.7. Tag is added directly by the tool without an AbuseFilter. (edit)	Applied manually by users and bots	Yes	0 changes	delete   deactivate
openrefine-3.6	openrefine-3.6 (edit)	used for edits made by OpenRefine 3.6. Tag is added directly by the tool without an AbuseFilter. (edit)	Applied manually by users and bots	Yes	0 changes	delete   deactivate
mw-contentmodelchange	content model change (edit)	Edits that change the content model of a page (edit)	Defined by the software	Yes	0 changes	
mw-removed-redirect	Removed redirect (edit)	Edits that change an existing redirect to a non-redirect (edit)	Defined by the software	Yes	0 changes	
mw-changed-redirect-target	Redirect target changed (edit)	Edits that change the target of a redirect (edit)	Defined by the software	Yes	0 changes	
mw-replace	Replaced (edit)	Edits that remove more than 90% of the content of a page (edit)	Defined by the software	Yes	0 changes	
mw-rollback	Rollback (edit)	Edits that roll back previous edits using the rollback link (edit)	Defined by the software	Yes	0 changes	
mw-undo	Undo (edit)	Edits that undo previous edits using the undo link (edit)	Defined by the software	Yes	0 changes	
mw-server-side-upload	Server-side upload (edit)	Media files that were uploaded via a maintenance script (edit)	Defined by the software	Yes	0 changes	
mw-add-media	(hidden) (edit)	Edits that add media (edit)	Defined by the software	No	0 changes	
mw-remove-media	(hidden) (edit)	Edits that remove media (edit)	Defined by the software	No	0 changes	

Add these tags manually to the **Special:Tags** page of your Wikibase



```
14 "properties": {
15   "instance_of": "P3",
16   "subclass_of": "P2"
17 }
```

<https://kbttestwikibase.wikibase.cloud/wiki/Special:Tags>

# 3) Configure your Wikibase instance

If you don't add them, OpenRefine will give you errors like these

```
:52:43.073 [refine] GET /command/core/get-csrf-token (42893ms)
:52:43.073 [refine] POST /command/wikidata/perform-wikibase-edits (0ms)
:52:43.073 [..WikibaseEditsOperation] Performing edits (0ms)
:52:43.073 [..ting.EditBatchProcessor] requesting documents (0ms)
:52:43.951 [..ting.EditBatchProcessor] MediaWiki error while editing [badtags]: The tag "Edit using OpenRefine-3.7" is not allowed to be manually applied. (878ms)
:52:44.223 [..ting.EditBatchProcessor] MediaWiki error while editing [badtags]: The tag "Edit using OpenRefine-3.7" is not allowed to be manually applied. (272ms)
:52:44.477 [..ting.EditBatchProcessor] MediaWiki error while editing [badtags]: The tag "Edit using OpenRefine-3.7" is not allowed to be manually applied. (254ms)
:52:44.755 [..ting.EditBatchProcessor] MediaWiki error while editing [badtags]: The tag "Edit using OpenRefine-3.7" is not allowed to be manually applied. (278ms)
:52:45.018 [..ting.EditBatchProcessor] MediaWiki error while editing [badtags]: The tag "Edit using OpenRefine-3.7" is not allowed to be manually applied. (263ms)
:52:45.294 [..ting.EditBatchProcessor] MediaWiki error while editing [badtags]: The tag "Edit using OpenRefine-3.7" is not allowed to be manually applied. (276ms)
:52:45.679 [refine] GET /command/core/get-history (385ms)
:52:45.711 [refine] GET /command/core/get-project-metadata (32ms)
:52:45.726 [refine] GET /command/core/get-models (15ms)
:52:45.742 [refine] POST /command/core/get-all-preferences (16ms)
:52:45.742 [refine] POST /command/core/get-rows (0ms)
:52:45.795 [refine] POST /command/core/compute-facets (53ms)
:52:45.842 [refine] GET /command/core/get-preference (47ms)
:52:45.880 [refine] GET /command/core/get-csrf-token (38ms)
:52:45.895 [refine] POST /command/wikidata/preview-wikibase-schema (15ms)
```

mw-changed-redirect-target	Redirect target changed (edit)	Edits that change the target of a redirect (edit)	Defined by the software	Yes	0 changes
mw-replace	Replaced (edit)	Edits that remove more than 90% of the content of a page (edit)	Defined by the software	Yes	0 changes
mw-rollback	Rollback (edit)	Edits that roll back previous edits using the rollback link (edit)	Defined by the software	Yes	0 changes
mw-undo	Undo (edit)	Edits that undo previous edits using the undo link (edit)	Defined by the software	Yes	0 changes
mw-server-side-upload	Server-side upload (edit)	Media files that were uploaded via a maintenance script (edit)	Defined by the software	Yes	0 changes
mw-add-media	(hidden) (edit)	Edits that add media (edit)	Defined by the software	No	0 changes
mw-remove-media	(hidden) (edit)	Edits that remove media (edit)	Defined by the software	No	0 changes

```
14 "properties": {
15   "instance_of": "P3",
16   "subclass_of": "P2"
17 }
```

<https://kbttestwikibase.wikibase.cloud/wiki/Special:Tags>

# 3 easy steps: approach documented by KB

Search or jump to... Pull requests Issues Codespaces Marketplace Explore

KBNLresearch / OpenRefine-Wikibase Public Edit Pins Unwatch 5 Fork 1 Star 0

<> Code Issues Pull requests Actions Projects Wiki Security Insights Settings

main 1 branch 0 tags Go to file Add file Code

ookgezellig Update README.md e44f343 8 days ago 44 commits

images	Add files via upload	9 days ago
README.md	Update README.md	8 days ago
kb-test-wikibase-cloud-config.py	Rename config.py to kb-test-wikibase-cloud-config.py	last month
kb-test-wikibase-cloud-manifest...	Update kb-test-wikibase-cloud-manifest.json	2 months ago
kbga-test-azure-config.py	Update kbga-test-azure-config.py	last month
kbga-test-azure-manifest.json	Update kbga-test-azure-manifest.json	last month

README.md

## OpenRefine-Wikibase

Files for interaction between OpenRefine and KB Wikibases.

Purpose: reconciling and uploading data to Wikibases of the KB, using Openrefine

### Wikibases of the KB

- [https://kbttestwikibase.wikibase.cloud/wiki/Main\\_Page](https://kbttestwikibase.wikibase.cloud/wiki/Main_Page) (informal test/prototyping WB instance, not part of the KB IT infrastructure)
- [http://kbga-wiki-test.westeurope.azurecontainer.io/wiki/Main\\_Page](http://kbga-wiki-test.westeurope.azurecontainer.io/wiki/Main_Page) (formal test WB instance, part of the IT-infra of the KB)

Setting up a reconciliation service for your Wikibase

General

About: Files for interaction between OpenRefine and KB Wikibases. 0 stars, 5 watching, 1 fork.

Releases: No releases published. Create a new release.

Packages: No packages published. Publish your first package.

Languages: Python 100.0%

<https://github.com/KBNLresearch/OpenRefine-Wikibase>

# Text summary of this slidedeck

- Reconciliation service for Wikidata - <https://wikidata.reconci.link> + <https://wikidata.reconci.link/en/api>
- Reconciliation service API specification - <https://reconciliation-api.github.io/specs/latest/>
- Documentation OpenRefine-Wikibase for developers - <https://openrefine-wikibase.readthedocs.io/en/latest/>
- Wikidata reconciliation service endpoints (as needed by OpenRefine) - <https://wikidata.reconci.link/en/api> // <https://wikidata.reconci.link/nl/api> // <https://wikidata.reconci.link/fr/api> // Etc..
- Reconciliation service for Wikimedia Commons - <https://commonsreconcile.toolforge.org> + <https://commonsreconcile.toolforge.org/en/api>
- Reconciliation service for Rhizome Artbase - <https://reconciliation.artbase.rhizome.org> + <https://reconciliation.artbase.rhizome.org/en/api>
- Reconciliation service for EU Knowledge Graph - <https://openrefine-reconciliation.linkedopendata.eu> + <https://openrefine-reconciliation.linkedopendata.eu/en/api>
- Reconciliation services overview - <https://reconciliation-api.github.io/testbench/#/>

## **Making a recon service for your own Wikibase and use it in OpenRefine in 3 easy steps**

### *1) Install the recon service for your Wikibase*

- <https://github.com/wetneb/openrefine-wikibase>
- Installing with Docker - <https://openrefine-wikibase.readthedocs.io/en/latest/install.html>
- *config.py* for <https://kbttestwikibase.wikibase.cloud> - <https://github.com/KBNLresearch/OpenRefine-Wikibase/blob/main/kb-test-wikibase-cloud-config.py>
- Reconciliation service for Kunstmuseum Den Haag - <https://reconciliation.kunstmuseum.nl> + <https://reconciliation.kunstmuseum.nl/nl/api>

### *2) Connect OpenRefine to your Wikibase instance*

- <https://docs.openrefine.org/manual/wikibase/configuration>
- Wikibase example manifests - <https://github.com/OpenRefine/wikibase-manifests>
- Wikimedia Commons manifest - <https://github.com/OpenRefine/wikibase-manifests/blob/master/wikimedia-commons-manifest.json>
- Write a Wikibase manifest - <https://github.com/OpenRefine/OpenRefine/wiki/Write-a-Wikibase-manifest>
- Manifest.json for Kunstmuseum Den Haag - [https://api.kunstmuseum.nl/wiki/Kunstmuseum\\_API#Openrefine](https://api.kunstmuseum.nl/wiki/Kunstmuseum_API#Openrefine)

### *3) Configure your Wikibase instance*

- <https://github.com/KBNLresearch/OpenRefine-Wikibase/blob/main/kb-test-wikibase-cloud-manifest.json>
- Add tags 'openrefine-3.4', 'openrefine-3.5', 'openrefine-3.6' and 'openrefine-3.7' manually to the Special:Tags page of your Wikibase - <https://kbttestwikibase.wikibase.cloud/wiki/Special:Tags>

3 easy steps: approach documented by KB - <https://github.com/KBNLresearch/OpenRefine-Wikibase>