

Abgleich & Anreicherung eigener Daten mit der GND

OpenRefine Reconciliation mit lobid-gnd

Fabian Steeg & Adrian Pohl, Offene Infrastruktur, hbz NRW



Workshop im Rahmen des DINI-KIM, 2021-10-07

Diese Präsentation:

<http://slides.lobid.org/2021-kim-reconcile/>



Überblick

Datenquellen und Oberfläche von lobid-gnd

Datenabgleich und -anreicherung in OpenRefine

Möglichkeiten zur Verbesserung des Matching

Standardisierung & W3C Community Group

Eigene Daten in OpenRefine verfügbar machen

Überblick

Datenquellen und Oberfläche von lobid-gnd

Datenabgleich und -anreicherung in OpenRefine

Möglichkeiten zur Verbesserung des Matching

Standardisierung & W3C Community Group

Eigene Daten in OpenRefine verfügbar machen

Wir haben versucht, Themenwünsche zu berücksichtigen, soweit wir uns damit auskennen

Datenquellen und Oberfläche von lobid-gnd

Datenabgleich und -anreicherung in OpenRefine

Möglichkeiten zur Verbesserung des Matching

Standardisierung & W3C Community Group

Eigene Daten in OpenRefine verfügbar machen





lobid steht für *Linking Open Bibliographic Data*



lobid steht für *Linking Open Bibliographic Data*

Unter lobid.org publiziert das hbz seit 2011 Linked Open Data (LOD), inzwischen in 3 Diensten



lobid steht für *Linking Open Bibliographic Data*

Unter lobid.org publiziert das hbz seit 2011 Linked Open Data (LOD), inzwischen in 3 Diensten

Stellen web-basierte Rechercheoberflächen und Schnittstellen bereit





Gemeinsame Normdatei für Bibliotheken im
deutschsprachigen Raum



Gemeinsame Normdatei für Bibliotheken im
deutschsprachigen Raum

Datensätze für Personen, Körperschaften, Konferenzen &
Veranstaltungen, Geografika, Schlagwörter, Werke



Gemeinsame Normdatei für Bibliotheken im
deutschsprachigen Raum

Datensätze für Personen, Körperschaften, Konferenzen &
Veranstaltungen, Geografika, Schlagwörter, Werke

Für die formale Erfassung einer Ressource und für die
inhaltliche Erschließung



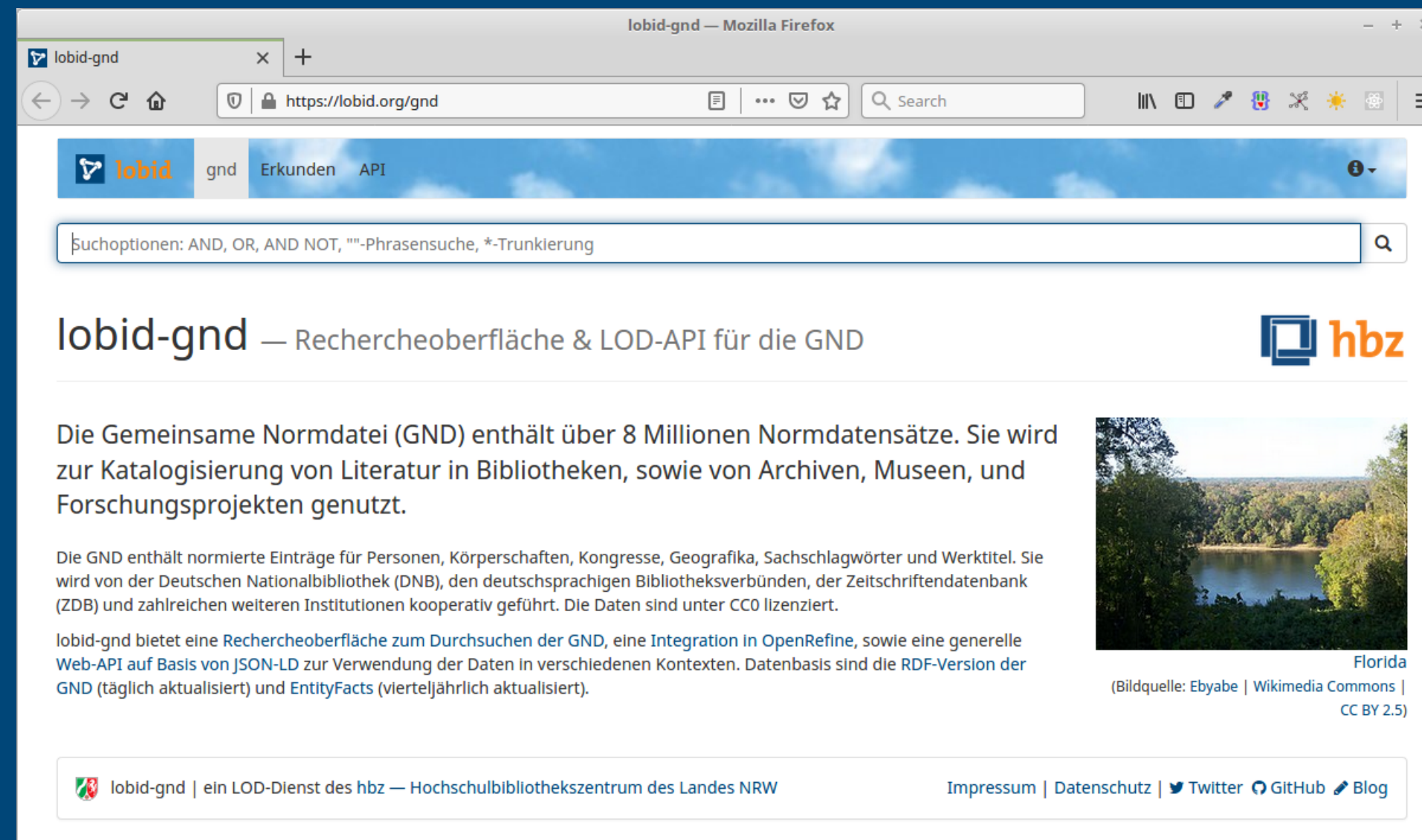
Gemeinsame Normdatei für Bibliotheken im
deutschsprachigen Raum

Datensätze für Personen, Körperschaften, Konferenzen &
Veranstaltungen, Geografika, Schlagwörter, Werke

Für die formale Erfassung einer Ressource und für die
inhaltliche Erschließung

In den letzten Jahren wird die GND vermehrt auch von
Archiven, Museen und Wissenschaftler:innen genutzt

lobid-gnd



The screenshot shows the Mozilla Firefox browser window with the address bar displaying `https://lobid.org/gnd`. The page features a navigation bar with the lobid logo, 'gnd', 'Erkunden', and 'API' links. A search bar contains the text 'Suchoptionen: AND, OR, AND NOT, ""-Phrasensuche, *-Trunkierung'. The main heading is 'lobid-gnd — Rechercheoberfläche & LOD-API für die GND', accompanied by the hbz logo. The content area explains that the GND contains over 8 million norm data records and is used for cataloging literature. It also mentions that lobid-gnd provides a research interface, OpenRefine integration, and a JSON-LD Web-API. A photograph of a lake in Florida is shown with the caption '(Bildquelle: Ebyabe | Wikimedia Commons | CC BY 2.5)'. The footer includes the text 'lobid-gnd | ein LOD-Dienst des hbz — Hochschulbibliothekszentrum des Landes NRW' and links to 'Impressum', 'Datenschutz', 'Twitter', 'GitHub', and 'Blog'.

lobid-gnd — Rechercheoberfläche & LOD-API für die GND

Die Gemeinsame Normdatei (GND) enthält über 8 Millionen Normdatensätze. Sie wird zur Katalogisierung von Literatur in Bibliotheken, sowie von Archiven, Museen, und Forschungsprojekten genutzt.

Die GND enthält normierte Einträge für Personen, Körperschaften, Kongresse, Geografika, Sachschlagwörter und Werktitel. Sie wird von der Deutschen Nationalbibliothek (DNB), den deutschsprachigen Bibliotheksverbünden, der Zeitschriftendatenbank (ZDB) und zahlreichen weiteren Institutionen kooperativ geführt. Die Daten sind unter CC0 lizenziert.

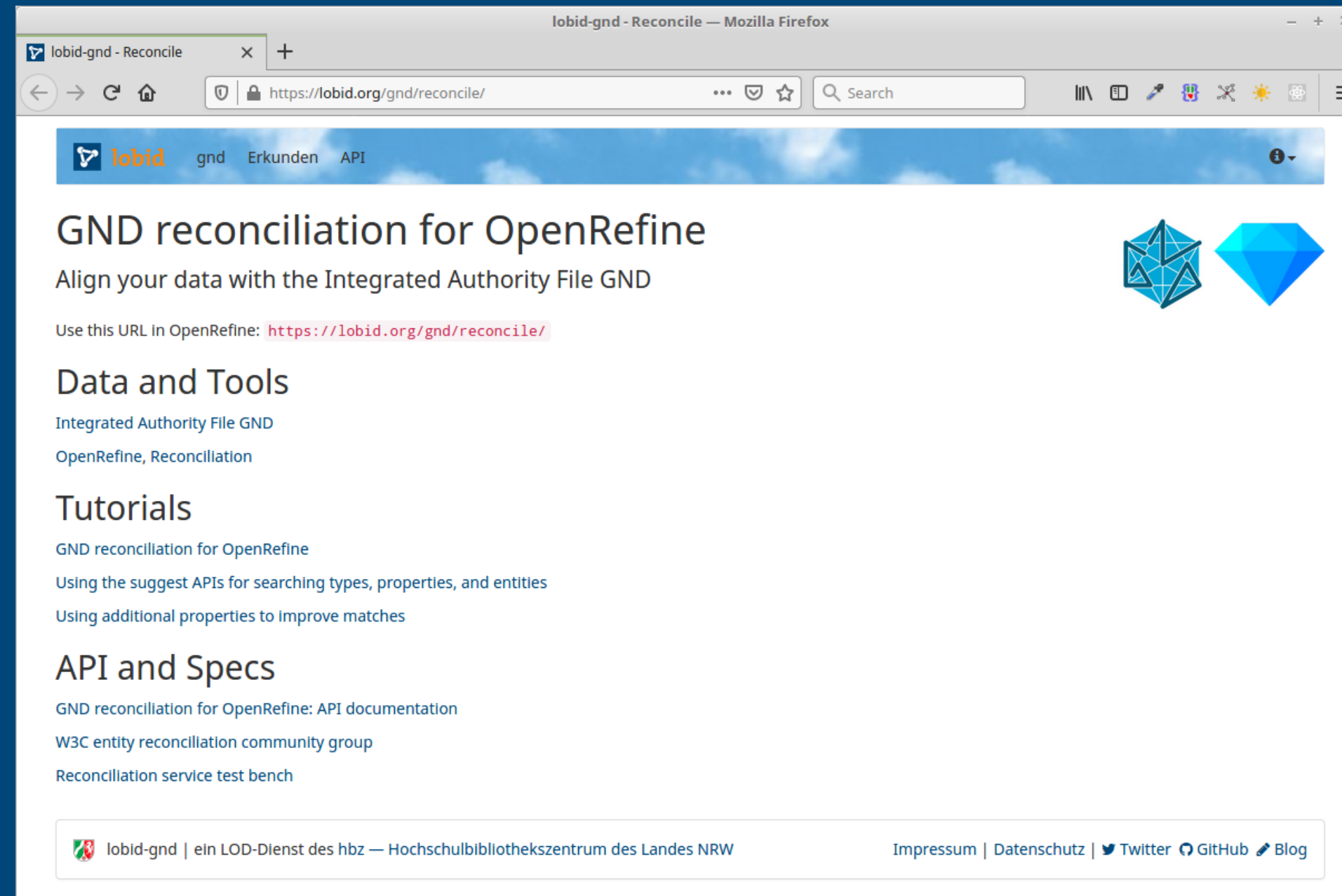
lobid-gnd bietet eine Rechercheoberfläche zum Durchsuchen der GND, eine Integration in OpenRefine, sowie eine generelle Web-API auf Basis von JSON-LD zur Verwendung der Daten in verschiedenen Kontexten. Datenbasis sind die RDF-Version der GND (täglich aktualisiert) und EntityFacts (vierteljährlich aktualisiert).

Florida
(Bildquelle: Ebyabe | Wikimedia Commons | CC BY 2.5)

lobid-gnd | ein LOD-Dienst des hbz — Hochschulbibliothekszentrum des Landes NRW

Impressum | Datenschutz | Twitter | GitHub | Blog

/reconcile



The screenshot shows a web browser window with the title "lobid-gnd - Reconcile — Mozilla Firefox". The address bar displays "https://lobid.org/gnd/reconcile/". The page content includes a header with the lobid logo and navigation links "gnd", "Erkunden", and "API". The main heading is "GND reconciliation for OpenRefine", followed by the subtitle "Align your data with the Integrated Authority File GND". A note states: "Use this URL in OpenRefine: <https://lobid.org/gnd/reconcile/>". Below this are sections for "Data and Tools" (with links to "Integrated Authority File GND" and "OpenRefine, Reconciliation") and "Tutorials" (with links to "GND reconciliation for OpenRefine", "Using the suggest APIs for searching types, properties, and entities", and "Using additional properties to improve matches"). The "API and Specs" section contains links to "GND reconciliation for OpenRefine: API documentation", "W3C entity reconciliation community group", and "Reconciliation service test bench". The footer includes the text "lobid-gnd | ein LOD-Dienst des hbz — Hochschulbibliothekszentrum des Landes NRW" and links for "Impressum", "Datenschutz", "Twitter", "GitHub", and "Blog".

lobid-gnd - Reconcile — Mozilla Firefox

lobid gnd Erkunden API

GND reconciliation for OpenRefine

Align your data with the Integrated Authority File GND

Use this URL in OpenRefine: <https://lobid.org/gnd/reconcile/>

Data and Tools

[Integrated Authority File GND](#)
[OpenRefine, Reconciliation](#)

Tutorials

[GND reconciliation for OpenRefine](#)
[Using the suggest APIs for searching types, properties, and entities](#)
[Using additional properties to improve matches](#)

API and Specs

[GND reconciliation for OpenRefine: API documentation](#)
[W3C entity reconciliation community group](#)
[Reconciliation service test bench](#)

lobid-gnd | ein LOD-Dienst des hbz — Hochschulbibliothekszentrum des Landes NRW

[Impressum](#) | [Datenschutz](#) | [Twitter](#) | [GitHub](#) | [Blog](#)

Abgleich und Anreicherung lokaler Daten mit der GND

OpenRefine kpe csv Permalink

Open... Export Help

Facet / Filter Undo / Redo 1 / 1

Refresh Reset All Remove All

Name: judgment change

3 choices Sort by: name count

(blank) 129
matched 184
none 162

Facet by choice counts

Name: best candidate's score change reset

40.00 — 160.00

☒ Numeric 345 ☐ Non-numeric 0 ☒ Blank 130 ☐ Error 0

475 rows

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

Extensions: Wikidata

« first < previous 1 - 50 next > last »

			All	GND	Name	DateOfBirth	DateOfDeath	Type
☆	🗨	1.	KPE 186170	Neumann, Jsrael Ben	1887	18.04.1961	personal	
				☒ Neumann, Paul (65) ☒ Neumann, Erich (65) ☒ Ben-Elī'ezer (61) ☒ Neumann, Hans Paul ☒ Hartmann, Jsrael (58) ☒ Israeli, Ben-Zion (56) ☒ Pinḥas ben Y. (56) ☒ Ben-Aroya, Avraham ☒ Rôm, Mōše Ben Yôḥa ☒ Ben-Yīśra'el, Asher (50) ☒ Create new item				
☆	🗨	2.	KPE 243353	Yamawaki, Michiko	13.07.1910	2000	personal	
				Choose new match				
☆	🗨	3.	KPE 255684	Ackermann, Richard	1905	1977	personal	
				Choose new match				
☆	🗨	4.	KPE 255687	Agatz, August	1904	1945	personal	
				☒ Agatz, Wilhelm (85) ☒ Agatz, August (82) ☒ Agatz, Alfred (60)				

Match this Cell Match All Identical Cells Cancel

[Neumann, Paul](#) (1146685785)

1887-

Individualisierte Person

Search for match

https://lobid.org/gnd/1146685785

Datenquellen

Datenquellen

Die GND wird erstellt und gepflegt durch die GND-Kooperative (Verbünde, Bibliotheken, DNB, ZDB, Agenturen)

Datenquellen

Die GND wird erstellt und gepflegt durch die GND-Kooperative (Verbünde, Bibliotheken, DNB, ZDB, Agenturen)

Die DNB publiziert die GND u.a. als **Linked Open Data**

Datenquellen

Die GND wird erstellt und gepflegt durch die GND-Kooperative (Verbünde, Bibliotheken, DNB, ZDB, Agenturen)

Die DNB publiziert die GND u.a. als **Linked Open Data**

Zusätzliche Anreicherungen, z.B. Links, in **EntityFacts**

Datenquellen

Die GND wird erstellt und gepflegt durch die GND-Kooperative (Verbünde, Bibliotheken, DNB, ZDB, Agenturen)

Die DNB publiziert die GND u.a. als **Linked Open Data**

Zusätzliche Anreicherungen, z.B. Links, in **EntityFacts**

Diese Daten bilden die Grundlage für lobid-gnd

Daten erkunden

Daten erkunden

Rechercheoberfläche ermöglicht ein manuelles Erkunden
der Daten vor einem automatischen Abgleich

Daten erkunden

Rechercheoberfläche ermöglicht ein manuelles Erkunden der Daten vor einem automatischen Abgleich

Zum Ausprobieren verlinken die folgenden Screenshots auf die entsprechenden Seiten in lobid-gnd

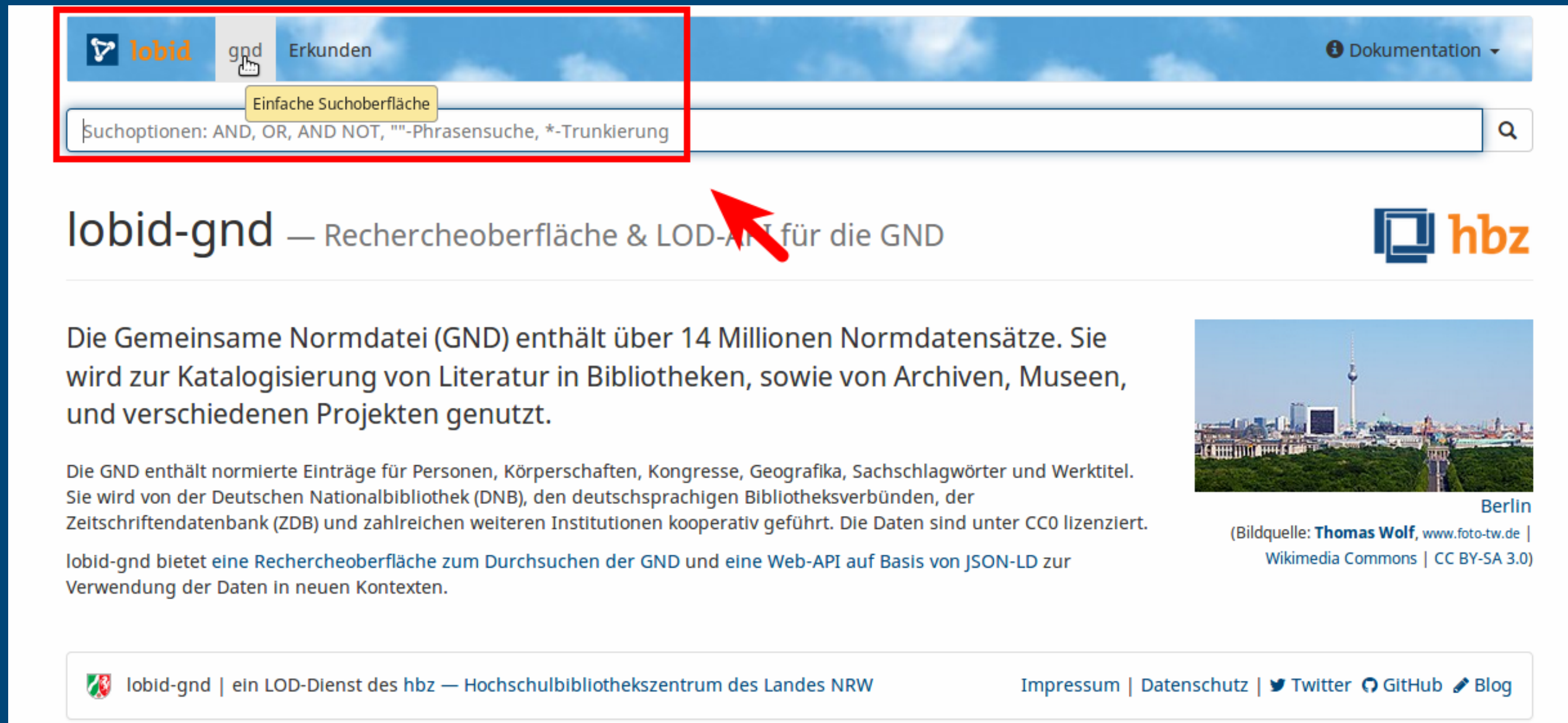
Daten erkunden

Rechercheoberfläche ermöglicht ein manuelles Erkunden der Daten vor einem automatischen Abgleich

Zum Ausprobieren verlinken die folgenden Screenshots auf die entsprechenden Seiten in lobid-gnd

Siehe auch das Tutorial [lobid-gnd: Suche und Navigation](#)

Oberfläche



lobid-gnd — Rechercheoberfläche & LOD-API für die GND

Die Gemeinsame Normdatei (GND) enthält über 14 Millionen Normdatensätze. Sie wird zur Katalogisierung von Literatur in Bibliotheken, sowie von Archiven, Museen, und verschiedenen Projekten genutzt.

Die GND enthält normierte Einträge für Personen, Körperschaften, Kongresse, Geografika, Sachschlagwörter und Werktitel. Sie wird von der Deutschen Nationalbibliothek (DNB), den deutschsprachigen Bibliotheksverbünden, der Zeitschriftendatenbank (ZDB) und zahlreichen weiteren Institutionen kooperativ geführt. Die Daten sind unter CC0 lizenziert.

lobid-gnd bietet eine Rechercheoberfläche zum Durchsuchen der GND und eine Web-API auf Basis von JSON-LD zur Verwendung der Daten in neuen Kontexten.

Berlin

(Bildquelle: **Thomas Wolf**, www.foto-tw.de | Wikimedia Commons | CC BY-SA 3.0)

lobid-gnd | ein LOD-Dienst des hbz — Hochschulbibliothekszentrum des Landes NRW

Impressum | Datenschutz | [Twitter](#) | [GitHub](#) | [Blog](#)

Auto Suggest

The screenshot shows a search interface with a text input field containing 'köl'n' and a magnifying glass icon. Below the input field is a list of suggestions. The third suggestion, 'Dom Köln (Köl'n) | 1248-1880', is highlighted with a red rectangular box. Each suggestion is followed by a category label in a dark blue box.

Suggestion	Category
Köl'n	Gebietskörperschaft oder Verwaltungseinheit
Köl'n Bongartz, Marianne	Werk
Dom Köln (Köl'n) 1248-1880	Bauwerk oder Denkmal
Künstlergruppe Köln Eins (Köl'n) Köl'n	Körperschaft
Kunsträume Köln (Galerie, Köl'n)	Körperschaft
Photofreunde Köln Rechtsrheinisch (Köl'n) Köl'n	Körperschaft
Dom Köln	Schlagwort

Ergebnisliste

köln

Treffer pro Seite: 10 30 50 100

5 Treffer, zeige 1 bis 5:

Teoman, Sibel Susann	Individualisierte Person	Schriftstellerin, Germanistin, Assistentin	13150276X 1975 –
Sevindim, Asli	Individualisierte Person	Schriftstellerin, Journalistin, Moderatorin	130454966 1973 –
Meray, Sema	Individualisierte Person	SchauspielerIn, Dramatikerin	136115519 1968 –
Özdoğan, Selim	Individualisierte Person	Schriftsteller	115370800 1971 –
Akhanlı, Doğan	Individualisierte Person	Schriftsteller	124063802 1957 –

← 1 →

Ergebnisse eingrenzen:

Entitätstyp

Person Filter entfernen ✕ +

GND-Sachgruppe +

Personen zu Literaturgeschichte (Schriftsteller) Filter entfernen ✕

Personen zu Betriebswirtschaftslehre (1)

Personen zu Theater, Tanz (1)

Personen zu Film (1)

Personen zu Rundfunk, Neuen Medien (1)

Ländercode

Deutschland Filter entfernen ✕

Türkei Filter entfernen ✕

Beruf oder Beschäftigung +

Schriftsteller (2)

Schriftstellerin (2)

Assistentin (1)

Journalistin (1)

Dramatikerin (1)

Einzeltreffer

Helgoländer Bucht

Geografikum

Natürlich geografische Einheit

Felder

Beziehungen

Entitätstyp	Natürlich geografische Einheit
GND-Nummer	4393546-1
Ländercode	Deutschland
GND-Sachgruppe	Physische Geographie Hydrologie, Meereskunde
Oberbegriff partitiv	Nordsee
Oberbegriff instantiell	Bucht
Siehe auch	Wikipedia (English) Wikipedia (Deutsch) Wikidata Virtual International Authority File (VIAF) Deutsche Nationalbibliothek (DNB)

Weitere Einträge mit Oberbegriff partitiv "Nordsee" suchen

Datenquelle: GND und EntityFacts | CC0

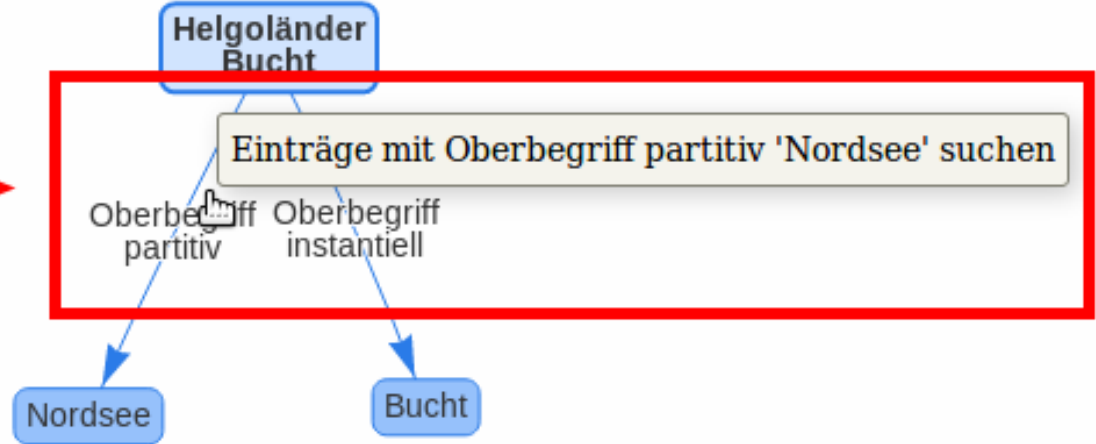


Bildquelle: NASA | Wikimedia Commons | Public domain

Beziehungsgraph

Helgoländer Bucht Geografikum Natürlich geografische Einheit 

Felder Beziehungen 



The graph shows 'Helgoländer Bucht' at the top, connected to 'Nordsee' and 'Bucht' via 'Oberbegriff partitiv' and 'Oberbegriff instanziell' respectively. A red box highlights the search bar 'Einträge mit Oberbegriff partitiv 'Nordsee' suchen' and the two relationship types. A red arrow points to the search bar. Zoom controls (-, +) are at the bottom right of the graph area.



Bildquelle: NASA | Wikimedia Commons | Public domain

Datenquelle: GND und EntityFacts | CC0

Datenquellen und Oberfläche von lobid-gnd

Datenabgleich und -anreicherung in OpenRefine

Möglichkeiten zur Verbesserung des Matching

Standardisierung & W3C Community Group

Eigene Daten in OpenRefine verfügbar machen

Was ist OpenRefine?

Was ist OpenRefine?

"A powerful tool for working with messy data"

Was ist OpenRefine?

"A powerful tool for working with messy data"

"cleaning it; transforming it from one format into another;
and extending it with web services and external data"

Was ist OpenRefine?

"A powerful tool for working with messy data"

"cleaning it; transforming it from one format into another;
and extending it with web services and external data"

Oberfläche wie Tabellenkalkulation

Was ist OpenRefine?

"A powerful tool for working with messy data"

"cleaning it; transforming it from one format into another;
and extending it with web services and external data"

Oberfläche wie Tabellenkalkulation

Läuft im Browser

Wo ist OpenRefine?

Wo ist OpenRefine?

Lokal, downloaden und installieren:

<https://openrefine.org/download.html>

Wo ist OpenRefine?

Lokal, downloaden und installieren:

<https://openrefine.org/download.html>

Gehostet, z.B. zum Ausprobieren:

<http://openrefine.labs.lobid.org>

Wofür ist
OpenRefine?

Wofür ist OpenRefine?

Gut geeignet wo UI/Tabellenkalkulation passt

Wofür ist OpenRefine?

Gut geeignet wo UI/Tabellenkalkulation passt

Fachliche Arbeit

Wofür ist OpenRefine?

Gut geeignet wo UI/Tabellenkalkulation passt

Fachliche Arbeit

Manuelle Optimierung

Wofür ist OpenRefine?

Gut geeignet wo UI/Tabellenkalkulation passt

Fachliche Arbeit

Manuelle Optimierung

Überschaubare Datenmengen

Wofür ist OpenRefine (noch)
nicht?

Wofür ist OpenRefine (noch) nicht?

Bisher nicht geeignet für große Dateien, aber eine neue
Architektur für größere Datenmengen ist fast fertig

Wofür ist OpenRefine (noch) nicht?

Bisher nicht geeignet für große Dateien, aber eine neue
Architektur für größere Datenmengen ist fast fertig

Automatisierung ohne UI nicht mit Bordmitteln, aber mit
Zusatztool möglich: `openrefine-batch`

Was ist Reconciliation?

Was ist Reconciliation?

Abgleichen / Matchen eigener Daten (z.B. Namen von Personen) auf GND-Einträge

Was ist Reconciliation?

Abgleichen / Matchen eigener Daten (z.B. Namen von Personen) auf GND-Einträge

Anreicherung durch Übernahme von spezifischen Feldern der gematchten GND-Einträge

Was ist Reconciliation?

Abgleichen / Matchen eigener Daten (z.B. Namen von Personen) auf GND-Einträge

Anreicherung durch Übernahme von spezifischen Feldern der gematchten GND-Einträge

Fokus hier ist Reconciliation, auch Grundlagen von OpenRefine, aber es gibt noch viel mehr

Projekt anlegen

OpenRefine - Mozilla Firefox

OpenRefine

127.0.0.1:3333

Search

OpenRefine

A power tool for working with messy data.

Create Project

Open Project

Import Project

Language Settings

Version 3.0-beta [TRUNK]

Preferences

Help

About

Create a project by importing data. What kinds of data files can I import?

TSV, CSV, *SV, Excel (.xls and .xlsx), JSON, XML, RDF as XML, and Google Data documents are all supported. Support for other formats can be added with OpenRefine extensions.

Get data from

This Computer

Web Addresses (URLs)

Clipboard

Data Package (JSON URL)

Database

Google Data

Locate one or more files on your computer to upload:

Browse...

No files selected.

Next »

23

OpenRefine - Mozilla Firefox

OpenRefine

127.0.0.1:3333

Search

OpenRefine

A power tool for working with messy data.

Create Project

Open Project

Import Project

Language Settings

Version 3.0-beta [TRUNK]

Preferences

Help

About

Create a project by importing data. What kinds of data files can I import?

TSV, CSV, *SV, Excel (.xls and .xlsx), JSON, XML, RDF as XML, and Google Data documents are all supported. Support for other formats can be added with OpenRefine extensions.

Get data from

This Computer

Web Addresses (URLs)

Clipboard

Data Package (JSON URL)

Database

Google Data

Paste data from clipboard here:

name;beruf;ort
J. Weizenbaum;Informatiker;Berlin
Twain, Mark;Schriftsteller;
Kumar, Lalit;;
Jemand;;

Next »

24

OpenRefine - Mozilla Firefox

OpenRefine

127.0.0.1:3333

Search

OpenRefine

A power tool for working with messy data.

Create Project

Open Project

Import Project

Language Settings

« Start Over

Configure Parsing Options

Project name

clipboard

Tags

Create Project »

	name	beruf	ort
1.	J. Weizenbaum	Informatiker	Berlin
2.	Twain, Mark	Schriftsteller	
3.	Kumar, Lalit		
4.	Jemand		

Parse data as

CSV / TSV / separator-based files

Line-based text files

Fixed-width field text files

PC-Axis text files

JSON files

MARC files

RDF/N3 files

Wikitext

XML files

Open Document Format spreadsheets

Character encoding

Update Preview

Columns are separated by

commas (CSV)

tabs (TSV)

custom ;

Escape special characters with \

Ignore first 0 line(s) at beginning of file

Parse next 1 line(s) as column headers

Discard initial 0 row(s) of data

Load at most 0 row(s) of data

Use " character to enclose cells containing column separators

Parse cell

Store blank

Version 3.0-beta [TRUNK]

Preferences

Help

About

clipboard - OpenRefine

+

openrefine.labs.lobid.org/proj

120%

Search

Download

Print

Copy

Share

Fullscreen

Help

OpenRefine clipboard [Permalink](#)

Open... Export ▾ Help

Facet / Filter

Undo / Redo 0 / 0

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](#)

4 rows

Extensions: Wikidata ▾

Show as: rows records

Show: 5 10 25 50 rows

« first < previous 1 - 4 next > last »

			All	name	beruf	ort
☆	🗨	1.	J. Weizenbaum	Informatiker	Berlin	
☆	🗨	2.	Twain, Mark	Schriftsteller		
☆	🗨	3.	Kumar, Lalit			
☆	🗨	4.	Jemand			

javascript:()

Projekt anlegen

openrefine.labs.lobid.org | localhost:3333

```
name;beruf;ort  
J. Weizenbaum;Informatiker;Berlin  
Twain, Mark;Schriftsteller;  
Kumar, Lalit;;  
Jemand;;
```

(Daten und URL per CTRL+C in die Zwischenablage kopieren)

Abgleich / Reconciliation

clipboard - OpenRefine - Mozilla Firefox

clipboard - OpenRefine x +

127.0.0.1:3333/project?project=2221362040394

OpenRefine clipboard Permalink

Open... Export Help

Facet / Filter Undo / Redo 0 / 0

4 rows

Extensions: Wikidata

Show as: rows records Show: 5 10 25 50 rows « first < previous 1 - 4 next > last »

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.

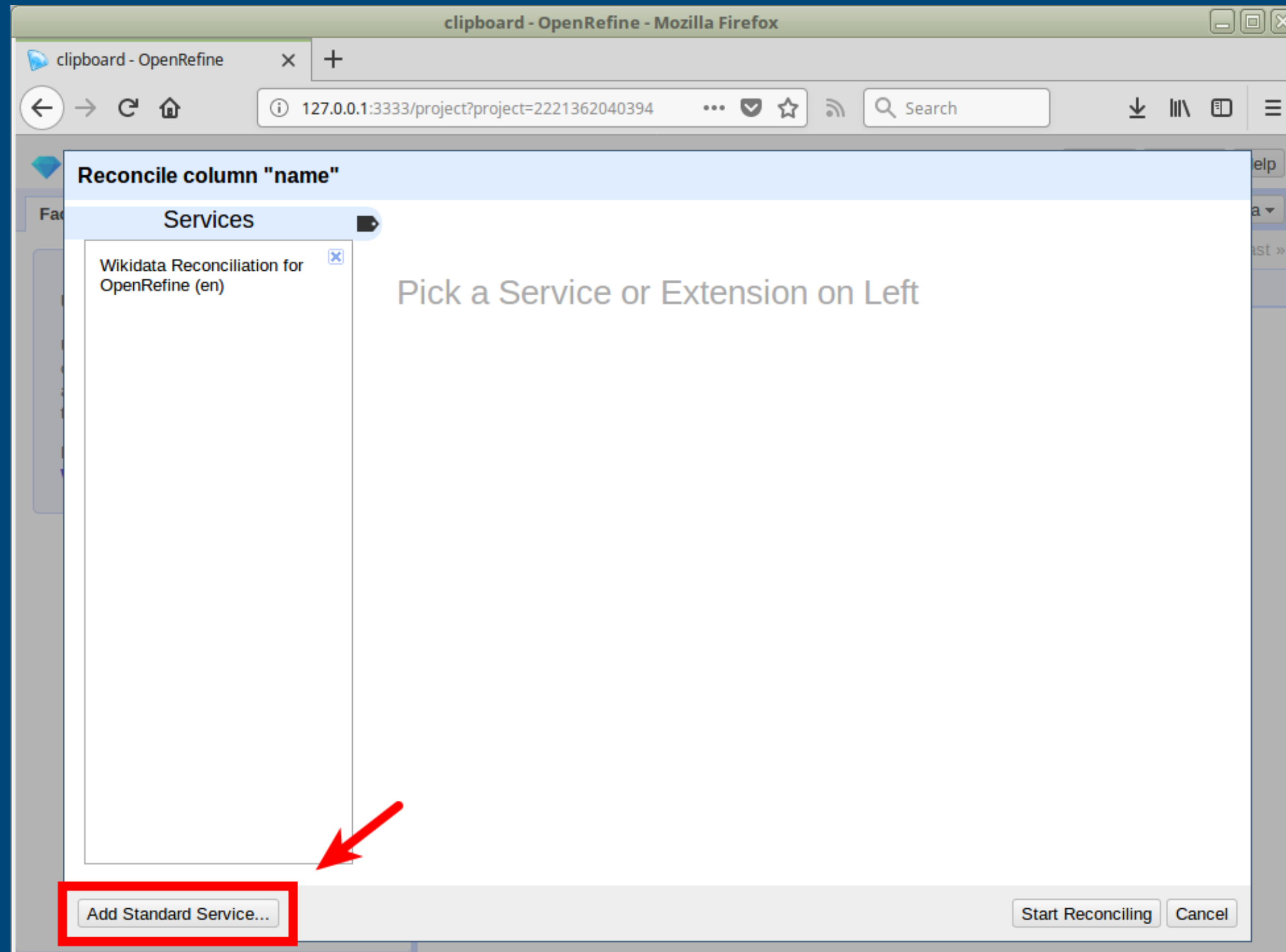
Not sure how to get started?
[Watch these screencasts](#)

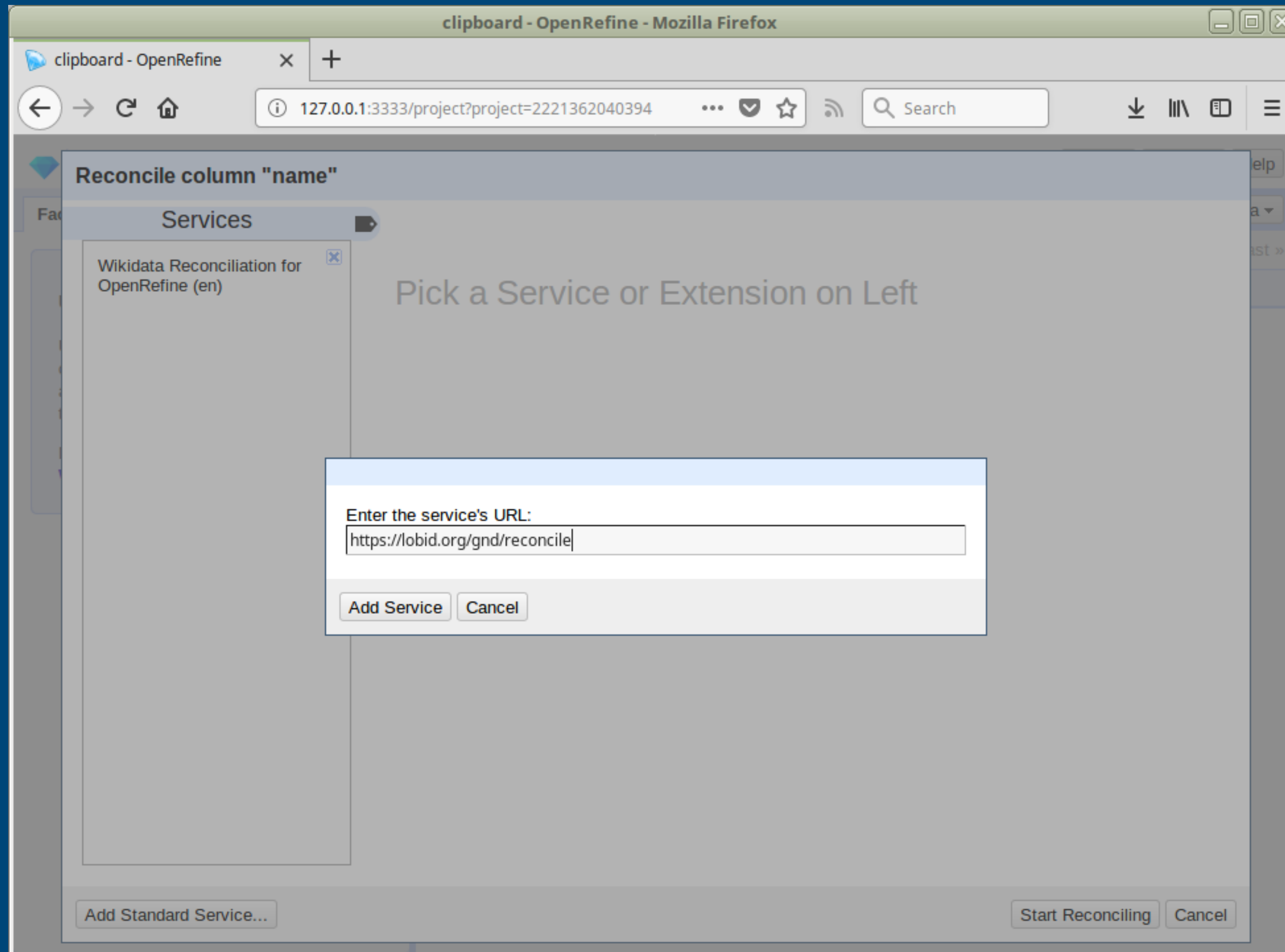
1. Facet
2. Text filter
3. Edit cells
4. Edit column
Transpose
Sort...
View
Reconcile

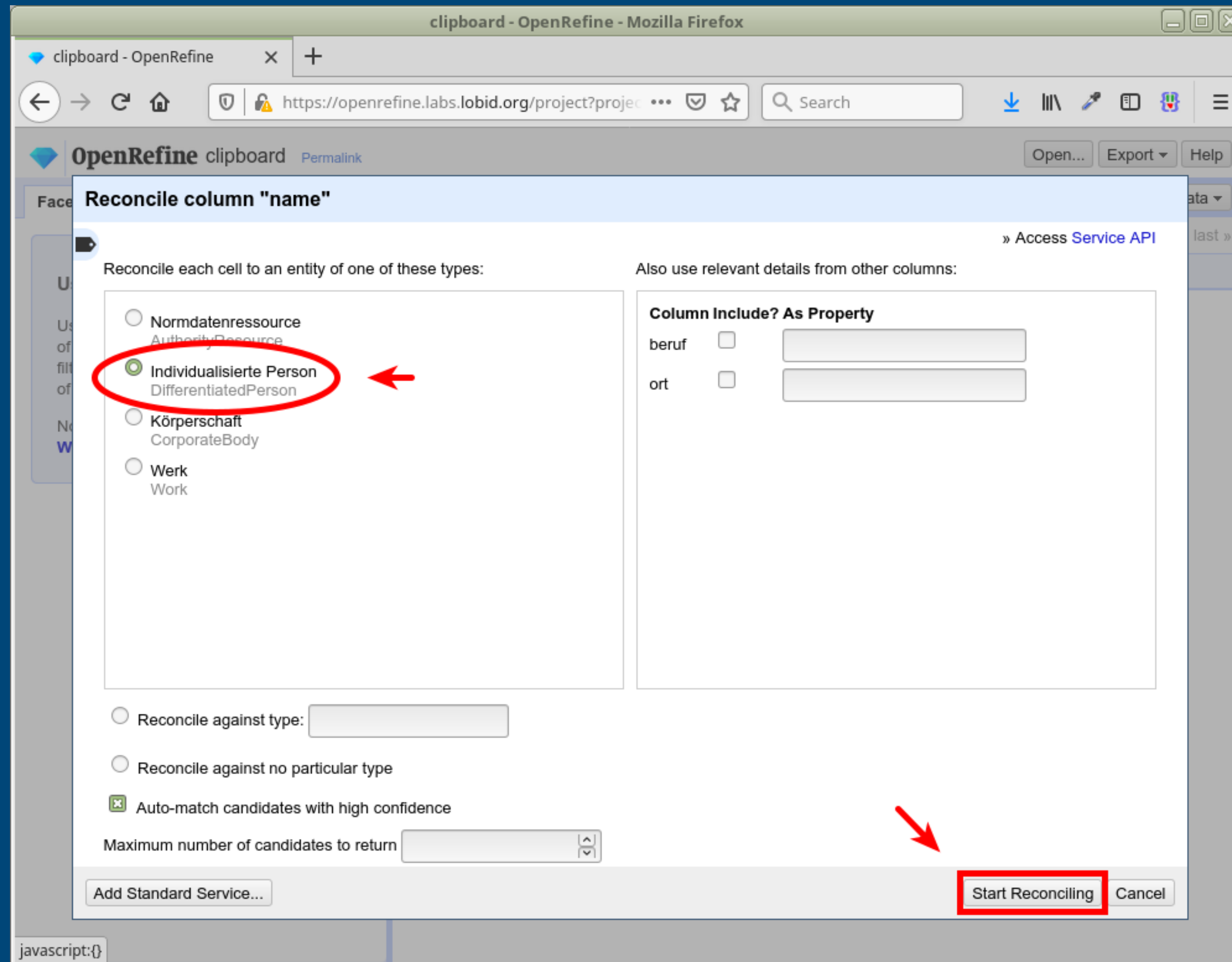
Start reconciling...

Facets
Actions
Copy reconciliation data...

The screenshot shows the OpenRefine web application running in a Mozilla Firefox browser. The browser's address bar displays the URL '127.0.0.1:3333/project?project=2221362040394'. The OpenRefine interface has a top navigation bar with 'Open...', 'Export', and 'Help' buttons. Below this, a sidebar on the left contains a 'Facet / Filter' section with a 'Using facets and filters' guide. The main area shows a table with 4 rows. The first column is 'name', the second is 'beruf', and the third is 'ort'. The 'name' column has a dropdown menu open, showing options like 'Facet', 'Text filter', 'Edit cells', 'Edit column', 'Transpose', 'Sort...', 'View', and 'Reconcile'. The 'Reconcile' option is highlighted, and a sub-menu is visible with 'Start reconciling...' selected. A red arrow points to the 'Reconcile' option in the main menu, and another red arrow points to the 'Start reconciling...' option in the sub-menu.







clipboard - OpenRefine

+

openrefine.labs.lobid.org/proj

110%

Search

Download

Print

Copy

Share

Fullscreen

Settings

Menu

OpenRefine

clipboard

Permalink

Open...

Export

Help

Facet / Filter

Undo / Redo 3 / 3

Refresh

Reset All

Remove All

name: judgment

change

2 choices

Sort by: name count

matched 2

none 2

Facet by choice counts

name: best candidate's score

change

reset

26.00 — 82.00

4 rows

Extensions: Wikidata

Show as: rows records

Show: 5 10 25 50 rows

« first < previous 1 - 4 next > last »

	All	name	beruf	ort
☆	🗨	1. Weizenbaum, Joseph Choose new match	Informatiker	Berlin
☆	🗨	2. Twain, Mark Choose new match	Schriftsteller	
☆	🗨	3. Kumar, Lalit ☒ ☒ Kumar, Lalit (80) ☒ ☒ Kumar, Lalit (80) ☒ ☒ Lalit Kumar (79) ☒ ☒ Kumar, Lalit (79) ☒ ☒ Banerjee, Lalit Kumar (66) ☒ ☒ Pandit, Lalit Kumar (66) ☒ ☒ Modi, Lalit Kumar (66) ☒ ☒ Singh, Lalit Kumar (65) ☒ ☒ Sen, Lalit Kumar (65) ☒ ☒ Barua, Lalit Kumar (65) ☒ ☒ Create new item Search for match		
☆	🗨	4. Jemand ☒ ☒ Langewiesche, Wilhelm (26) ☒ ☒ Trömer, Johann Christian (11) ☒ ☒ Suttner, Bertha von (5) ☒ ☒ Create new item Search for match		

34

Abgleich / Reconciliation

openrefine.labs.lobid.org | localhost:3333

```
https://lobid.org/gnd/reconcile
```

(Auf openrefine.labs.lobid.org muss der Dienst nicht ergänzt werden)

Anreicherung

Anreicherung

Namen wurden mit GND-IDs assoziiert

Anreicherung

Namen wurden mit GND-IDs assoziiert

Felder der GND-Entitäten stehen so zur Verfügung

Anreicherung

Namen wurden mit GND-IDs assoziiert

Felder der GND-Entitäten stehen so zur Verfügung

Mit diesen können wir unsere lokalen Daten anreichern

Anreicherung

Namen wurden mit GND-IDs assoziiert

Felder der GND-Entitäten stehen so zur Verfügung

Mit diesen können wir unsere lokalen Daten anreichern

d.h. konkret: neue Spalten in unserer Tabelle anlegen

clipboard - OpenRefine - Mozilla Firefox

clipboard - OpenRefine x +

127.0.0.1:3333/project?project=2221362040394

OpenRefine clipboard Permalink

Open... Export Help

Facet / Filter Undo / Redo 2 / 2

Refresh Reset All Remove All

name: judgment change

2 choices Sort by: name count

matched 1

none 3

Facet by choice counts

name: best candidate's score change reset

19.00 — 41.00

4 rows

Show as: rows records Show: 5 10 25 50 rows « first < previous 1 - 4 next > last »

☒ All ☒ name ☐ beruf ☐ ort

1. ☐ ☐ 1. Informatiker Berlin

2. ☐ ☐ 2. Schriftsteller

3. ☐ ☐ 3. Split into several columns...
Add column based on this column...
Add column by fetching URLs...
Add columns from reconciled values...
Rename this column
Remove this column
Move column to beginning
Move column to end
Move column left
Move column right

4. Jemand
☒ ☒ Neigenfind, Rulf (
☒ ☒ Langewiesche, W
☒ ☒ Rantzau, Adeline
☒ ☒ Blaha, Marjana (10
☒ ☒ Laechel, Kurt (10)
☒ ☒ Kukulski, Jaroslav (10)
☒ ☒ Heidepeter, Lothar (10)
☒ ☒ Cammerhoff, Georg (10)
☒ ☒ Ney, Friederike (10)
☒ ☒ Greiszing, Alfred Harro (10)
☒ ☒ Create new item

clipboard - OpenRefine - Mozilla Firefox

clipboard - OpenRefine × +

localhost:3333/project?project=232754877

OpenRefine clipboard Permalink

Open... Export ▾ Help

Facet / Filter

Refresh

name: ju

2 choices

matched 3

none 1

Facet by choice

name: b

Reset

Add columns from reconciled column name

Add Property Preview

Ort name

Select an item from the list:

Ort	place
Wirkungsort	placeOfActivity
Geburtsort	placeOfBirth
Geburtsort (Literal)	placeOfBirthAsLiteral
Veranstaltungsort	placeOfConferenceOrEvent
Aufbewahrungsort	placeOfCustody
Sterbeort	placeOfDeath
Sterbeort (Literal)	placeOfDeathAsLiteral
Fundort	placeOfDiscovery
Exilort	placeOfExile
Herstellungsort	placeOfManufacture
Vorheriges Schlagwort	precedingSubject
In Beziehung stehendes Schlagwort	relatedSubjectHeadline

Geburtsort remove configure

Berlin

Florida, Mo.

OK Cancel

clipboard - OpenRefine - Mozilla Firefox

clipboard - OpenRefine

127.0.0.1:3333/project?project=2221362040394

Search

OpenRefine

Facet / Filter

Refresh

name: jud

2 choices

matched 1

none 3

Facet by choice

name: bes

Wikidata

next > last >>

Help

Add columns from reconciled column name

Add Property

Preview

Reset

Suggested Properties

Adelstitel

Adelstitel (Literal)

Affiliation

Affiliation (Literal)

Akademischer Grad

Alte Normnummer

Beruf oder Beschäftigung

Beruf oder Beschäftigung (Literal)

Beruflich Beziehung

Bevorzugter Name

Beziehung, Bekanntschaft, Freundschaft

Biografische oder historische Angaben

name	Beruf oder Beschäftigung remove configure	Geburtsort remove configure	Sterbeort remove configure	Ländercode remove configure
Weizenbaum, Joseph	Informatiker	Berlin	Berlin	USA
	Mathematiker			Deutschland
Twain, Mark	Lotse	Florida, Mo.	Redding, Conn.	USA
	Schriftsteller			
	Drucker			
	Journalist			
	Soldat			
Lalit Kumar				Sonstiges Ausland, Unbekanntes Land
<not reconciled>				

OK Cancel

clipboard - OpenRefine - Mozilla Firefox

127.0.0.1:3333/project?project=1790589415416

OpenRefine clipboard

Facet / Filter

Refresh

name: judgme

2 choices Sort by:

matched 2

none 2

Facet by choice count

name: best can

20

Add columns from reconciled column name

Add Property Preview

Ländercode

Suggested Properties

- Adelstitel
- Adelstitel (Literal)
- Affiliation
- Affiliation (Literal)
- Akademischer Grad
- Alte Normnummer
- Beruf oder Beschäftigung
- Beruf oder Beschäftigung (Literal)
- Berufliche Beziehung
- Bevorzugter Name
- Beziehung, Bekanntschaft, Freundschaft
- Biografische oder

name	Beruf oder Beschäftigung remove configure	Geburtsort remove configure	Sterbeort remove configure	Ländercode remove configure
Weizenbaum, Joseph	Informatiker	Berlin	Berlin	USA
	Mathematiker			Deutschland
				USA
				Sonstiges Ausland, Unbekanntes Land

Reset

Settings for geographicAreaCode

Limit: 0

Content:

☐ ID

☒ Literal

OK Cancel

OK Cancel

Anreicherung

```
id,name  
1,Haddaway  
2,Johannes Geßner  
3,Judith Kuckart  
4,Rosa Luxemburg  
5,Albert Einstein  
6,Hazel Brugger  
7,Anne Cuneo  
8,Vitus Huonder
```

Projekt anlegen, Namen abgleichen und Berufe anreichern

Anreicherung mit GND-Entitäten

Anreicherung mit GND-Entitäten

Angereicherte Felder sind z.T. GND-Entitäten

Anreicherung mit GND-Entitäten

Angereicherte Felder sind z.T. GND-Entitäten

Sind automatisch abgeglichen / reconciled

Anreicherung mit GND-Entitäten

Angereicherte Felder sind z.T. GND-Entitäten

Sind automatisch abgeglichen / reconciled

Können als Grundlage für weitere Anreicherung dienen

clipboard - OpenRefine

×

+

←

→

↺

🏠

127.0.0.1:3333/project?project=2221362040394

⋮

🔒

☆

📶

🔍 Search

⬇

|||

📄

☰

📦

OpenRefine clipboard

Permalink

Open...

Export ▾

Help

Facet / Filter

Undo / Redo 6 / 6

Refresh

Reset All

Remove All

✕ name: judgment

change

2 choices

Sort by: name count

(blank) 5

matched 3

Facet by choice counts

✕ name: best candidate's score

change reset

0

100

20.00 — 44.00

☑ Numeric

4

☐ Non-numeric

0

☑ Blank

2

☐ Error

0

3 matching records (4 total)

Extensions: Wikidata ▾

Show as: rows records

Show: 5 10 25 50 records

« first < previous 1 - 3 next > last »

▼ All	▼ name	▼ Beruf oder Besc	▼ Geburtsort	▼ Sterbeort	▼ Ländercode
☆ 🗨 1.	Weizenbaum, Joseph Choose new match	Informatiker Choose new match	Berlin Choose new match	Berlin Choose new match	USA
☆ 🗨		Mathematiker Choose new match			Deutschland
☆ 🗨 2.	Twain, Mark Choose new match	Lotse Choose new match	Florida, Mo. Choose new match	Redding, Conn. Choose new match	USA
☆ 🗨		Schriftsteller Choose new match			
☆ 🗨		Drucker Choose new match			
☆ 🗨		Journalist Choose new match			
☆ 🗨		Soldat Choose new match			
☆ 🗨 3.	Lalit Kumar Choose new match				Sonstiges Ausland, Unbekanntes Land

Anreicherung mit GND-Entitäten

Wirkungsort und Darstellung des Wirkungsortes anreichern

Open...

Export ▾

Help

8 records

Extensions: Wikidata ▾

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

« first < previous 1 - 8 next > last »

▼ All			▼ id	▼ name	▼ Beruf oder Beschäft	▼ Wirkungsort	▼ Darstellung des Wirkungsortes
☆	🗨	1.	1	Haddaway Choose new match	Komponist Choose new match	Köln Choose new match	http://commons.wikimedia.org/wiki/Special:FilePath/DEU%20Koeln%20COA.svg
☆	🗨				Sänger Choose new match	Kitzbühel Choose new match	http://commons.wikimedia.org/wiki/Special:FilePath/AUT%20Kitzb%20C3%BChel%20COA.svg
☆	🗨				Musiker Choose new match		
☆	🗨	2.	2	Geßner, Johannes Choose new match	Arzt Choose new match	Zürich Choose new match	http://commons.wikimedia.org/wiki/Special:FilePath/Zurich-coat%20of%20arms.svg
☆	🗨				Mathematiker Choose new match		
☆	🗨				Naturwissenschaftler Choose new match		
☆	🗨	3.	3	Kuckart, Judith Choose new match	Schriftstellerin Choose new match	Berlin Choose new match	https://commons.wikimedia.org/wiki/Special:FilePath/Cityscape_Berlin.jpg
☆	🗨				Regisseurin Choose new match	Zürich Choose new match	http://commons.wikimedia.org/wiki/Special:FilePath/Zurich-coat%20of%20arms.svg
☆	🗨				Tänzerin Choose new match		
☆	🗨				Choreografin Choose new match		

Datenquellen und Oberfläche von lobid-gnd

Datenabgleich und -anreicherung in OpenRefine

Möglichkeiten zur Verbesserung des Matching

Standardisierung & W3C Community Group

Eigene Daten in OpenRefine verfügbar machen

Ergebnisse
verbessern

Ergebnisse verbessern

Einschränkung der Typen der Zielentitäten

Ergebnisse verbessern

Einschränkung der Typen der Zielentitäten

Verwendung zusätzlicher lokaler Daten

Ergebnisse verbessern

Einschränkung der Typen der Zielentitäten

Verwendung zusätzlicher lokaler Daten

Gezielte Suche nach Treffern

Einschränkung der Entitätstypen

clipboard - OpenRefine - Mozilla Firefox

clipboard - OpenRefine

https://openrefine.labs.lobid.org/project?project=...

Search

Open... Export Help

OpenRefine clipboard

Reconcile column "name"

» Access Service API

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

☐ Normdatenressource
AuthorityResource

☐ Individualisierte Person
DifferentiatedPerson

☐ Körperschaft
CorporateBody

☐ Werk
Work

Also use relevant details from other columns:

Column Include? As Property

beruf ☐

ort ☐

☒ Reconcile against type

☐ Reconcile against no parent

☒ Auto-match candidates

Maximum number of candidates

Select an item from the list:

Werk der Musik	MusicalWork
Musikalische Körperschaft	MusicalCorporateBody
Fassung eines Werks der Musik	VersionOfAMusicalWork

Add Standard Service...

Start Reconciling Cancel

Verwendung zusätzlicher lokaler Daten

clipboard - OpenRefine

+

←

→

↻

🏠

🔒

openrefine.labs.lobid.org/proj

120%

⋮

🛡️

☆

🔍

Search

⬇️

⌵

📄

🖋️

🌐

✂️

☀️

🔗

☰

🔹

OpenRefine

clipboard

Permalink

Open...

Export ▾

Help

Facet / Filter

Undo / Redo 0 / 0

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](#)

4 rows

Extensions: Wikidata ▾

Show as: rows records

Show: 5 10 25 50 rows

« first < previous 1 - 4 next > last »

▼ All	▼ name	▼ beruf	▼ ort
☆ 🗨️ 1.	J. Weizenbaum	Informatiker	Berlin
☆ 🗨️ 2.	Twain, Mark	Schriftsteller	
☆ 🗨️ 3.	Kumar, Lalit		
☆ 🗨️ 4.	Jemand		

clipboard - OpenRefine - Mozilla Firefox

clipboard - OpenRefine

localhost:3333/project?project=232754877

Search

Open... Export Help

OpenRefine clipboard

Reconcile column "name"

» Access Service API

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

Also use relevant details from other columns:

☐ Normdatenressource
AuthorityResource

☐ Individualisierte Person
DifferentiatedPerson

☐ Werk
Work

☐ Körperschaft
CorporateBody

Column Include? As Property

beruf ☐ Beruf

Select an item from the list:

Beruf oder BeschäftigungprofessionOrOccupation

Beruf oder Beschäftigung (Literal)professionOrOccupationAsLiteral

Berufliche BeziehungprofessionalRelationship

☒ Reconcile against type: Person

☐ Reconcile against no particular type

☒ Auto-match candidates with high confidence

Maximum number of candidates to return

Add Standard Service...Start ReconcilingCancel

Gezielte Suche nach Treffern

clipboard - OpenRefine

+

←

→

↻

🏠

localhost:3333/project?project=232754877

...

🔒

☆

🔍

Search

⬇

|||

✍

📄

🌐

☰

🔷

OpenRefine clipboard

Permalink

Open...

Export ▾

Help

Facet / Filter

Undo / Redo 1 / 1

Refresh

Reset All

Remove All

✕ name: judgment

change

2 choices

Sort by: name count

matched 1

none 3

Facet by choice counts

✕ name: best candidate's score

change

19.00 — 41.00

4 rows

Extensions: Wikidata ▾

Show as: rows records

Show: 5 10 25 50 rows

« first < previous 1 - 4 next > last »

▼ All

▼ name

▼ beruf

▼ ort

1. J. Weizenbaum

✓ ✓ Weizenbaum, Joseph (36)

✓ ✓ Create new item

Search for match

2. Twain, Mark

Choose new match

Schriftsteller

Search for match

Search for "J. Weizenbaum"

☒ Match other cells with same content

☐ Match this cell only

Match

New Item

Don't Reconcile

J. Weizenbaum

Select an item from the list:

Weizenbaum, Joseph118820591

1923-2008 | Informatiker

WEIZENBAUM, J.187478929

4. Jemand

✓ ✓ Neigenfind, Rulf (20)

✓ ✓ Langewiesche, Wilhelm (15)

✓ ✓ Trömer, Johann Christian (10)

✓ ✓ Günther, Walter Wolfgang (10)

✓ ✓ Rapley, Belinda (10)

✓ ✓ Ullrich, Anne Kathrin (10)

✓ ✓ Heidepeter, Lothar (10)

✓ ✓ Kiffe, Johan (10)

✓ ✓ Cammerhoff, Georg (10)

✓ ✓ Seifert, Katharina (10)

✓ ✓ Create new item

Search for match

Beispiele

Beispiele

Verwendung zusätzlicher lokaler Daten

Beispiele

Verwendung zusätzlicher lokaler Daten

1. Beispiel: Lebensdaten

Beispiele

Verwendung zusätzlicher lokaler Daten

1. Beispiel: Lebensdaten

2. Beispiel: Beruf & Parteizugehörigkeit

1. Beispiel: Lebensdaten

 **OpenRefine** *A power tool for working with messy data.*

Create Project

Open Project

Import Project

Language Settings

Create a project by importing data. What kinds of data files can I import?

TSV, CSV, *SV, Excel (.xls and .xlsx), JSON, XML, RDF as XML, and Google Data documents with OpenRefine extensions.

Get data from

This Computer

Web Addresses (URLs)

Clipboard

Database

Google Data

Enter one or more web addresses (URLs) pointing to data to download:

https://blog.lobid.org/data/persons.csv

Add Another URL

Next »

CSV lokal speichern oder direkt per URL:
<https://blog.lobid.org/data/persons.csv>

OpenRefine - Mozilla Firefox

OpenRefine

localhost:3333

Search

OpenRefine

A power tool for working with messy data.

Create Project

Open Project

Import Project

Language Settings

« Start Over

Configure Parsing Options

Project name

persons csv

Tags

Create Project »

	Name	DateOfBirth	DateOfDeath	Type
1.	Keidel, Karl	1902	1981	personal
2.	Beckmann, Hannes	1909	1977	personal
3.	Klopper, Paul	1876	1967	personal
4.	Frank, Ellen	1904	1999	personal
5.	Lange, Annemarie	1907	1976	personal
6.	Hoff, August	1892	1971	personal
7.	Seböck, Stefan	1901	1942	personal
8.	Kunowski, Gertrud von	1877	1960	personal
9.	Tappolet, Bertha	1897	1947	personal
10.	Benscheidt, Karl	1888	1975	personal
11.	Reimann, Friedrich	1812		personal

Parse data as

CSV / TSV / separator-based files

Line-based text files

Fixed-width field text files

PC-Axis text files

JSON files

MARC files

JSON-LD files

RDF/N3 files

RDF/N-Triples files

RDF/Turtle files

Character encoding

Update Preview

Columns are separated by

☒ commas (CSV)

☐ tabs (TSV)

☐ custom: ,

Escape special characters with \

☐ Ignore first 0 line(s) at beginning of file

☒ Parse next 1 line(s) as column headers

☐ Discard initial 0 row(s) of data

☐ Load at most 0 row(s) of data

☒ Use " to enclose cells containing column separators

☐ Parse cell text

☒ Store blank rows

☐ Column names (comma separated):

Version 3.3 [dbc732c]

Preferences

Help

About

persons csv - OpenRefine - Mozilla Firefox

persons csv - OpenRefine

localhost:3333/project?project=2565457314865

OpenRefine persons csv

Reconcile column "Name"

» Access Service API

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

Also use relevant details from other columns:

Column	Include?	As Property
DateOfBirth	☐	
DateOfDeath	☐	
Type	☐	

☐ Normdatenressource
AuthorityResource

☒ Individualisierte Person
DifferentiatedPerson

☐ Reconcile against type:

☐ Reconcile against no particular type

☒ Auto-match candidates with high confidence

Maximum number of candidates to return

Add Standard Service...

Start Reconciling Cancel

25. Göhl, Ernst 1907 1944 personal

Project anlegen und Namen abgleichen

openrefine.labs.lobid.org | localhost:3333

persons.csv

 **OpenRefine** persons csv [Permalink](#)

Facet / Filter [Undo / Redo](#) 8 / 8

[Refresh](#) [Reset All](#) [Remove All](#)

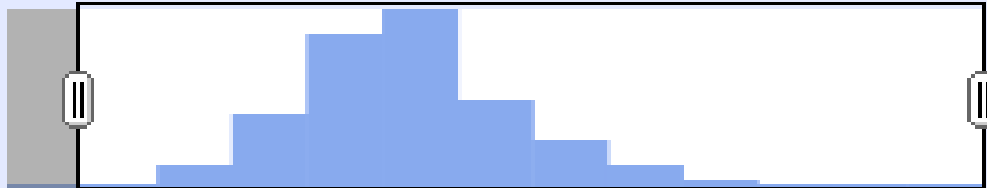
Name: judgment [change](#)

2 choices Sort by: [name](#) [count](#)

matched 976
none 409

[Facet by choice counts](#)

Name: best candidate's score [change](#) [reset](#)


40.00 — 160.00

1385 matching rows (1386 total)

Show as: [rows](#) [records](#) Show: [5](#) [10](#) [25](#) [50](#) rows

[All](#) [Name](#)

		1.	Keidel, Karl
☒	☒		Keidel, Carl (65)
☒	☒		Keidel, Friedrich (64)
☒	☒		Keidel, Walter (53)
☒	☒		Keidel, Christine (53)
☒	☒		Keidel, Margaret (53)
☒	☒		Keidel, Stephan (53)
☒	☒		Keidel, Sabine (53)
☒	☒		Keidel, Albert (53)
☒	☒		Keidel, Verena (53)
☒	☒		Keidel, Anke (53)
☒	☒		Create new item

[Search for match](#)

Ein erstes Ergebnis:
70% automatisch abgeglichen

Wir wollen die Lebensdaten
zur Disambiguierung nutzen

Wir wollen die Lebensdaten zur Disambiguierung nutzen

In unseren abzugleichenden Daten z.B. "1904"

Wir wollen die Lebensdaten zur Disambiguierung nutzen

In unseren abzugleichenden Daten z.B. "1904"

Womit gleichen wir ab? z.B. **Ellen Frank**

Wir wollen die Lebensdaten zur Disambiguierung nutzen

In unseren abzugleichenden Daten z.B. "1904"

Womit gleichen wir ab? z.B. **Ellen Frank**

```
"dateOfDeath": [  
  "1999-09-17"  
],  
"dateOfBirth": [  
  "1904-03-09"  
]
```


persons csv - OpenRefine

+

← → ↻ 🏠

localhost:3333/project?project=2565457314865

... 🛡️ ☆ 🔍 Search

⬇️ ||| 🔧 📖 🇺🇦 ☰

📊 OpenRefine persons csv [Permalink](#)

Open... Export ▾ Help

Facet / Filter Undo / Redo 1 / 1

Refresh Reset All Remove All

✕ Name: judgment change

2 choices Sort by: name count

matched 911

none 475

Facet by choice counts

✕ Name: best candidate's score change reset



20.00 — 150.00

1386 rows

Extensions: Wikidata ▾

Show as: rows records Show: 5 10 25 50 rows « first < previous 1 - 25 next > last »

		▼ All	▼ Name	▼ DateOfBirth	▼ DateOfDeath	▼ Type
☆	🗨	1.	Keidel, Karl ✓ ✓ Keidel, Carl (62) ✓ ✓ Keidel, Friedrich (61) ✓ ✓ Trutzer, Karl (50) ✓ ✓ Keidel, Margaret (50) ✓ ✓ Keidel, Sabine (50) ✓ ✓ Keidel, Stephan (50) ✓ ✓ Keidel, Albert (50) ✓ ✓ Keidel, Verena (50) ✓ ✓ Keidel, Walter (50) ✓ ✓ Keidel, Florian (50) ✓ ✓ Create new item Search for match			personal
☆	🗨	2.	Beckmann, Hannes ✓ ✓ Beckmann, Hannes (75) ✓ ✓ Beckmann, Hannes (75) ✓ ✓ Beckmann, ... (50) ✓ ✓ Beckmann (50) ✓ ✓ Beckmann, ... (50) ✓ ✓ Beckmann (50) ✓ ✓ Beckmann, ... (50) ✓ ✓ Beckmann, ... (50) ✓ ✓ Beckmann, ... (50) ✓ ✓ Beckmann, ... (50) ✓ ✓ Create new item Search for match	1909	1977	
☆	🗨	3.	Klopfer, Paul Hugo ✓ ✓ Create new item Choose new match	1876	1967	personal
☆	🗨	4.	Frank, Ellen ✓ ✓ Frank, Ellen (59) ✓ ✓ Frank, Ellen (59) ✓ ✓ Ellen, Ellen (44) ✓ ✓ Ellen, R. F. (39) ✓ ✓ Frank (33)	1904	1999	personal

javascript: {}

▼ DateOfBirth

Facet

Text filter

Edit cells ▾ Transform... ←

Edit column ▾ Common transforms ▾

Transpose ▾ Fill down

Sort... Blank down

View ▾ Split multi-valued cells...

Reconcile ▾ Join multi-valued cells...

Cluster and edit...

Replace

Expression Language

Expression Language

Für die Transformation verschiedene Sprachen

Expression Language

Für die Transformation verschiedene Sprachen

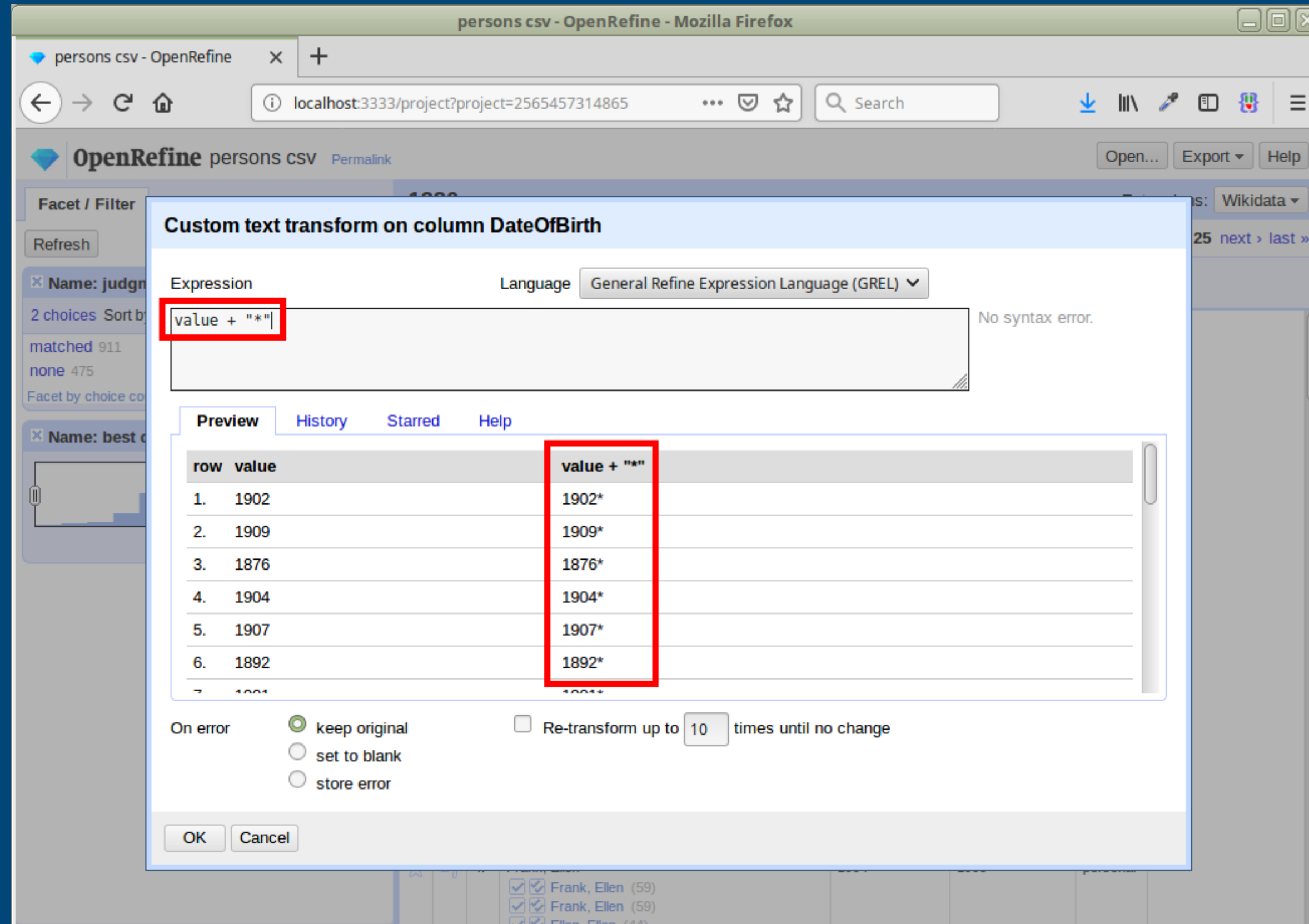
General Refine Expression Language (GREL)

Expression Language

Für die Transformation verschiedene Sprachen

General Refine Expression Language (GREL)

Python/Jython & Clojure



Wildcards: ? für ein Zeichen, * für beliebig viele Zeichen

Daten bereit für neuen,
verbesserten Abgleich

OpenRefine persons csv Permalink

Open... Export ▾ Help

Facet / Filter Undo / Redo 1 / 1

Refresh Reset All Remove All

1386 rows

Show as: rows records Show: 5 10 25 50 rows « first < previous 1 - 25 next > last »

Extensions: Wikidata ▾

Facet / Filter

Undo / Redo 1 / 1

Refresh Reset All Remove All

1. Name: judgment change

2 choices Sort by: name count

matched 911

none 475

Facet by choice counts

2. Name: best candidate's score change reset

20.00 — 150.00

1. Name

Facet

Text filter

Edit cells

Edit column

Transpose

Sort...

View

Reconcile

Start reconciling...

Facets

Actions

Copy reconciliation data...

Use values as identifiers

Match each cell to its best candidate

Create a new item for each cell

Create one new item for similar cells

Match all filtered cells to...

Discard reconciliation judgments

Clear reconciliation data

2. Beckmann, Hannes

Beckmann, Hanne

Beckmann, Hanne

Beckmann, ... (50)

Beckmann, ... (50)

Beckmann, ... (50)

Beckmann, ... (50)

Beckmann, ... (50)

Beckmann, ... (50)

Beckmann, ... (50)

Beckmann, ... (50)

Beckmann, ... (50)

Create new item

Vorherigen Abgleich zurücksetzen

persons csv - OpenRefine - Mozilla Firefox

persons csv - OpenRefine x +

localhost:3333/project?project=2565457314865

Reconcile column "Name" » Access Service API

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

☐ Normdatenressource
AuthorityResource

☒ Individualisierte Person
DifferentiatedPerson

Also use relevant details from other columns:

Column	Include?	As Property
DateOfBirth	☒	Geburtsdatum
DateOfDeath	☒	Ster

Select an item from the list:

Geburts- und Sterbedatum	dateOfBirthAndDeath
Sterbedatum	dateOfDeath
Östlichster Längengrad	easternmostLongitude
Erster Künstler	firstArtist
Erster Komponist	firstComposer
Nördlichster Breitengrad	northernmostLatitude
Sterbeort	placeOfDeath
Sterbeort (Literal)	placeOfDeathAsLiteral
Südlichster Breitengrad	southernmostLatitude
Westlichster Längengrad	westernmostLongitude

☐ Reconcile against type:

☐ Reconcile against no particular type

☒ Auto-match candidates with high confidence

Maximum number of candidates to return

Add Standard Service... Start Reconciling Cancel


OpenRefine persons csv [Permalink](#)

Facet / Filter
Undo / Redo 10 / 10

Refresh
Reset All
Remove All

✕
Name: judgment
change

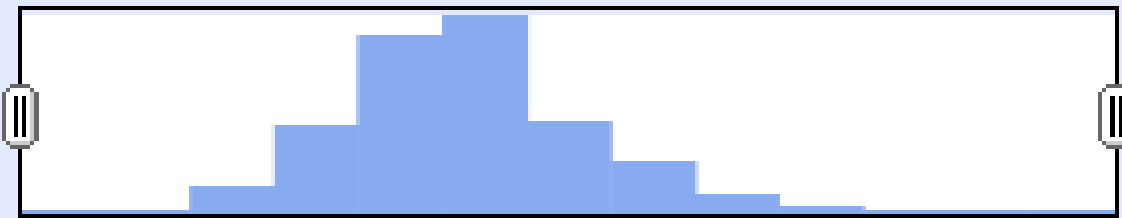
2 choices Sort by: name count

matched 1331
none 55

←

Facet by choice counts

✕
Name: best candidate's score
change reset



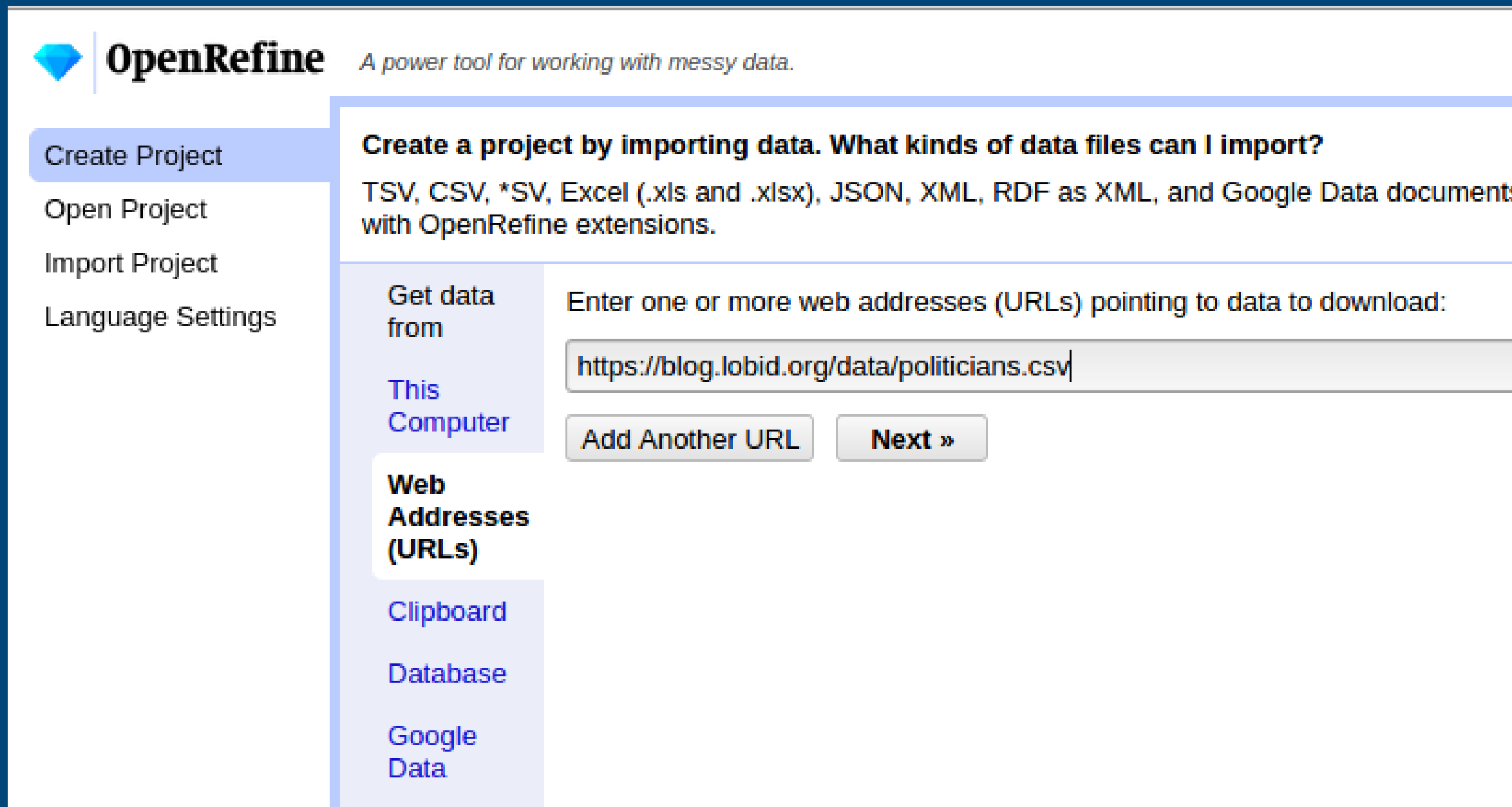
1386 rows

Show as: rows records Show: 5 10 25 50

▼ All		▼ Name		▼ DateOfB
☆	🗨	1.	Keidel, Carl Choose new match	1902*
☆	🗨	2.	Beckmann, Hannes Choose new match	1909*
☆	🗨	3.	Klopfer, Paul Choose new match	1876*
☆	🗨	4.	Frank, Ellen Choose new match	1904*
☆	🗨	5.	Lange, Annemarie Choose new match	1907*
☆	🗨	6.	Hoff, August Choose new match	1892*
☆	🗨	7.	Sebök, Stefan Choose new match	1901*

>95% automatisch abgeglichen

2. Beispiel: Beruf & Parteizugehörigkeit



OpenRefine *A power tool for working with messy data.*

Create Project
Open Project
Import Project
Language Settings

Create a project by importing data. What kinds of data files can I import?
TSV, CSV, *SV, Excel (.xls and .xlsx), JSON, XML, RDF as XML, and Google Data documents with OpenRefine extensions.

Get data from

This Computer

Web Addresses (URLs)

Clipboard

Database

Google Data

Enter one or more web addresses (URLs) pointing to data to download:

<https://blog.lobid.org/data/politicians.csv>

Add Another URL Next »

CSV lokal speichern oder direkt per URL:
<https://blog.lobid.org/data/politicians.csv>

OpenRefine - Mozilla Firefox

OpenRefine

localhost:3333

Search

OpenRefine

A power tool for working with messy data.

Create Project

Open Project

Import Project

Language Settings

« Start Over

Configure Project name

politicians csv

Tags

Create Project »

	label	parliamentary_group	GND_Kennung	Wikidata_ID
1.	Günter Rexrodt	NA	NA	NA
2.	Hans-Eberhard Urbaniak	SPD	NA	NA
3.	Ernst Schwanhold	SPD	NA	NA
4.	Kurt Rossmanith	CDU	NA	NA
5.	Margareta Wolf	GRUENE	1103834029	Q1894924
6.	Rita Süßmuth	NA	NA	NA
7.	Otto Graf Lambsdorff	FDP	NA	NA
8.	Rolf Kutzmutz	PDS	NA	NA
9.	Paul Krüger	CDU	NA	NA
10.	Rolf Schwanitz	SPD	NA	NA
11.	Christa Luft	PDS	NA	NA
12.	Anke Fuchs	SPD	NA	NA

Parse data as

CSV / TSV / separator-based files

Line-based text files

Fixed-width field text files

PC-Axis text files

JSON files

MARC files

JSON-LD files

RDF/N3 files

RDF/N-Triples files

RDF/Turtle files

Character encoding

Columns are separated by

☒ commas (CSV)

☐ tabs (TSV)

☐ custom: ,

Escape special characters with \

Update Preview

☐ Ignore first

0 line(s) at beginning of file

☒ Parse next

1 line(s) as column headers

☐ Discard initial

0 row(s) of data

☐ Load at most

0 row(s) of data

☒ Use character

" to enclose cells containing

Version 3.3 [cf578cf]

Preferences

Help

About

politicians csv - OpenRefine - Mozilla Firefox

politicians csv - OpenRefine

localhost:3333/project?project=2116118209

Search

Open... Export Help

Reconcile column "label"

» Access Service API

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

☐ Normdatenressource
AuthorityResource

☒ Individualisierte Person
DifferentiatedPerson

Also use relevant details from other columns:

Column	Include?	As Property
parliamentary_group	☐	
GND_Kennung	☐	
Wikidata_ID	☐	

☐ Reconcile against type:

☐ Reconcile against no particular type

☒ Auto-match candidates with high confidence

Maximum number of candidates to return

Add Standard Service... Start Reconciling Cancel

javascript:{}

25. Volker Kauder CDU NA NA

Project anlegen und Namen abgleichen

openrefine.labs.lobid.org | localhost:3333

[politicians.csv](#)

OpenRefine politicians csv [Permalink](#)

Facet / Filter [Undo / Redo 1 / 1](#)

[Refresh](#) [Reset All](#) [Remove All](#)

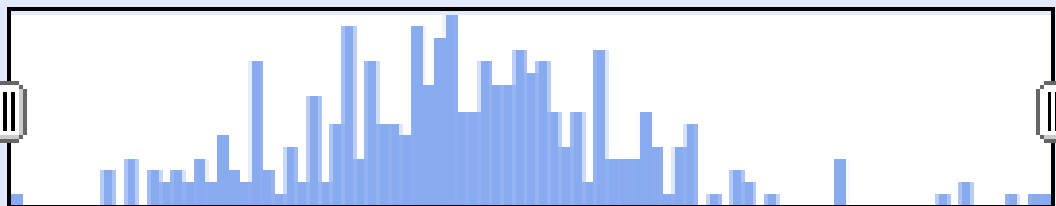
label: judgment [change](#)

2 choices Sort by: **name** count

matched 212
none 146

[Facet by choice counts](#)

label: best candidate's score [change](#) [reset](#)



41.00 — 132.00

☒ **Numeric** 357 ☐ **Non-numeric** 0 ☒ **Blank** 1

358 rows

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

All		label	
☆	🗨	1.	Rexrodt, Günter Choose new match
☆	🗨	2.	Urbaniak, Hans-Eberhard Choose new match
☆	🗨	3.	Schwanhold, Ernst Choose new match
☆	🗨	4.	Kurt Rossmanith ☒ ☒ Rossmanith, Kurt (85) ☒ ☒ Rossmanith, Kurt (82) ☒ ☒ Rossmanith, Alois (57) ☒ ☒ Rossmanith, Peter (57) ☒ ☒ Rossmanith, Sabine (57) ☒ ☒ Rossmanith, Christian (57) ☒ ☒ Rossmanith, Sandra (56) ☒ ☒ Rossmanith, Jonas (56)

Ein erstes Ergebnis:

< 60% automatisch abgeglichen

Wir wollen die
Parteizugehörigkeit nutzen

Wir wollen die Parteizugehörigkeit nutzen

In unseren abzugleichenden Daten z.B. SPD

Wir wollen die Parteizugehörigkeit nutzen

In unseren abzugleichenden Daten z.B. SPD

Womit gleichen wir ab? z.B. **Willy Brandt**

Wir wollen die Parteizugehörigkeit nutzen

In unseren abzugleichenden Daten z.B. SPD

Womit gleichen wir ab? z.B. **Willy Brandt**

```
"affiliation": [  
  {  
    "id": "https://d-nb.info/gnd/2022139-3",  
    "label": "Sozialdemokratische Partei Deutschlands"  
  }  
]
```

Abkürzungen auf eindeutige Identifizier mappen

Welche Abkürzungen kommen in den Daten vor?

OpenRefine politicians.csv Permalink

Open... Export Help

Facet / Filter Undo / Redo 1 / 5

Refresh Reset All Remove All

parliamentary_group change

8 choices Sort by: name count Cluster

CDU 118
FDP 50
fraktionslos 1
GRUENE 30
LINKE 7
NA 31
PDS 11
SPD 110

Facet by choice counts

358 rows

Show as: rows records Show: 5 10 25 50 rows « first < previous 1 - 10 next > last »

label parliamentary_g GND_Kennung Wikidata_ID

Facet Text facet

Text filter Numeric facet

Edit cells Timeline facet

Edit column Scatterplot facet

Transpose Custom text facet...

Sort... Custom Numeric Facet...

View Customized facets

Reconcile

1.	Rexrodt, Günter			
Choose new match				
2.	Hans-Eberhard Urbaniak			
	Urbaniak, Regina (53)			
	Urbaniak, Jan (52)			
	Urbaniak, Beata (52)			
	Urbaniak, Tamina (52)			
	Urbaniak, Rafal (52)			
	Urbaniak, Petra (52)			
	Urbaniak, Gertrud (51)			
	Urbaniak, Thomas (51)			
	Urbaniak, James (51)			
	Urbaniak, Michal (51)			
	Create new item			
Search for match				
3.	Schwanhold, Ernst	SPD	NA	NA
Choose new match				
4.	Kurt Rossmanith	CDU	NA	NA
	Rossmanith, Kurt (79)			
	Rossmanith, Kurt (78)			
	Rossmanith, Jonas (53)			
	Rossmanith, Frieda (53)			
	Roßmanith, Birgit (53)			
	Rossmanith, Sandra (53)			
	Rossmanith, Peter (52)			
	Rossmanith, Alois (52)			
	Rossmanith, Sabine (52)			
	Roßmanith, Richard (52)			
	Create new item			
Search for match				

Wie finden wir die GND-Nummern? z.B. SPD:

 gnd Erkunden API

sozialdemokratische partei deutschlands

1

Treffer pro Seite: 10 30 50 100

665 Treffer, zeige 1 bis 10:

	Sozialdemokratische Partei Deutschlands. Parteivorstand	Körperschaft	Deutschland	4479218-9
	Sozialdemokratische Partei Deutschlands. Parteitag	Kongressfolge oder Veranstaltungsfolge	Deutschland	7576523-8
	Unabhängige Sozialdemokratische Partei Deutschlands	Körperschaft	Deutschland, Deutsches Reich	2023731-5
	Sozialdemokratische Partei Deutschlands	Körperschaft	Deutschland, Deutsches Reich, Deutschland	2022139-3

Ergebnisse eingrenzen:

Entitätstyp

 Körperschaft (630) 

 Organ einer Körperschaft (21)

 Konferenz oder Veranstaltung (25) 

 Kongressfolge oder Veranstaltungsfolge (2)

 Werk (10)

GND-Sachgruppe 

Parteien, Politische Organisationen

Filter entfernen 

Gruppe, Organisationssoziologie, Interaktion (11)

Innenpolitik (3)

Schule (2)

Bevölkerung, Sozialstruktur, Soziale Situation, Soziale Bewegungen (2)

2

3

Gesuchte Mappings

- SPD → <https://d-nb.info/gnd/???>
- CDU → <https://d-nb.info/gnd/???>
- GRUENE → <https://d-nb.info/gnd/???>
- FDP → <https://d-nb.info/gnd/???>
- LINKE → <https://d-nb.info/gnd/???>
- PDS → <https://d-nb.info/gnd/???>

Mappings

- SPD → <https://d-nb.info/gnd/2022139-3>
- CDU → <https://d-nb.info/gnd/7230-8>
- GRUENE → <https://d-nb.info/gnd/2124337-2>
- FDP → <https://d-nb.info/gnd/37037-X>
- LINKE → <https://d-nb.info/gnd/10173685-X>
- PDS → <https://d-nb.info/gnd/5010217-5>

In der Text-Facette können wir alle Vorkommen der Werte entsprechend unseres Mappings editieren:

 **OpenRefine** politicians csv [Permalink](#)

Facet / Filter [Undo / Redo 3 / 3](#)

[Refresh](#) [Reset All](#) [Remove All](#)

parliamentary_group [change](#)

8 choices Sort by: name count [Cluster](#)

fraktionslos 1

GRUENE 30 [edit](#) [include](#)

[http://d-nb.info/gnd/37037-X](#) 50

[http://d-nb.info/gnd/7230-8](#) 118

LINKE 7

NA 31

PDS 11

SPD 110

Facet by choice counts

358 rows

Show as: rows records Show: 5 10 25 50

All **label**

☐	☐	1.	Rexrodt, Günter Choose new match
☐	☐	2.	Hans-Eberhard Urbaniak ☒ ☒ Urbaniak, Regina (53) ☒ ☒ Urbaniak, Jan (52) ☒ ☒ Urbaniak, Beata (52) ☒ ☒ Urbaniak, Tamina (52) ☒ ☒ Urbaniak, Rafal (52) ☒ ☒ Urbaniak, Petra (52) ☒ ☒ Urbaniak, Gertrud (51) ☒ ☒ Urbaniak, Thomas (51) ☒ ☒ Urbaniak, James (51) ☒ ☒ Urbaniak, Michał (51) ☒ ☒ Create new item Search for match

Diese manuelle Schritte können in der Undo/Redo-History extrahiert werden:

OpenRefine politicians csv [Permalink](#)

Facet / Filter **Undo / Redo** 8 / 8

Extract... Apply...

Filter:

0. Create project

1. Reconcile cells in column label to type **DifferentiatedPerson**

2. Create new column affiliation based on column parliamentary_group by filling 358 rows with **grel:value**

318 matching rows (358 total)

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

▼ All ▼ label

☆	🗨	1.	Rexrodt, Günter Choose new match
☆	🗨	2.	Urbaniak, Hans-Eberhard Choose new match
☆	🗨	3.	Schwanhold, Ernst Choose new match
☆	🗨	4.	Kurt Rossmanith ☒ ☒ Rossmanith, Kurt (87) ☒ ☒ Rossmanith, Kurt (85) ☒ ☒ Rossmanith, Kurt (86)

Wir bekommen die Editierhistorie als JSON:

```
[
  {
    "op": "core/column-addition",
    "engineConfig": {
      "facets": [],
      "mode": "row-based"
    },
    "baseColumnName": "parliamentary_group",
    "expression": "grel:value",
    "onError": "set-to-blank",
    "newColumnName": "affiliation",
    "columnInsertIndex": 2,
    "description": "Create column affiliation at index 2 based
on column parliamentary_group using expression grel:value"
  },
  ...
]
```

Wir bekommen die Editierhistorie als JSON:

```
[
  {
    "op": "core/column-addition",
    "engineConfig": {
      "facets": [],
      "mode": "row-based"
    },
    "baseColumnName": "parliamentary_group",
    "expression": "grel:value",
    "onError": "set-to-blank",
    "newColumnName": "affiliation",
    "columnInsertIndex": 2,
    "description": "Create column affiliation at index 2 based
on column parliamentary_group using expression grel:value"
  },
]
```

(Hier setzt auch die erwähnte Möglichkeit zur Automatisierung an: die manuell von der Fachebene in der Oberfläche erstellten Schritte können so ohne Oberfläche automatisiert werden)

Operationen können angepasst und reproduziert werden:

The screenshot shows the OpenRefine web interface for a project named 'politicians csv'. The top bar includes the OpenRefine logo, the project name, and a 'Permalink' link. Below this, there are tabs for 'Facet / Filter' and 'Undo / Redo 1 / 1'. A red arrow points to the '358 rows' count, and a red box highlights the 'Apply...' button next to the 'Extract...' button. The left sidebar shows a list of operations: '0. Create project' and '1. Reconcile cells in column label to type DifferentiatedPerson'. The main area displays a table of data with columns for 'All' and 'label'. The table shows three rows of data, each with a star icon, a comment icon, and a list of matches. The first row is 'Rexrodt, Günter' with a 'Choose new match' link. The second row is 'Hans-Eberhard Urbaniak' with a list of matches including 'Urbaniak, Regina (53)', 'Urbaniak, Jan (52)', 'Urbaniak, Beata (52)', 'Urbaniak, Tamina (52)', 'Urbaniak, Rafal (52)', 'Urbaniak, Petra (52)', 'Urbaniak, Gertrud (51)', 'Urbaniak, Thomas (51)', 'Urbaniak, James (51)', 'Urbaniak, Michał (51)', and 'Create new item'. The third row is 'Schwanhold, Ernst' with a 'Choose new match' link.

OpenRefine politicians csv [Permalink](#)

Facet / Filter Undo / Redo 1 / 1

Extract... Apply...

358 rows

Show as: rows records Show: 5 10 25 50

Filter:

0. Create project

1. Reconcile cells in column label to type DifferentiatedPerson

All	label
☆	1. Rexrodt, Günter Choose new match
☆	2. Hans-Eberhard Urbaniak ☒ ☒ Urbaniak, Regina (53) ☒ ☒ Urbaniak, Jan (52) ☒ ☒ Urbaniak, Beata (52) ☒ ☒ Urbaniak, Tamina (52) ☒ ☒ Urbaniak, Rafal (52) ☒ ☒ Urbaniak, Petra (52) ☒ ☒ Urbaniak, Gertrud (51) ☒ ☒ Urbaniak, Thomas (51) ☒ ☒ Urbaniak, James (51) ☒ ☒ Urbaniak, Michał (51) ☒ ☒ Create new item Search for match
☆	3. Schwanhold, Ernst Choose new match

politicians csv - OpenRefine - Mozilla Firefox

politicians csv - OpenRefine × +

localhost:3333/project?project=2116118209

Open... Export ▾ Help

Facet / Filter

Wikidata ▾

next > last >>

Filter:

0. Create project

1. Reconcile

Different

ata_ID

Apply Operation History

Paste an extracted JSON history of operations to perform:

```
[{"fromError": false, "to": "http://d-nb.info/gnd/7230-8"}, {"from": ["SPD"], "fromBlank": false, "fromError": false, "to": "http://d-nb.info/gnd/2022139-3"}, {"from": ["LINKE"], "fromBlank": false, "fromError": false, "to": "http://d-nb.info/gnd/10173685-X"}, {"from": ["PDS"], "fromBlank": false, "fromError": false, "to": "http://d-nb.info/gnd/5010217-5"}], {"description": "Mass edit cells in column affiliation"}]
```

Perform Operations Cancel

6. Süßmuth, Rita
Choose new match

NA NA NA

Daten bereit für neuen,
verbesserten Abgleich

politicians csv - OpenRefine - Mozilla Firefox

politicians csv - OpenRefine × +

localhost:3333/project?project=2116118209

OpenRefine politicians csv Permalink

Open... Export Help

Facet / Filter Undo / Redo 3 / 7

Refresh Reset All Remove All

358 rows

Show as: rows records Show: 5 10 25 50 rows « first < previous 1 - 25 next > last »

Extensions: Wikidata

label: judgment change

2 choices Sort by: name count

matched 185

none 173

Facet by choice counts

label: best candidate's score change reset

28.00 — 109.00

Numeric 357 Non-numeric 0 Blank 1

Error 0

All label parliamentary_g affiliation GND_Kennung

1. NA NA NA

2. SPD http://d-nb.info/gnd/2022139-3 NA

3. Schwanhold, Ernst

4. Kurt Rossmannith

5. Wolf, Margareta GRUENE http://d-nb.info/gnd/2124337-2 1103834029

6. Süßmuth, Rita NA NA

Facet

Text filter

Edit cells

Edit column

Transpose

Sort...

View

Reconcile

Start reconciling...

Facets

Actions

Copy reconciliation data...

Use values as identifiers

Match each cell to its best candidate

Create a new item for each cell

Create one new item for similar cells

Match all filtered cells to...

Discard reconciliation judgments

Clear reconciliation data

politicians csv - OpenRefine - Mozilla Firefox

politicians csv - OpenRefine × +

localhost:3333/project?project=2116118209

Open... Export ▾ Help

Reconcile column "label"

» Access Service API

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

- ☐ Normdatenressource
AuthorityResource
- ☒ Individualisierte Person
DifferentiatedPerson

Also use relevant details from other columns:

Column	Include?	As Property
parliamentary_group	☐	
affiliation	☒	af

Select an item from the list:

- Beziehung, Bekanntschaft, Freundschaft acquaintanceshipOrFriendship
- Affiliation** affiliation
- Affiliation (Literal) affiliationAsLiteral
- Biografische oder historische Angaben biographicalOrHistoricalInformatio
- In Beziehung stehende Körperschaft relatedCorporateBody
- In Beziehung stehendes Geografikum relatedPlaceOrGeographicName

☐ Reconcile against type:

☐ Reconcile against no particular type

☒ Auto-match candidates with high confidence

Maximum number of candidates to return

Add Standard Service... Start Reconciling Cancel

javascript:{} 24. Ingrid Matthäus-Maier SPD http://d-nb.info/gnd/2022139-3 NA NA

 **OpenRefine** politicians csv [Permalink](#)

Facet / Filter [Undo / Redo 5 / 5](#)

[Refresh](#) [Reset All](#) [Remove All](#)

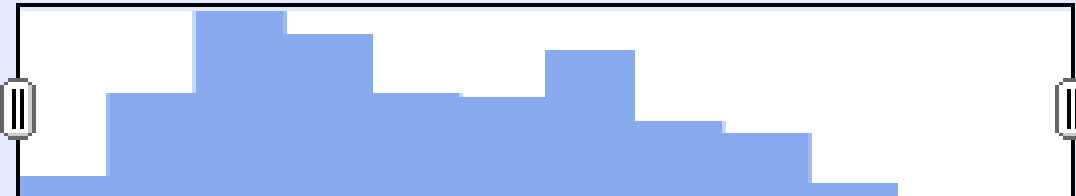
label: judgment [change](#)

2 choices Sort by: [name](#) [count](#)

matched 264
none 94

[Facet by choice counts](#)

label: best candidate's score [change](#) [reset](#)



50.00 — 170.00

☒ **Numeric** 357 ☐ **Non-numeric** 0 ☒ **Blank** 1

358 rows

Show as: [rows](#) [records](#) Show: [5](#) [10](#) [25](#) [50](#) rows

All **label**

		1.	Rexrodt, Günter Choose new match
		2.	Urbaniak, Hans-Eberhard Choose new match
		3.	Schwanhold, Ernst Choose new match
		4.	Kurt Rossmanith ☒ ☒ Rossmanith, Kurt (85) ☒ ☒ Rossmanith, Kurt (82) ☒ ☒ Rocker, Kurt (64) ☒ ☒ Pfeiffer, Kurt (63) ☒ ☒ Thürk, Kurt (63) ☒ ☒ Schmücker, Kurt (63) ☒ ☒ Malangré, Kurt (63) ☒ ☒ Kiesinger, Kurt Georg (59)

Zwischenergebnis: > 70% automatisch abgeglichen

Implizites Wissen nutzen

Implizites Wissen nutzen

Hintergrundwissen, das so gar nicht in den Daten ist

Implizites Wissen nutzen

Hintergrundwissen, das so gar nicht in den Daten ist

Formal ergänzen für die Disambiguierung

Implizites Wissen nutzen

Hintergrundwissen, das so gar nicht in den Daten ist

Formal ergänzen für die Disambiguierung

Hier z.B. wissen wir, dass es sich um Politiker:innen handelt

Beruf ergänzen

Beruf ergänzen

D.h. wir ergänzen eine neue Spalte mit dem Beruf

Beruf ergänzen

D.h. wir ergänzen eine neue Spalte mit dem Beruf

Mit dem Wert `Politik*`

Beruf ergänzen

D.h. wir ergänzen eine neue Spalte mit dem Beruf

Mit dem Wert `Politik*`

Über die Wildcard decken wir mehrere Formen ab: Politiker, Politikerin, Politikwissenschaftler, Politikwissenschaftlerin

 **OpenRefine** politicians.csv [Permalink](#)

Facet / Filter

Undo / Redo 5 / 5

Refresh

Reset All

Remove All

×

label: judgment

change

2 choices Sort by: name count

matched 264

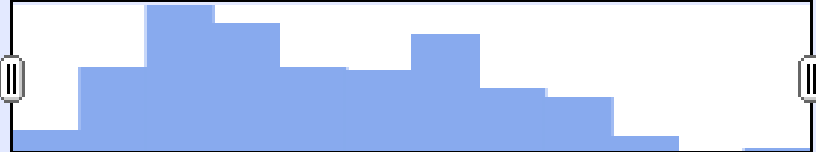
none 94

Facet by choice counts

×

label: best candidate's score

change reset



50.00 — 170.00

☒ Numeric 357

☐ Non-numeric 0

☒ Blank 1

358 rows

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

▼ All	▼ label	▼ parliamentary_g	▼ affiliation
☆ 1.		NA	
☆ 2.		SPD	https://d-nb.info/g
☆ 3.			b.info/g
☆ 4.			b.info/g

Facet

Text filter

Edit cells

Edit column

Transpose

Sort...

View

Reconcile

Split into several columns...

Join columns...

Add column based on this column...

Add column by fetching URLs...

Add columns from reconciled values...

Rename this column

Remove this column

94

politicians csv - OpenRefine - Mozilla Firefox

politicians csv - OpenRefine × +

localhost:3333/project?project=2116118209

Search

Open... Export ▾ Help

Facet / Filter

Using facets a

Use facets and fi

of your data to ac

and filter methods

the top of each d

Not sure how to g

Watch these scr

Add column based on column label

New column name

On error ☒ set to blank ☐ store error ☐ copy value from original column

Expression Language | | | |
| --- | --- | --- |
| 1. | Günter Rexrodt | Politik* |
| 2. | Hans-Eberhard Urbaniak | Politik* |
| 3. | Ernst Schwanhold | Politik* |
| 4. | Kurt Rossmanith | Politik* |
| 5. | Margareta Wolf | Politik* |
| 6. | Rita Süßmuth | Politik* |
| 7. | Olaf Scholz | Politik* |

OK Cancel

5. Wolf, Margareta
Choose new match

6. Süßmuth, Rita

GRUENE

http://d-nb.info/gnd/2124337-2

1103834029

NA

Daten bereit für neuen,
verbesserten Abgleich

Daten bereit für neuen, verbesserten Abgleich

Vorherigen Abgleich zurücksetzen

Daten bereit für neuen, verbesserten Abgleich

Vorherigen Abgleich zurücksetzen

Beruf & Affiliation mitschicken

politicians csv - OpenRefine - Mozilla Firefox

politicians csv - OpenRefine × +

localhost:3333/project?project=2116118209

OpenRefine politicians csv Permalink

Open... Export ▾ Help

Reconcile column "label"

» Access Service API

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

- ☐ Normdatenressource
AuthorityResource
- ☒ Individualisierte Person
DifferentiatedPerson

Also use relevant details from other columns:

Column	Include?	As Property
occupation	☒	Beruf oder Beschäftigung
parliamentary_group	☐	
affiliation	☒	af

Select an item from the list:

- Beziehung, Bekanntschaft, Freundschaft acquaintanceshipOrFriendship
- Affiliation** affiliation
- Affiliation (Literal) affiliationAsLiteral
- Biografische oder historische Angaben biographicalOrHistoricalInformatio
- In Beziehung stehende Körperschaft relatedCorporateBody
- In Beziehung stehendes Geografikum relatedPlaceOrGeographicName

☐ Reconcile against type:

☐ Reconcile against no particular type

☒ Auto-match candidates with high confidence

Maximum number of candidates to return

Add Standard Service... Start Reconciling Cancel

javascript:{} 18. Gisela Babel Politik* FDP http://d-nb.info 122410262

 **OpenRefine** politicians csv [Permalink](#)

Facet / Filter [Undo / Redo 8 / 8](#)

[Refresh](#) [Reset All](#) [Remove All](#)

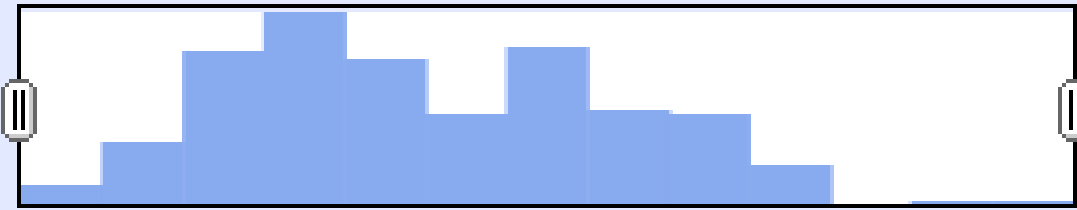
label: judgment [change](#)

2 choices Sort by: [name](#) [count](#)

matched 292
none 66

[Facet by choice counts](#)

label: best candidate's score [change](#) [reset](#)



50.00 — 180.00

☒ **Numeric** 357 ☐ **Non-numeric** 0 ☒ **Blank** 1

358 rows

Show as: **rows** [records](#) Show: **5** [10](#) [25](#) [50](#) rows

All		label	
☆	🗨	1.	Rexrodt, Günter Choose new match
☆	🗨	2.	Urbaniak, Hans-Eberhard Choose new match
☆	🗨	3.	Schwanhold, Ernst Choose new match
☆	🗨	4.	Kurt Rossmanith ☒ ☒ Rossmanith, Kurt (87) ☒ ☒ Rossmanith, Kurt (85) ☒ ☒ Rocker, Kurt (69) ☒ ☒ Pfeiffer, Kurt (68) ☒ ☒ Thürk, Kurt (68) ☒ ☒ Schmücker, Kurt (68) ☒ ☒ Malangré, Kurt (68) ☒ ☒ Kiesinger, Kurt Georg (64)

> 80% automatisch abgeglichen

Export

Export

Zahlreiche eingebaute, anpassbare Exportformate

Export

Zahlreiche eingebaute, anpassbare Exportformate

Erweiterte Funktionalität durch Extensions

Export

Zahlreiche eingebaute, anpassbare Exportformate

Erweiterte Funktionalität durch Extensions

z.B. für Wikidata, vergleichbar für GND denkbar

Open...Export ▼Help

data ▼last »

8 records

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

▼ All	▼ id	▼ name	▼ Beruf oder Beschäftigung	▼ Wirkungsort
☆	1.	1	Haddaway Choose new match	Köln Choose new match
☆			Sänger Choose new match	Kitzbühel Choose new match
☆			Musiker Choose new match	
☆	2.	2	Geßner, Johannes Choose new match	Zürich Choose new match
☆			Mathematiker Choose new match	
☆			Naturwissenschaftler Choose new match	
☆	3.	3	Kuckart, Judith Choose new match	Berlin Choose new match
☆			Regisseurin Choose new match	Zürich Choose new match
☆			Tänzerin Choose new match	
☆			Choreografin Choose new match	

OpenRefine project archive to file

Tab-separated value

Comma-separated value

HTML table

Excel (.xls)

Excel 2007+ (.xlsx)

ODF spreadsheet

Custom tabular exporter...

SQL Exporter...

Templating...

OpenRefine project archive to Google Drive...

Google Sheets

Wikidata edits...

QuickStatements file

Wikidata schema

/Zurich-coat%20of%20arms.svg

Open...Export ▾Help

8 records

Extensions: Wikidata ▾

Show as: rows recordsShow: 5102550 records

▼ All	▼ id	▼ name	▼ Beruf oder Beschäftigung	▼ Wirkungsort	▼ D	
★	1.	1	Haddaway Choose new match	Komponist Choose new match	Köln Choose new match	http://DEU
★			Sänger Choose new match	Kitzbühel Choose new match	http://AUT	
★			Musiker Choose new match			
★	2.	2	Geßner, Johannes Choose new match	Arzt Choose new match	Zürich Choose new match	http://commons.wikimedia.org/wiki/Special:FilePath/Zurich-coat%20of%20arms.svg
★			Mathematiker Choose new match			
★			Naturwissenschaftler Choose new match			
★	3.	3	Kuckart, Judith Choose new match	Schriftstellerin Choose new match	Berlin Choose new match	https://commons.wikimedia.org/wiki/Special:FilePath/Cityscape_Berlin.jpg
★			Regisseurin Choose new match		Zürich Choose new match	http://commons.wikimedia.org/wiki/Special:FilePath/Zurich-coat%20of%20arms.svg
★			Tänzerin Choose new match			
★			Choreografin Choose new match			

Edit Wikidata schema

Manage Wikidata account

Import schema

Export schema

Upload edits to Wikidata

Export to QuickStatements

Flexibler Export mit Templating

nk

Open... Export ▼ Help

318 matching rows (358 total)

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

▼ All	▼ label	▼ occupation	▼ parliamentary_c
☆	1. Rexrodt, Günter Choose new match	Politik*	NA
☆	2. Urbaniak, Hans-Eberhard Choose new match	Politik*	SPD
☆	3. Schwanhold, Ernst Choose new match	Politik*	SPD
☆	4. Rossmanith, Kurt Choose new match	Politik*	CDU
☆	5. Wolf, Margareta Choose new match	Politik*	GRUENE
☆	6. Süßmuth, Rita Choose new match	Politik*	NA
☆	8. Kutzmutz, Rolf Choose new match	Politik*	PDS
☆	9. Krüger, Paul Choose new match	Politik*	CDU
☆	10. Schwanitz, Rolf Choose new match	Politik*	SPD
☆	11. Luft, Christa Choose new match	Politik*	PDS

OpenRefine project archive to file

Tab-separated value

Comma-separated value

HTML table

Excel (.xls)

Excel 2007+ (.xlsx)

ODF spreadsheet

Custom tabular exporter...

SQL Exporter

Templating...

OpenRefine project archive to Google Drive...

Google Sheets

Wikidata edits...

QuickStatements file

Wikidata schema

data ▼

last »

ung ▼

NA

NA

NA

NA

Q188

NA

NA

NA

NA

https://d-nb.info/gnd/5010217-5

NA

NA

Facet / Filter

Refresh

✕ label: judge

2 choices Sort by

matched 252

none 66

Facet by choice count

✕ label: best

☒ Numeric
357☐ Error
0

Templating Export

Prefix

1

```
{  
  "rows" : [  

```

Row Template

2

```
{  
  "label" : {{jsonize(cells["label"].value)}},  
  "occupation" : {{jsonize(cells["occupation"].value)}},  
  "parliamentary_group" : {{jsonize(cells["parliamentary_group"].value)}},  
  "affiliation" : {{jsonize(cells["affiliation"].value)}},  
  "GND_Kennung" : {{jsonize(cells["GND_Kennung"].value)}},  
  "Wikidata_ID" : {{jsonize(cells["Wikidata_ID"].value)}}  
}
```

Row Separator

3

```
,
```

Suffix

4

```
]  
}
```

Reset Template

5

```
{  
  "rows" : [  
    {  
      "label" : "Günter Rexrodt",  
      "occupation" : "Politik*",  
      "parliamentary_group" : "NA",  
      "affiliation" : "",  
      "GND_Kennung" : "NA",  
      "Wikidata_ID" : "NA"  
    },  
    {  
      "label" : "Hans-Eberhard Urbaniak",  
      "occupation" : "Politik*",  
      "parliamentary_group" : "SPD",  
      "affiliation" : "https://d-nb.info/gnd/2022139-3",  
      "GND_Kennung" : "NA",  
      "Wikidata_ID" : "NA"  
    },  
    {  
      "label" : "Ernst Schwanhold",  
      "occupation" : "Politik*",  
      "parliamentary_group" : "SPD",  
      "affiliation" : "https://d-nb.info/gnd/2022139-3",  
      "GND_Kennung" : "NA",  
      "Wikidata_ID" : "NA"  
    },  
    {  
      "label" : "Kurt Rossmann",  
      "occupation" : "Politik*",  
      "parliamentary_group" : "CDU",  
      "affiliation" : "https://d-nb.info/gnd/7230-8",  
      "GND_Kennung" : "NA",  
      "Wikidata_ID" : "NA"  
    }  
  ]  
}
```

Export

Cancel

☒ Krüger, Paul (35)☒ Create new item

Search for match

Datenquellen und Oberfläche von lobid-gnd

Datenabgleich und -anreicherung in OpenRefine

Möglichkeiten zur Verbesserung des Matching

Standardisierung & W3C Community Group

Eigene Daten in OpenRefine verfügbar machen

Standardisierung & W3C Community Group

Standardisierung & W3C Community Group

W3C Entity Reconciliation Community Group

Standardisierung & W3C Community Group

W3C Entity Reconciliation Community Group

<https://www.w3.org/community/reconciliation/>

Standardisierung & W3C Community Group

W3C Entity Reconciliation Community Group

<https://www.w3.org/community/reconciliation/>

<https://github.com/reconciliation-api>

Ziele

Ziele

Dokumentation und Verbesserung des Protokolls der
OpenRefine Reconciliation API

Ziele

Dokumentation und Verbesserung des Protokolls der
OpenRefine Reconciliation API

Protokoll als Standard für den Datenabgleich im Web
vorantreiben

Ziele

Dokumentation und Verbesserung des Protokolls der
OpenRefine Reconciliation API

Protokoll als Standard für den Datenabgleich im Web
vorantreiben

Werkzeuge rund um das Protokoll bereitstellen

Testbench

Testbench

Ein Werkzeug zum Testen von Reconciliation-Services

Testbench

Ein Werkzeug zum Testen von Reconciliation-Services

Ein zentrales Verzeichnis verfügbarer Reconciliation-Services (aktuell 38)

Testbench

Ein Werkzeug zum Testen von Reconciliation-Services

Ein zentrales Verzeichnis verfügbarer Reconciliation-Services (aktuell 38)

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



Reconciliation service test bench

[Services](#)[Test bench](#)

This table lists reconciliation services known to [Wikidata](#) and indicates which features of the [reconciliation API](#) they support.

If you are viewing this app over HTTPS it is likely that reconciliation services using HTTP are blocked as "mixed content". You can disable this protection in your browser to load HTTP reconciliation services.

Due to a [security risk inherent to JSONP](#), only endpoints supporting [CORS](#) are loaded by default. You can click the  button in each row to attempt to load the service via [JSONP](#). Note that a malicious endpoint could use JSONP to execute arbitrary JavaScript code in this page. If you trust all the reconciliation services listed here, you can also .

Name	Endpoint	CORS	JSONP	View entities	Suggest entities	Suggest types	Suggest properties	Preview entities	Extend data
Wikidata (docs , source) 	 https://wikidata.reconci.link/en/api	✓		✓	✓	✓	✓	✓	✓
Wikibase Registry (docs , source) 	 https://wikibase-registry.reconci.link/en/api	✓		✓	✓	✓	✓	✓	✓
Getty Vocabulary Program (docs) 	 https://services.getty.edu/vocab/reconcile/	✓		✓	✗	✗	✓	✓	✗
Thesaurus of Geographic Names LOD Dataset (docs) 	 https://services.getty.edu/vocab/reconcile/	✓		✓	✗	✗	✓	✓	✗
Scholars@BGU (docs) 	 https://scholars.bgu.ac.il/reconcile	✗							
ROR ID (source) 	 https://reconcile.ror.org/reconcile	✗							
OpenCorporates (docs) 	 https://opencorporates.com/reconcile	✓		✓	✓	✗	✓	✓	✗

Census

Census

Überblick über das Reconciliation-Ökosystem

Census

Überblick über das Reconciliation-Ökosystem

Verfügbare Services, Clients, Software-Bibliotheken

Census

Überblick über das Reconciliation-Ökosystem

Verfügbare Services, Clients, Software-Bibliotheken

<https://reconciliation-api.github.io/census/>

Andere Clients als OpenRefine

Alma Refine

< Physical Titles (1 - 1 of 1)

The myths of security what the computer security industry doesn't want you to know

Save Query

Institution

Community

Sort by : Rank Secondary Sort by : Rank Expand

1



The myths of security : what the computer security industry doesn't want you to know / John Viega with introduction by Elia Zafrani

Book By Viega, John. (Beijing ; Cambridge ; Sebastopol, CA : O'Reilly, ©2009.)

Subject: Internet industry--Security measures. Computer security. Data protection. and others

Modification Date: 18/08/2020 13:13

Creation Date: 03/02/2020 09:30

Edition: 1st ed.

Language: English

ISBN: 9780596523022

Record number: (OCoLC)263978539

MMS ID: 99229211100121

Edit Record

Holdings

Physical (2)

Other details

< Back to App List



Alma Refine

To begin refining records, navigate to a list of Bibliographic records or select a set.

☐ Set

☒ Displayed Records

☐ Select Records

Refine the 1 records on this page

Select a refine service

GND

Begin Refining

Cocoda

BK Basisklassifikation

Wohnungsbau

5 Ingenieurwissenschaften
56.80 Gebäudelehre: Allgemeines
56.81 Wohnungsbau

Info Labels Editorial Search Links

<http://uri.gbv.de/terminology/bk/56.81>
Modified: 29. Nov. 2018

Top Concepts Favorite Concepts

- 56.80 Gebäudelehre: Allgemeines
- 56.81 Wohnungsbau**
- 56.82 Verwaltungsbau. Geschäftshäuser
- 56.83 Industriebau
- 56.84 Schulbau. Bibliotheksbau. Museumsbau. Theaterbau
- 56.85 Sportbauten. Freizeitanlagen
- 56.86 Sozialbauten. Krankenhausbau
- 56.87 Hotelbau. Messearchitektur
- 56.88 Sakralbau
- 56.89 Gebäudelehre: Sonstiges
- 56.90 Architektur: Sonstiges

Mapping Editor

not saved

BK RVK

56.81 Wohnungsbau = ZH 5000 - ZH 5960 Wohnungsbau

Concordances Mapping Search Mapping Navigator

Mappings: C L W Recommendations: CR CC CO

From	To	Creator	Date
C Concordance Registry 4 Mappings			
DDC 690 Hausbau und Bauhandwerk	RVK ZH 5000 - ZH 5960 Wohnungsbau	VZG	2013
DDC 728 Wohnbauten und verwandte Gebäude	RVK ZH 5000 - ZH 5960 Wohnungsbau	VZG	2013
DDC 690 Hausbau und Bauhandwerk	BK 56.81 Wohnungsbau	VZG	
DDC 728 Wohnbauten und verwandte Gebäude	BK 56.81 Wohnungsbau	VZG	
L Local No Mappings			
W Wikidata-Mappings No Mappings			
CR coli-conc Recommendations 5 of 12 Mappings			
BK 56.81 Wohnungsbau	RVK ZH 5000 - ZH 5960 Wohnungsbau		
BK 56.81 Wohnungsbau	RVK ZH 5000 - ZH 5960 Wohnungsbau		
BK 56.81 Wohnungsbau	RVK PN 615 Wohnungsbaurecht		
BK 56.81 Wohnungsbau	RVK PN 981.615		

RVK Regensburg

Classification [0] PUBLICDOMAIN

Wohnungsbau

ZG - ZS Technik
ZH Architektur
ZH 5000 - ZH 5960 Wohnungsbau

Info Labels Search Links

<http://rvk.uni-regensburg.de/nt/ZH%205000%20-%20ZH%205960>
156168:
Created: 5. Juli 2012
Modified: 19. März 2019

ZH 5000 Allgemeines
ZH 5050 Wohngebäude, Geschichte
ZH 5100 Ein- und Zweifamilienhäuser, Bungalows

Top Concepts Favorite Concepts

- Baukunstarten
- ZH 4700 - ZH 4999 Baugeschichte
- ZH 5000 - ZH 5960 Wohnungsbau**
- ZH 6000 - ZH 6410 Öffentliche Gebäude
- ZH 6500 - ZH 6850 Wirtschaftsgebäude; gewerbliche Bauten allgemein
- ZH 7000 - ZH 7985 Gebäude
- ZH 8000 - ZH 8985 Innenarchitektur
- ZH 9000 - ZH 9450 Raumordnung und Städtebau
- ZH 9605 - ZH 9800 Landschaftsarchitektur/Landschaftsplanung
- ZI Baulingenieurwesen
- ZK Bergbau und Hüttenwesen
- ZL Maschinenbau
- ZM Werkstoffwissenschaften und

Specs

Specs

Spezifikation des Protokolls für Datenabgleich und -
anreicherung

Specs

Spezifikation des Protokolls für Datenabgleich und -
anreicherung

Ein Protokoll für den Datenabgleich im Web am Beispiel von
OpenRefine und der Gemeinsamen Normdatei (GND)

Specs

Spezifikation des Protokolls für Datenabgleich und -
anreicherung

Ein Protokoll für den Datenabgleich im Web am Beispiel von
OpenRefine und der Gemeinsamen Normdatei (GND)

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

Datenquellen und Oberfläche von lobid-gnd

Datenabgleich und -anreicherung in OpenRefine

Möglichkeiten zur Verbesserung des Matching

Standardisierung & W3C Community Group

Eigene Daten in OpenRefine verfügbar machen

Eigene Daten in OpenRefine verfügbar machen

Eigene Daten in OpenRefine verfügbar machen

1. Einzeldatensätze als JSON anbieten

Eigene Daten in OpenRefine verfügbar machen

1. Einzeldatensätze als JSON anbieten
2. Suche nach Datensätzen implementieren

Eigene Daten in OpenRefine verfügbar machen

1. Einzeldatensätze als JSON anbieten
2. Suche nach Datensätzen implementieren
3. Reconciliation-Protokoll implementieren

Eigene Daten in OpenRefine verfügbar machen

1. Einzeldatensätze als JSON anbieten
2. Suche nach Datensätzen implementieren
3. Reconciliation-Protokoll implementieren

Für Details zu 1. und 2. siehe [Workshop-Material From LOD to LOUD: building & using JSON-LD APIs](#)

1. Einzeldatensätze als JSON

1. Einzeldatensätze als JSON

z.B. http://localhost:9200/loc/work/c0000000011/_source

1. Einzeldatensätze als JSON

z.B. `http://localhost:9200/loc/work/c0000000011/_source`

Dann schon nutzbar in OpenRefine: Daten per URL holen

Exkurs: z.B. mit diesen Daten

```
http://id.loc.gov/resources/works/c000000011  
http://id.loc.gov/resources/works/c000000020  
http://id.loc.gov/resources/works/c000000026  
http://id.loc.gov/resources/works/c000000028  
http://id.loc.gov/resources/works/c000000100
```



Facet / Filter

Undo / Redo 0 / 0

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](#)

5 rows

Show as: **rows** records

Show: 5 10 25 50 rows

▼ All

▼ Column 1



1.

Facet



/works/c000000011



2.

Text filter



/works/c000000020



3.

Edit cells



/works/c000000026



4.

Edit column



/works/c000000028



5.

Transpose



Split into several columns...

Sort...

View

Reconcile

Add column based on this column...

Add column by fetching URLs...

Add columns from reconciled values...

Rename this column

Remove this column

Move column to beginning

Move column to end

Move column left

Move column right

ID rauspicken & URL erzeugen, unter der wir JSON
finden/selbst anbieten

ID rauspicken & URL erzeugen, unter der wir JSON
finden/selbst anbieten

<http://id.loc.gov/resources/works/c0000000011>

ID rauspicken & URL erzeugen, unter der wir JSON
finden/selbst anbieten

<http://id.loc.gov/resources/works/c0000000011> →
http://localhost:9200/loc/work/c0000000011/_source

ID rauspicken & URL erzeugen, unter der wir JSON
finden/selbst anbieten

`http://id.loc.gov/resources/works/c000000011` →
`http://localhost:9200/loc/work/c000000011/_source`

```
"http://localhost:9200/loc/work/" +  
split(value, '/')[-1] + "_source"
```

Facet / Filter

[Undo / Redo](#) 0 / 0

5 rows

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](#)

Add column by fetching URLs based on column Column 1

New column name Throttle delay
milliseconds

On error



set to blank



store error



Cache responses

HTTP headers to be used when fetching URLs: [Show](#)

Formulate the URLs to fetch:

Expression

Language

General Refine Expression Language (GREL) ▼

No syntax error.

Preview

[History](#)[Starred](#)[Help](#)

row	value	"http://localhost:9200/loc/wor ...
1.	http://id.loc.gov/resources/works/c000000011	http://localhost:9200/loc/work/c000000011/_source
2.	http://id.loc.gov/resources/works/c000000020	http://localhost:9200/loc/work/c000000020/_source
3.	http://id.loc.gov/resources/works/c000000026	http://localhost:9200/loc/work/c000000026/_source
4.	http://id.loc.gov/resources/works/c000000028	http://localhost:9200/loc/work/c000000028/_source

OK

Cancel

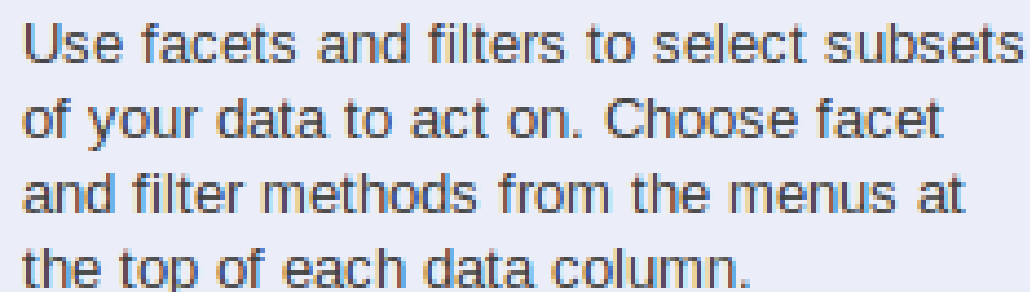
Das JSON können wir dann zum Anreichern nutzen



Facet / Filter

Undo / Redo 1 / 1

Using facets and filters



Not sure how to get started?
Watch these screencasts

5 rows

Show as: **rows** records

Show: 5 10 25 50 rows



▼ Column 1



[http://id.loc.gov
/resources
/works
/c000000011](http://id.loc.gov/resources/works/c000000011)

Facet

Text filter

Edit cells

Edit column

transpose

Sort...

View

Reconcile

```
1800.", "componentList": [{"and travel"}, {"type": "http://1800."}, {"id": "http://id.loc.(Mass.)"}], "title": {"type": "French. 1784", "titleSortKey": "farmer. French. 1784"}, {"id": "_:b592", "type": "Title 1784", "titleSortKey": "Lette 1784"}, {"label": "Letters from
```

- Split into several columns...
- Add column based on this column...
- Add column by fetching URLs...
- Add columns from reconciled values...

- Rename this column
- Remove this column

- Move column to beginning
- Move column to end
- Move column left
- Move column right

Facet / Filter

Undo / Redo 1 / 1

5 rows

Show as: rows records

Show: 5 10 25 50 rows

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](#)

Add column based on column json

New column name

label



set to blank



store error



copy value from original column

Expression

Language

General Refine Expression Language (GREL) ▼

value.parseJson().label

No syntax error.

Preview

History

Starred

Help

row value

value.parseJson().label

1.	{ "@context": "http://localhost:3000/context.json", "id": "http://id.loc.gov/resources/works/c000000011", "type": ["Work", "Text"], "adminMetadata": [{"type": "AdminMetadata", "encodingLevel": [{"type": "EncodingLevel", "code": "5"}], "changeDate": {"type": "http://www.w3.org/2001/XMLSchema#dateTime", "@value": "2004-11-19T10:19:25"}, "creationDate": {"type": "http://www.w3.org/2001/XMLSchema#date", "@value": "1993-10-26"}, "descriptionConventions": [{"type": "DescriptionConventions", "code": "aacr"}].	Letters from an American farmer. French. 1784
----	---	---

OK

Cancel

clipboard - OpenRefine

127.0.0.1:3333/project?project=1657999837942

Search

Open...ExportHelp

OpenRefine clipboardPermalink

Facet / FilterUndo / Redo 5 / 5

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

5 rows

Show as: rows recordsShow: 5 10 25 50 rows

		uri	author	label
☆		1. http://id.loc.gov/resources/works/c000000011	St. John de Crèvecoeur, J. Hector, 1735-1813.	Letters from an American farmer. French. 1784
☆		2. http://id.loc.gov/resources/works/c000000020	Howells, William Dean, 1837-1920.	Italian journeys /
☆		3. http://id.loc.gov/resources/works/c000000026	Ryan, George, active 1811.	Report of the trial of George Ryan : before the S May term 1811, for highway robbery.
☆		4. http://id.loc.gov/resources/works/c000000028	Miscellaneous Pamphlet Collection (Library of Congress)	The devil upon crutches in England; or, Night sc
☆		5. http://id.loc.gov/resources/works/c000000100	Colorado State Teachers' Association.	Education in Colorado : 1861-1885 : a brief histo the history of the State Teachers' Association, a

Export project

Project data package

Tab-separated value

Comma-separated value

HTML table

Excel (.xls)

Excel 2007+ (.xlsx)

ODF spreadsheet

Custom tabular exporter...

SQL Exporter...

Templating...

QuickStatements

2. Suche nach Datensätzen implementieren

2. Suche nach Datensätzen implementieren

z.B. `http://localhost:9200/loc/work/_search?q=*`

2. Suche nach Datensätzen implementieren

z.B. `http://localhost:9200/loc/work/_search?q=*`

d.h. Indexierung der Daten in Suchmaschine, z.B.
Elasticsearch, Solr, Lucene

z.B. mit diesen Daten

Letters from an American farmer
Italian journeys
Report of the trial of George Ryan
The devil upon crutches in England
Educational interests of Colorado

z.B. mit diesen Daten

```
Letters from an American farmer  
Italian journeys  
Report of the trial of George Ryan  
The devil upon crutches in England  
Educational interests of Colorado
```

```
"http://localhost:9200/loc/work/_search?q=\" +  
  replace(value, " ", "+") + "\""
```

z.B. mit diesen Daten

```
Letters from an American farmer  
Italian journeys  
Report of the trial of George Ryan  
The devil upon crutches in England  
Educational interests of Colorado
```

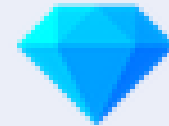
```
"http://localhost:9200/loc/work/_search?q=\" +  
    replace(value, " ", "+") + "\""
```

```
value.parseJson().hits.hits[0]._source.id
```

**Facet / Filter**

Undo / Redo 16 / 16

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](#)

5 rowsShow as: **rows** records

Show: 5 10 25 50 rows

▼ All		▼ title		▼ uri
☆	🗨	1.	Letters from an American farmer	http://id.loc.gov/resources/works/11#Work240-15
☆	🗨	2.	Italian journeys	http://id.loc.gov/resources/works/c000000020
☆	🗨	3.	Report of the trial of George Ryan	http://id.loc.gov/resources/works/c000000026
☆	🗨	4.	The devil upon crutches in England	http://id.loc.gov/resources/works/c000000028
☆	🗨	5.	Educational interests of Colorado	http://id.loc.gov/resources/works/c000000100

3. Reconciliation-Protokoll implementieren

3. Reconciliation-Protokoll implementieren

Mit speziellen Reconciliation-Tools/Frameworks

3. Reconciliation-Protokoll implementieren

Mit speziellen Reconciliation-Tools/Frameworks

Mit generischen Web-Entwicklungs-Tools/Frameworks
(JSON API)

3. Reconciliation-Protokoll implementieren

Mit speziellen Reconciliation-Tools/Frameworks

Mit generischen Web-Entwicklungs-Tools/Frameworks
(JSON API)

Testen mit Testbench



Reconciliation Service API

A protocol for data matching on the Web



ReSpec

Draft Community Group Report 16 September 2021

Latest editor's draft:

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

Editors:

[Antonin Delpeuch](#) (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 © 2021 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.





§ 2. Core Concepts

ReSpec

This section documents the data model behind the reconciliation API. A reconciliation service lets users match their data against [entities](#) exposed by the service. Matching can be refined by filtering by [type](#) or [properties](#) with [property values](#). The purpose of this section is to define these notions.

§ 2.1 Entities

An **entity** is a record in the data source exposed by the service. It comprises the following fields:

id

an identifier, which is a non-empty string. This identifier must be unique among all entities;

name

a name, which is also a non-empty string;

description

an optional description as a human-readable string;

type

a list of [types](#), possibly empty;

Moreover, for each [property](#) it contains a set of associated [property values](#), possibly empty.

Reconciliation services can define in their [service manifest](#) a **view template** for entities, which associates to each entity a corresponding [URI](#), by inserting its identifier in the template. A view template is a string which contains the `{{id}}` substring. For each entity, replacing `{{id}}` in the template by the entity's identifier *MUST* result in a valid [URI](#). To guarantee the correctness of the formed [URI](#), clients *MUST* percent-encode component-delimiting reserved characters in the identifier, i.e. encode it as a [URI](#) component [RFC3986].



A **reconciliation query batch** is a set of [reconciliation queries](#) indexed by string identifiers.

ReSpec

Minimal example of a [reconciliation query batch](#) with mandatory fields only:

```
{
  "q1": {
    "query": "Hans-Eberhard Urbaniak"
  },
  "q2": {
    "query": "Ernst Schwanhold"
  }
}
```

Full example of a [reconciliation query batch](#) with all optional fields:

```
{
  "q0": {
    "query": "Christel Hanewinckel",
    "type": "DifferentiatedPerson",
    "limit": 5,
    "properties": [
      {
        "pid": "professionOrOccupation",
        "v": "Politik*"
      },
      {
        "pid": "affiliation",
        "v": "http://d-nb.info/gnd/2022139-3"
      }
    ]
  }
}
```


In Testbench
abgleichen

In Testbench abgleichen

Einzel-Abgleich simulieren für "Brandt, Willy" mit Beruf und
Affiliation (wie im 2. Beispiel vorhin)

In Testbench abgleichen

Einzel-Abgleich simulieren für "Brandt, Willy" mit Beruf und
Affiliation (wie im 2. Beispiel vorhin)

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

This form lets you test a reconciliation endpoint interactively.

Endpoint:

https://test.lobid.org/gnd/reconcile



Reconcile

Suggest

Preview

Extend

Name:

Brandt, Willy

Reconcile

Type:

☐ Normdatenressource

AuthorityResource

☐ Körperschaft

CorporateBody

☐ Konferenz oder Veranstaltung

ConferenceOrEvent

☐ Schlagwort

SubjectHeading

☐ Werk

Work

☐ Geografikum

PlaceOrGeographicName

☒ Individualisierte Person

DifferentiatedPerson

☐ Familie

Family

☐ Custom:

☐ Reconcile against no particular type

Properties:

Beruf oder Beschäftigung

Politik*



Affiliation

tps://d-nb.info/gnd/2022



query: "Brandt, Willy"

type: "DifferentiatedPerson"

▼ properties:

▼ 0:

pid: "professionOrOccupation"

v: "Politik*"

▼ 1:

pid: "affiliation"

v: "https://d-nb.info/gnd/2022139-3"

[View query results on the service](#)

Brandt, Willy

11851444X

107.3416

Normdatenressource (AuthorityResource),
Individualisierte Person (DifferentiatedPerson)

Brandt, Hugo

119034298

73.07417

Normdatenressource (AuthorityResource),
Individualisierte Person (DifferentiatedPerson)

Brandt, Willy

1095908332

71.96527

Normdatenressource (AuthorityResource),
Individualisierte Person (DifferentiatedPerson)

Brandt, Willy

1146796943

71.96527

Normdatenressource (AuthorityResource),
Individualisierte Person (DifferentiatedPerson)

Brandt, Willy

135738563

71.90846

Normdatenressource (AuthorityResource),
Individualisierte Person (DifferentiatedPerson)

Prölß, Willy

1051304008

69.98882

Normdatenressource (AuthorityResource),
Individualisierte Person (DifferentiatedPerson)

Zusammenarbeit von Fachebene & Entwicklung

Zusammenarbeit von Fachebene & Entwicklung

Schon erster Schritt, Bereitstellung von Einzeldaten als JSON
kann fachlich in OpenRefine getestet werden

Zusammenarbeit von Fachebene & Entwicklung

Schon erster Schritt, Bereitstellung von Einzeldaten als JSON
kann fachlich in OpenRefine getestet werden

Bis hin zur Evaluation der Scores beim Abgleich mit
disambiguierenden lokalen Daten in der Testbench

Zusammenarbeit von Fachebene & Entwicklung

Schon erster Schritt, Bereitstellung von Einzeldaten als JSON
kann fachlich in OpenRefine getestet werden

Bis hin zur Evaluation der Scores beim Abgleich mit
disambiguierenden lokalen Daten in der Testbench

Ermöglicht auf vielen Ebenen die unmittelbare
Zusammenarbeit von Fachebene & Entwicklung

Weitere Informationen

OpenRefine user manual

GND reconciliation for OpenRefine

Implementierung & Details in lobid-gnd

Danke an Felix Ostrowski und Florian Gilberg für die Beispiel-Datensets/-Anwendungsfälle