

# Abgleich & Anreicherung eigener Daten mit der GND

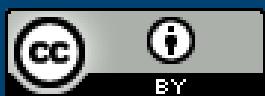
OpenRefine Reconciliation mit lobid-gnd

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](http://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](http://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 a Mozilla Firefox browser window with the title bar "lobid-gnd — Mozilla Firefox". The address bar displays the URL "https://lobid.org/gnd". The main content area shows the lobid-gnd homepage. At the top, there is a navigation bar with the "lobid" logo, "gnd", "Erkunden", and "API" buttons. Below the navigation bar is a search bar containing the placeholder text "Suchoptionen: AND, OR, AND NOT, ""-Phrasensuche, \*-Trunkierung" and a magnifying glass icon. The main heading "lobid-gnd — Rechercheoberfläche & LOD-API für die GND" is followed by a brief description of the GND (Gemeinsame Normdatei) and its usage. A paragraph explains the data source and licensing. Another section describes the services offered by lobid-gnd. On the right side, there is a small image of a lake surrounded by trees with the caption "Florida" and credits to Ebyabe and Wikimedia Commons under CC BY 2.5. The footer contains the lobid-gnd logo, a link to Impressum/Datenschutz/Twitter/GitHub/Blog, and the hbz logo.

# /reconcile

The screenshot shows a Mozilla Firefox browser window with the title bar "lobid-gnd - Reconcile — Mozilla Firefox". The address bar displays the URL "https://lobid.org/gnd/reconcile/". The main content area features a blue header bar with the "lobid" logo, the word "gnd", and menu items "Erkunden" and "API". Below the header, the text "GND reconciliation for OpenRefine" is prominently displayed, followed by "Align your data with the Integrated Authority File GND". A red link "Use this URL in OpenRefine: <https://lobid.org/gnd/reconcile/>" is shown. To the right of the text are two icons: a blue hexagonal mesh icon and a blue diamond icon. The main body of the page is divided into sections: "Data and Tools" (links to "Integrated Authority File GND" and "OpenRefine, Reconciliation"), "Tutorials" (links to "GND reconciliation for OpenRefine", "Using the suggest APIs for searching types, properties, and entities", and "Using additional properties to improve matches"), and "API and Specs" (links to "GND reconciliation for OpenRefine: API documentation", "W3C entity reconciliation community group", and "Reconciliation service test bench"). At the bottom, there is a footer bar with the "lobid-gnd" logo, the text "ein LOD-Dienst des hbz — Hochschulbibliothekszentrum des Landes NRW", and links to "Impressum", "Datenschutz", "Twitter", "GitHub", and "Blog".

# Abgleich und Anreicherung lokaler Daten mit der GND

The screenshot shows the OpenRefine interface with a project titled "kpe csv". The main view displays 475 rows of data. The columns include "Name", "DateOfBirth", "DateOfDeath", and "Type". A context menu is open over the first row, which contains the name "Neumann, Israel Ben". The menu options include "Match this Cell", "Match All Identical Cells", and "Cancel". Below the menu, there is a list of names from the GND matching the local data, such as "Neumann, Paul (65)", "Neumann, Erich (65)", and "Ben-Elī'ezer (61)". The interface also features facets on the left for filtering by "Name: judgment" and "Name: best candidate's score".

		Name	DateOfBirth	DateOfDeath	Type
1.	KPE 186170	Neumann, Israel Ben Neumann, Paul (65) Neumann, Erich (65) Ben-Elī'ezer (61) Neumann, Hans Paul Hartmann, Israel (58) Israeli, Ben-Zion (56) Pinħas ben Y. (56) Ben-Aroya, Avraham Rōm, Mōše Ben Yōḥā Ben-Yiśrā'el, Asher (50) Create new item	1887	18.04.1961	personal
2.	KPE 243353	Yamawaki, Michiko	13.07.1910	2000	personal
3.	KPE 255684	Ackermann, Richard	1905	1977	personal
4.	KPE 255687	Agatz, August Agatz, Wilhelm (85) Agatz, August (82) Agatz, Alfred (60)	1904	1945	personal

# 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



The screenshot shows the lobid-gnd search interface. At the top, there is a navigation bar with the lobid logo, a 'gnd' button with a hand cursor icon, and a 'Dokumentation' dropdown. Below the navigation bar is a search bar containing the placeholder 'Einfache Suchoberfläche' and the text 'Suchoptionen: AND, OR, AND NOT, ""-Phrasensuche, \*-Trunkierung'. A red box highlights this search area. The main title 'lobid-gnd — Rechercheoberfläche & LOD-API für die GND' is displayed below the search bar. To the right of the title is the hbz logo. Below the title, a text block describes the GND as a common normative database containing over 14 million entries used for cataloging literature and other projects. It also mentions the data's availability through various institutions and its CC0 license. A red arrow points from the text '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.' down towards the footer. On the right side of the page, there is a photograph of the Berlin skyline with the TV tower, labeled 'Berlin'. Below the photo, it says '(Bildquelle: Thomas Wolf, www.foto-tw.de | Wikimedia Commons | CC BY-SA 3.0)'. At the bottom of the page, there is a footer with the lobid-gnd logo, the text 'ein LOD-Dienst des hbz — Hochschulbibliothekszentrum des Landes NRW', and links for Impressum, Datenschutz, Twitter, GitHub, and Blog.

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.

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



# 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, Aslı	Individualisierte Person	Schriftstellerin, Journalistin, Moderatorin	130454966 1973 -
 Meray, Sema	Individualisierte Person	Schauspielerin, Dramatikerin	136115519 1968 -
 Özdogan, Selim	Individualisierte Person	Schriftsteller	115376230 1971 -
 Akhanlı, Doğan	Individualisierte Person	Schriftsteller	124063802 1957 -

← 1 →

Ergebnisse eingrenzen:

Entitätstyp

Person  

GND-Sachgruppe 

Personen zu Literaturgeschichte (Schriftsteller)   
Personen zu Betriebswirtschaftslehre (1)  
Personen zu Theater, Tanz (1)  
Personen zu Film (1)  
Personen zu Rundfunk, Neuen Medien (1)

Ländercode

Deutschland  

Türkei  

Beruf oder Beschäftigung 

Schriftsteller (2)  
Schriftstellerin (2)  
Assistentin (1)  
Journalistin (1)  
Dramatikerin (1)

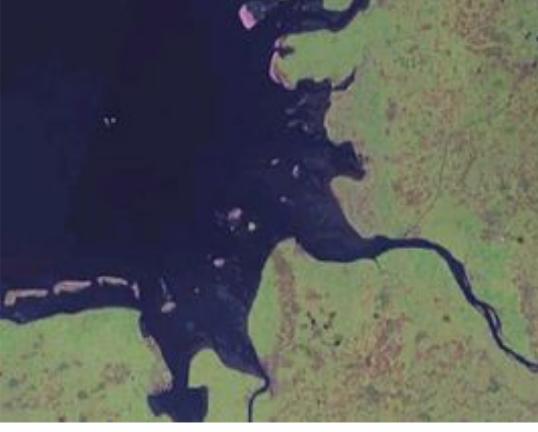
# Einzeltreffer

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  Weitere Einträge mit Oberbegriff partitiv "Nordsee" suchen
Siehe auch	<a href="#">w Wikipedia (English)</a>   <a href="#">w Wikipedia (Deutsch)</a>   <a href="#">Wikidata</a>   <a href="#">Virtual International Authority File (VIAF)</a>   <a href="#">Deutsche Nationalbibliothek (DNB)</a>

Datenquelle: GND und EntityFacts | CC0



Bildquelle: NASA | Wikimedia Commons | Public domain

# Beziehungsgraph

Helgoländer Bucht Geografikum Natürliche geografische Einheit { }

Felder Beziehungen

←

Helgoländer Bucht

Einträge mit Oberbegriff partitiv 'Nordsee' suchen

Oberbegriff partitiv → Oberbegriff instantiell

Nordsee      Bucht

- +

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

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)

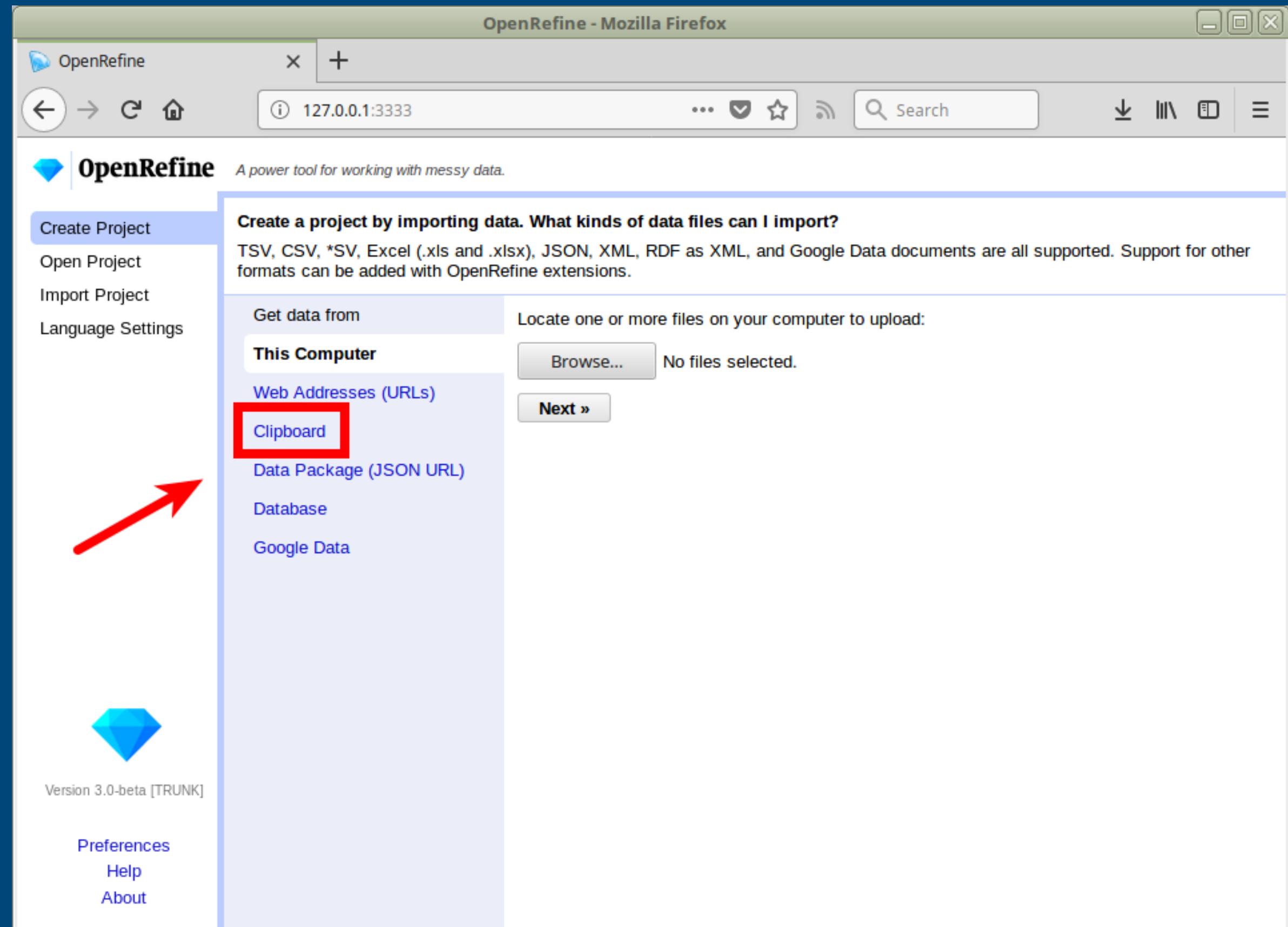
Database

Google Data

Locate one or more files on your computer to upload:

Browse... No files selected.

Next »



OpenRefine - Mozilla Firefox

OpenRefine 127.0.0.1:3333

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 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 »

The screenshot shows the 'Create Project' step of the OpenRefine wizard. A red box highlights the clipboard content: "name;beruf;ort" followed by several names and their roles and locations. Two red arrows point from the bottom right towards the 'Next >' button at the bottom left of the main panel.

OpenRefine - Mozilla Firefox

OpenRefine 127.0.0.1:3333

A power tool for working with messy data.

Create Project   « Start Over   Configure   Project name clipboard   Tags   Create Project »

Parsing Options

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

Parse data as

Character encoding  Update Preview

**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

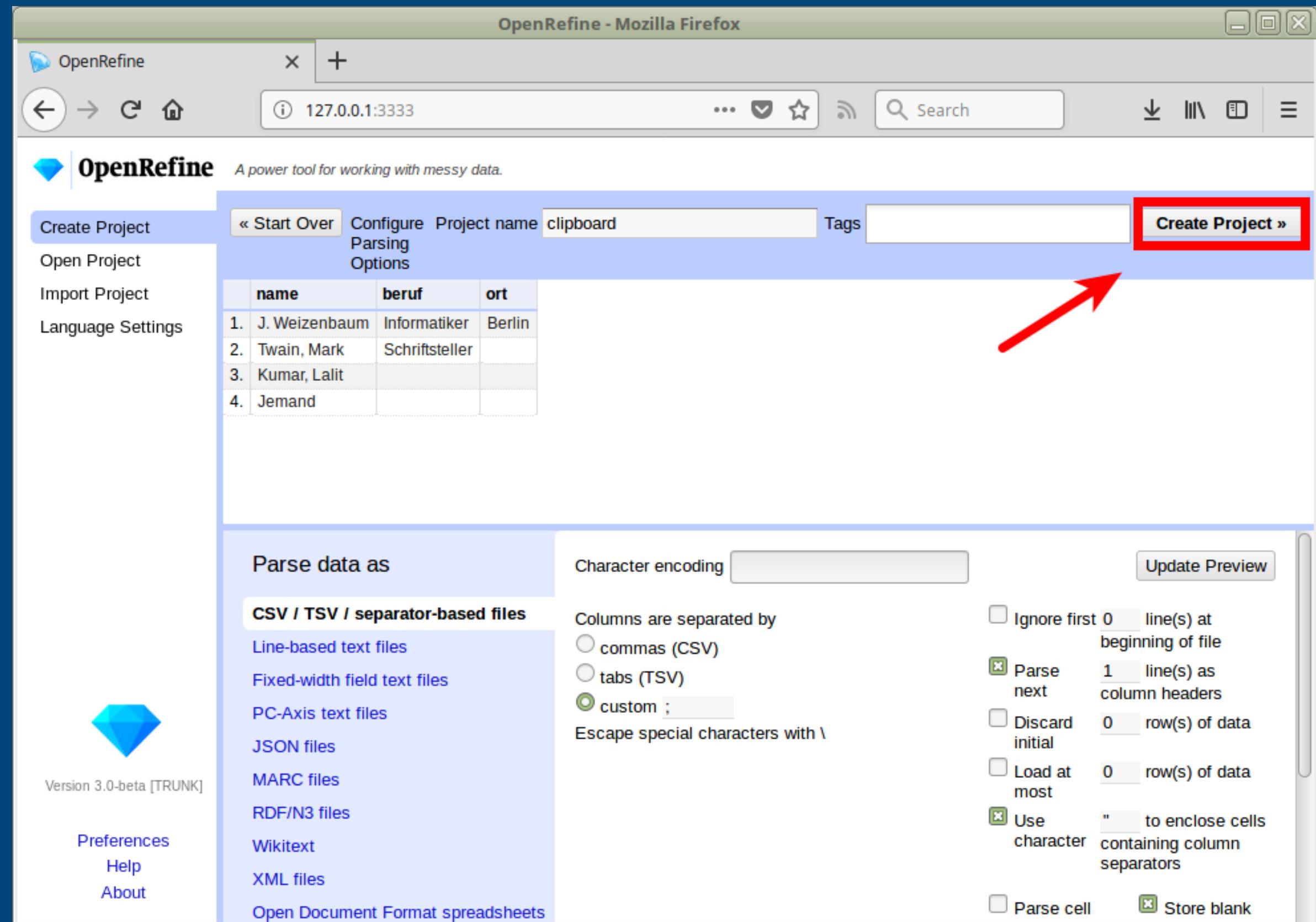
Columns are separated by  
 commas (CSV)  
 tabs (TSV)  
 custom ;

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 — Mozilla Firefox

clipboard - OpenRefine x +

openrefine.labs.lobid.org/proj 120% ... Search

OpenRefine clipboard Permalink

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 »

All name beruf ort

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

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](#)

javascript:{}

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

# Projekt anlegen

[openrefine.labs.lobid.org](http://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 127.0.0.1:3333/project?project=2221362040394 Open... Export Help

OpenRefine clipboard Permalink

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

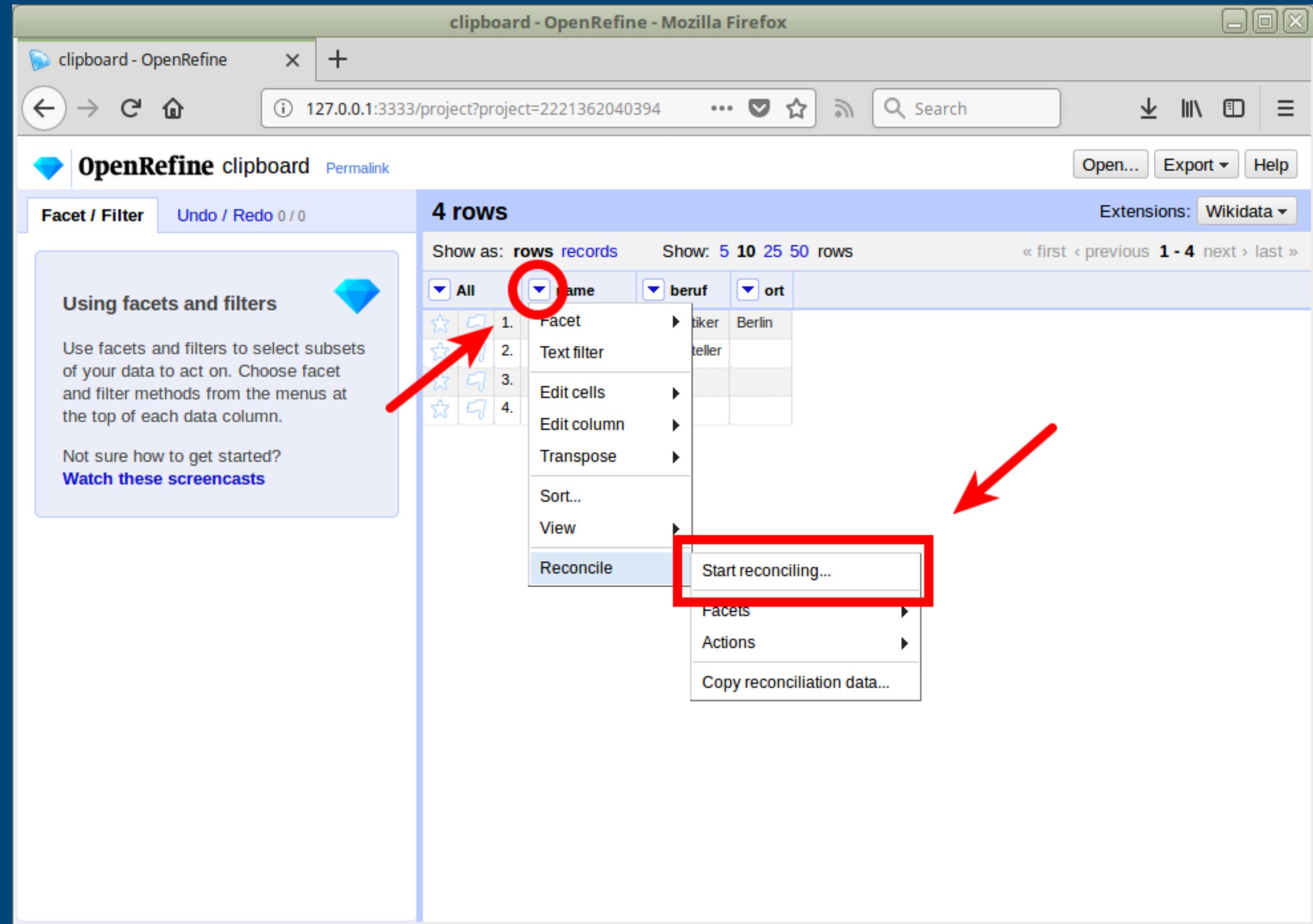
Show as: rows records Show: 5 10 25 50 rows « first < previous 1 - 4 next > last »

All name beruf ort

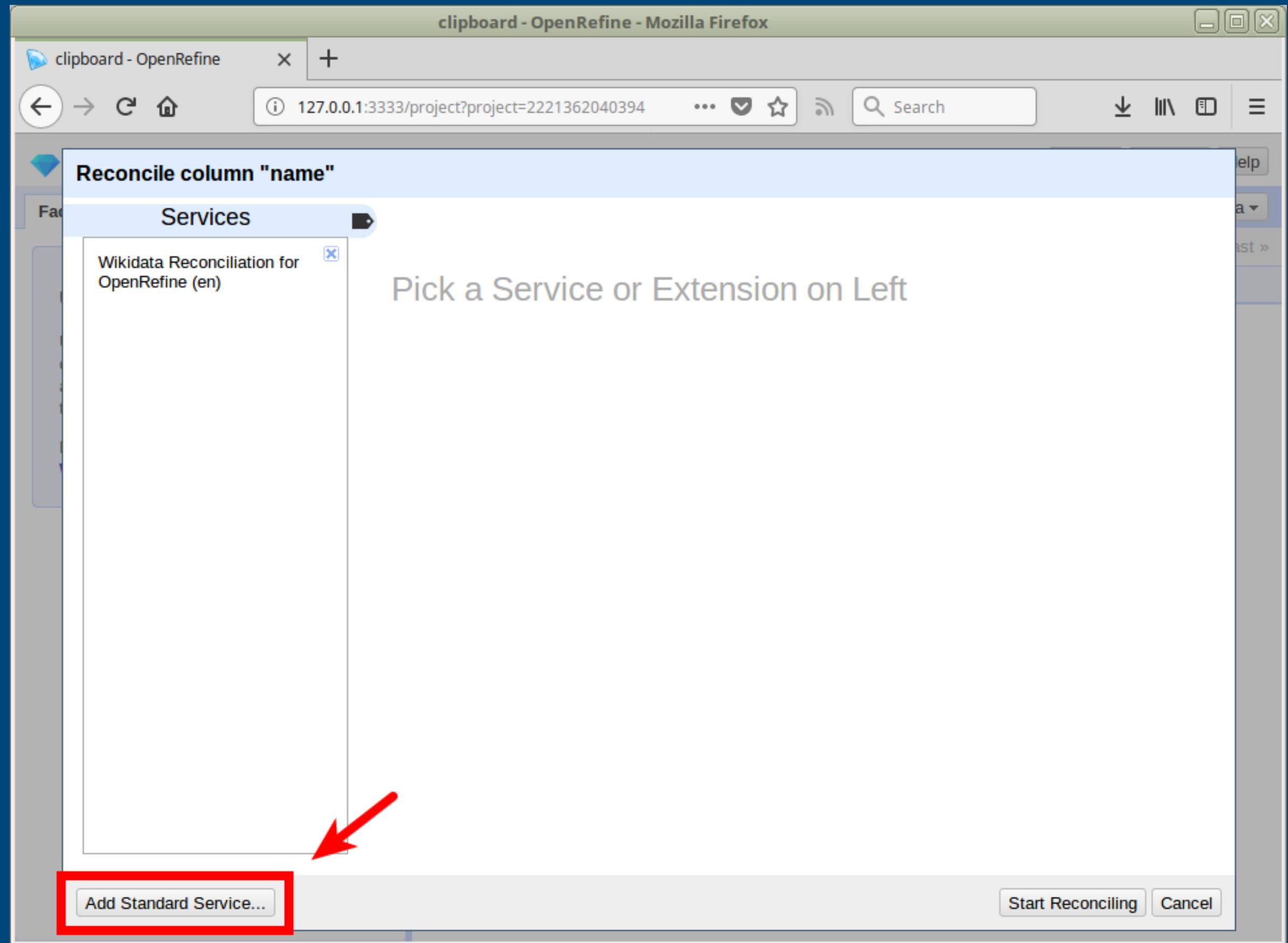
	name	beruf	ort
1.	Facet	tiker	Berlin
2.	Text filter	teller	
3.	Edit cells		
4.	Edit column		
	Transpose		
	Sort...		
	View		
	Reconcile		

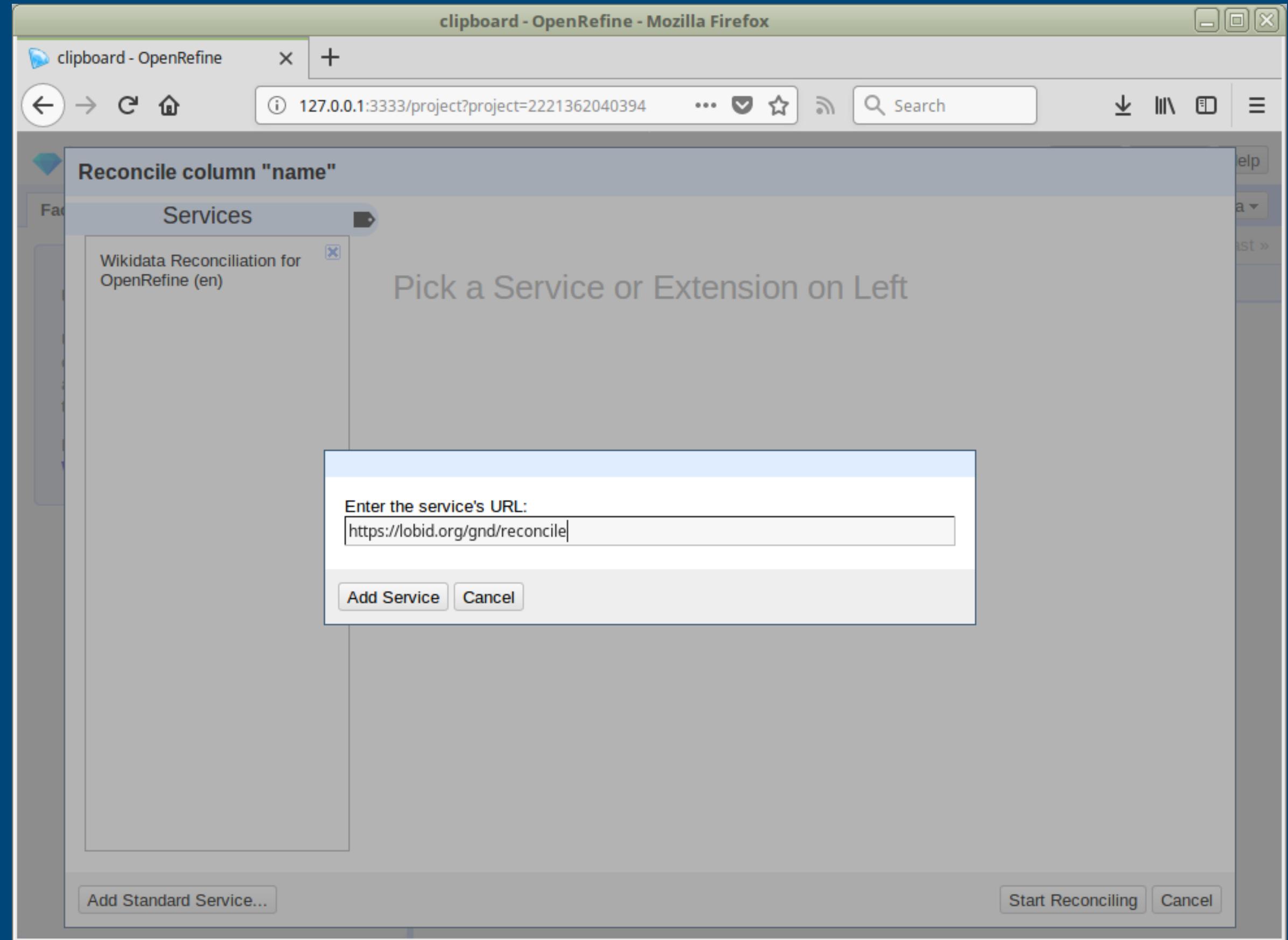
Start reconciling...

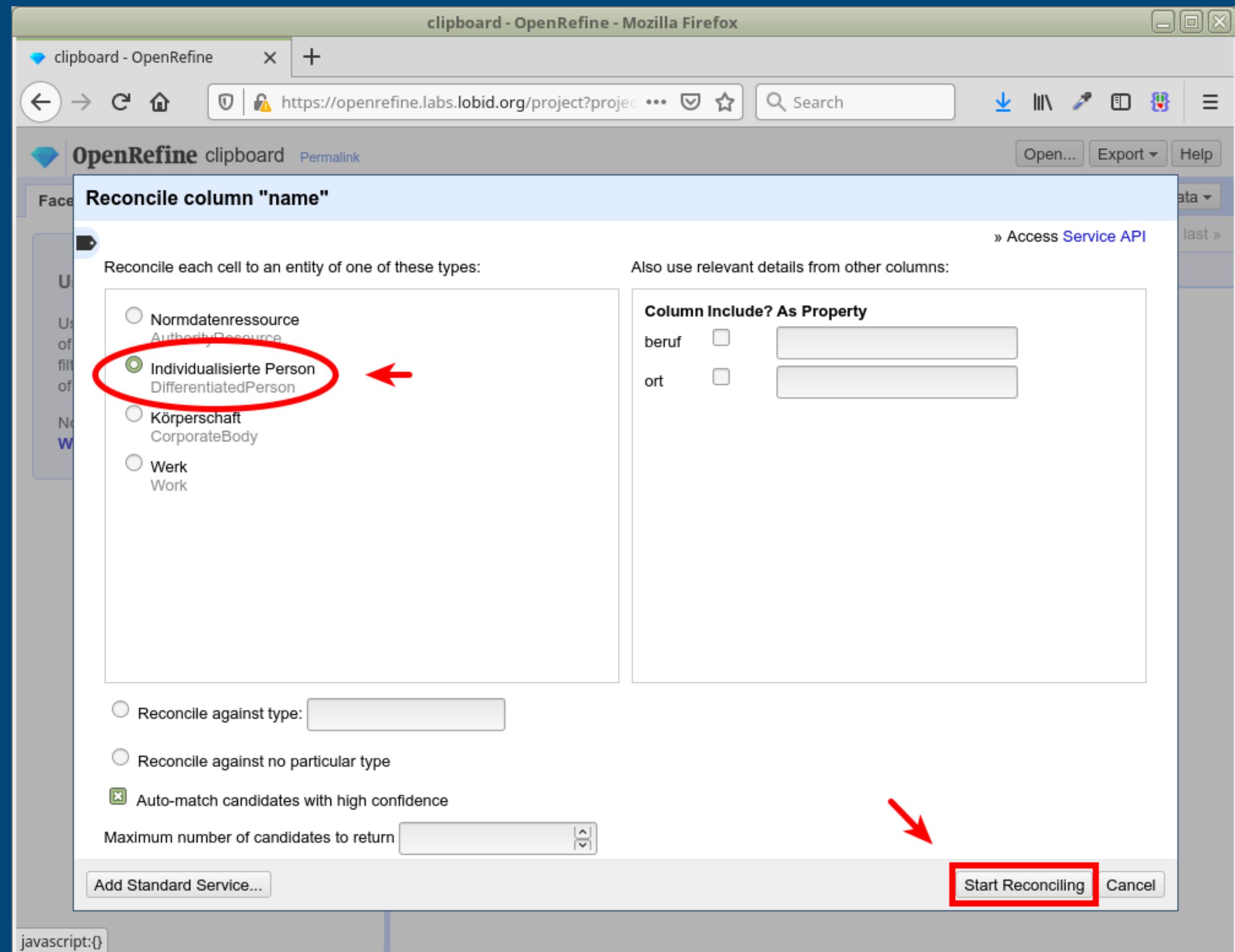
Facets Actions Copy reconciliation data...



The screenshot shows the OpenRefine interface in a browser window. A context menu is open over the 'name' column header. The 'Reconcile' option is circled in red. The 'Start reconciling...' submenu item is also circled in red and has a red arrow pointing to it, indicating the next step in the reconciliation process. The main interface shows 4 rows of data with columns for name, beruf, and ort. The 'name' column has four options: Facet, Text filter, Edit cells, and Edit column. The 'beruf' and 'ort' columns have their own dropdown menus. The 'Reconcile' option in the context menu is part of the Wikidata extension, as indicated by the 'Wikidata' dropdown in the top right.







clipboard - OpenRefine — Mozilla Firefox

clipboard - OpenRefine openrefine.labs.lobid.org/proj 110% ... Search Open... Export Help

OpenRefine clipboard Permalink

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		
2.	Twain, Mark Choose new match		
3.	Kumar, Lalit <input checked="" type="checkbox"/> Kumar, Lalit (80) <input checked="" type="checkbox"/> Kumar, Lalit (80) <input checked="" type="checkbox"/> Lalit Kumar (79) <input checked="" type="checkbox"/> Kumar, Lalit (79) <input checked="" type="checkbox"/> Banerjee, Lalit Kumar (66) <input checked="" type="checkbox"/> Pandit, Lalit Kumar (66) <input checked="" type="checkbox"/> Modi, Lalit Kumar (66) <input checked="" type="checkbox"/> Singh, Lalit Kumar (65) <input checked="" type="checkbox"/> Sen, Lalit Kumar (65) <input checked="" type="checkbox"/> Barua, Lalit Kumar (65) <input checked="" type="checkbox"/> Create new item Search for match		
4.	Jemand <input checked="" type="checkbox"/> Langwiesche, Wilhelm (26) <input checked="" type="checkbox"/> Trömer, Johann Christian (11) <input checked="" type="checkbox"/> Suttner, Bertha von (5) <input checked="" type="checkbox"/> Create new item Search for match		

clipboard - OpenRefine - Mozilla Firefox

clipboard - OpenRefine 127.0.0.1:3333/project?project=2221362040394 Search Extensions: Wikidata

OpenRefine clipboard Permalink

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 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. J. Weizenbaum  
Weizenbaum, Joseph  
WEIZENBAUM, J.  
Create new item

2. Twain, Mark  
Choose new match

3. Kumar, Lalit  
Lalit Kumar (40)  
Banerjee, Lalit Kumar (33)  
Pandit, Lalit Kumar (33)  
Modi, Lalit Kumar (33)  
Bansal, Lalit Kumar (33)  
Sen, Lalit Kumar (33)  
Barua, Lalit Kumar (33)  
Jha, Lalit Kumar (33)  
Shukla, Lalit Kumar (33)  
Ölgä (27)  
Create new item

4. Jemand  
Neigenfind, Rulf (19)  
Langewiesche, Wilhelm (15)  
Rantzau, Adeline Gfin zu (13)  
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

Match this Cell Match All Identical Cells Cancel  
Weizenbaum, Joseph (118820591)  
1923-2008 | Informatiker, Mathematiker  
Individualisierte Person

	All	name	beruf	ort
1.	J. Weizenbaum	Weizenbaum, Joseph WEIZENBAUM, J. Create new item	Informatiker	Berlin
2.	Twain, Mark	Choose new match		
3.	Kumar, Lalit	Lalit Kumar (40) Banerjee, Lalit Kumar (33) Pandit, Lalit Kumar (33) Modi, Lalit Kumar (33) Bansal, Lalit Kumar (33) Sen, Lalit Kumar (33) Barua, Lalit Kumar (33) Jha, Lalit Kumar (33) Shukla, Lalit Kumar (33) Ölgä (27) Create new item		
4.	Jemand	Neigenfind, Rulf (19) Langewiesche, Wilhelm (15) Rantzau, Adeline Gfin zu (13) 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		

# Abgleich / Reconciliation

[openrefine.labs.lobid.org](https://openrefine.labs.lobid.org) | localhost:3333

<https://lobid.org/gnd/reconcile>

(Auf [openrefine.labs.lobid.org](https://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 127.0.0.1:3333/project?project=2221362040394 Search Extensions: Wikidata

OpenRefine clipboard Permalink

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

	name	beruf	ort
1.	Facet	Informatiker	Berlin
2.	Text filter		
3.	Edit cells	Schriftsteller	
	Edit column	Split into several columns...	
	Transpose	Add column based on this column...	
	Sort...	Add column by fetching URLs...	
	View	Add columns from reconciled values...	
	Reconcile	Rename this column	
		Remove this column	
		Move column to beginning	
		Move column to end	
		Move column left	
		Move column right	

Extensions: Wikidata

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

	name	beruf	ort
1.	Facet	Informatiker	Berlin
2.	Text filter		
3.	Edit cells	Schriftsteller	
	Edit column	Split into several columns...	
	Transpose	Add column based on this column...	
	Sort...	Add column by fetching URLs...	
	View	Add columns from reconciled values...	
	Reconcile	Rename this column	
		Remove this column	
		Move column to beginning	
		Move column to end	
		Move column left	
		Move column right	

Extensions: Wikidata

Open... Export Help

clipboard - OpenRefine - Mozilla Firefox

clipboard - OpenRefine +

localhost:3333/project?project=232754877

Search

Open... Export Help

Wikidata

next > last »

### Add columns from reconciled column name

Add Property Preview Reset

Ort

Select an item from the list:

Ort	name
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	relatedSubjectHeading

Affiliation

Affiliation (Literal)

Akademischer Grad

OK Cancel

Facet / Filter

Refresh

x name: ju 2 choices S matched 3 none 1 Facet by choice

x name: b

clipboard - OpenRefine - Mozilla Firefox

clipboard - OpenRefine 127.0.0.1:3333/project?project=2221362040394 Search Help Wikidata next > last »

Add columns from reconciled column name

Add Property Preview Reset

Suggested Properties

	name	Beruf oder Beschäftigung	Geburtsort	Sterbeort	Ländercode
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

clipboard - OpenRefine 127.0.0.1:3333/project?project=1790589415416 ... Search Export Help 1 - 4 next > last »

Add columns from reconciled column name

Add Property Preview Reset

Ländercode name Beruf oder Geburtsort Sterbeort Ländercode  
remove configure remove configure remove configure remove configure

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

Facet / Filter Refresh

name: judgment 2 choices Sort by: matched 2 none 2 Facet by choice count

name: best can 26

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

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 - Mozilla Firefox

clipboard - OpenRefine 127.0.0.1:3333/project?project=2221362040394

OpenRefine clipboard Permalink

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

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 »

	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
2.	Twain, Mark Choose new match	Mathematiker Choose new match			Deutschland
3.	Lalit Kumar 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			
					Sonstiges Ausland, Unbekanntes Land

clipboard - OpenRefine - Mozilla Firefox

clipboard - OpenRefine 127.0.0.1:3333/project?project=2221362040394

### OpenRefine clipboard

Facet / Filter Undo / Redo 6 / 7

Refresh Reset All Remove All

**x name: judgment** change  
2 choices Sort by: name count  
(blank) 5  
matched 3  
Facet by choice counts

**x name: best candidate's score** change reset  
20.00 — 44.00  
Numeric 4 Non-numeric 0 Blank 2  
Error 0

**x name: judgment** change  
2 choices Sort by: name count

### 3 matching records (4 total)

Show as: rows records Show: 5 10 25 50 records « first < previous 1 - 3 next > last »

	name	Beruf oder Besc	Geburtsort	Sterbeort	Ländercode
1.	Weizenbaum, Joseph Choose new match	Informatiker Choose new match	Berlin Choose new match		
2.	Twain, Mark Choose new ma	Mathematiker			
3.	Lalit Kumar Choose new ma				Sonstiges Ausland, Unbekanntes Land

Extensions: Wikidata ▾

Facet ▾  
Text filter  
Edit cells ▾  
Edit column ▾  
Transpose ▾  
Sort...  
View ▾  
Reconcile ▾

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

Red arrows point to the "Add columns from reconciled values..." option in the context menu and the "Sterbeort" column header.

# 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 »

<input type="checkbox"/> All	<input type="checkbox"/> id	<input type="checkbox"/> name	<input type="checkbox"/> Beruf oder Beschäf	<input type="checkbox"/> Wirkungsort	<input type="checkbox"/> Darstellung des Wirkungsortes
		1. 1	Haddaway Choose new match	Komponist Choose new match	<a href="http://commons.wikimedia.org/wiki/Special:FilePath/DEU%20Koeln%20COA.svg">http://commons.wikimedia.org/wiki/Special:FilePath/DEU%20Koeln%20COA.svg</a>
				Sänger Choose new match	<a href="http://commons.wikimedia.org/wiki/Special:FilePath/AUT%20Kitzb%C3%B6chel%20COA.svg">http://commons.wikimedia.org/wiki/Special:FilePath/AUT%20Kitzb%C3%B6chel%20COA.svg</a>
				Musiker Choose new match	
		2. 2	Geßner, Johannes Choose new match	Arzt Choose new match	<a href="http://commons.wikimedia.org/wiki/Special:FilePath/Zurich-coat%20of%20arms.svg">http://commons.wikimedia.org/wiki/Special:FilePath/Zurich-coat%20of%20arms.svg</a>
				Mathematiker Choose new match	
				Naturwissenschaftler Choose new match	
		3. 3	Kuckart, Judith Choose new match	Schriftstellerin Choose new match	<a href="https://commons.wikimedia.org/wiki/Special:FilePath/Cityscape_Berlin.jpg">https://commons.wikimedia.org/wiki/Special:FilePath/Cityscape_Berlin.jpg</a>
				Regisseurin Choose new match	<a href="http://commons.wikimedia.org/wiki/Special:FilePath/Zurich-coat%20of%20arms.svg">http://commons.wikimedia.org/wiki/Special:FilePath/Zurich-coat%20of%20arms.svg</a>
				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

clipbaord - OpenRefine

OpenRefine clipboard Permalink

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	<input type="checkbox"/>
ort	<input type="checkbox"/>

Reconcile against type

Mus

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

\* 50

# Verwendung zusätzlicher lokaler Daten

clipboard - OpenRefine — Mozilla Firefox

clipboard - OpenRefine x +

openrefine.labs.lobid.org/proj 120% ... Search

OpenRefine clipboard Permalink

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 »

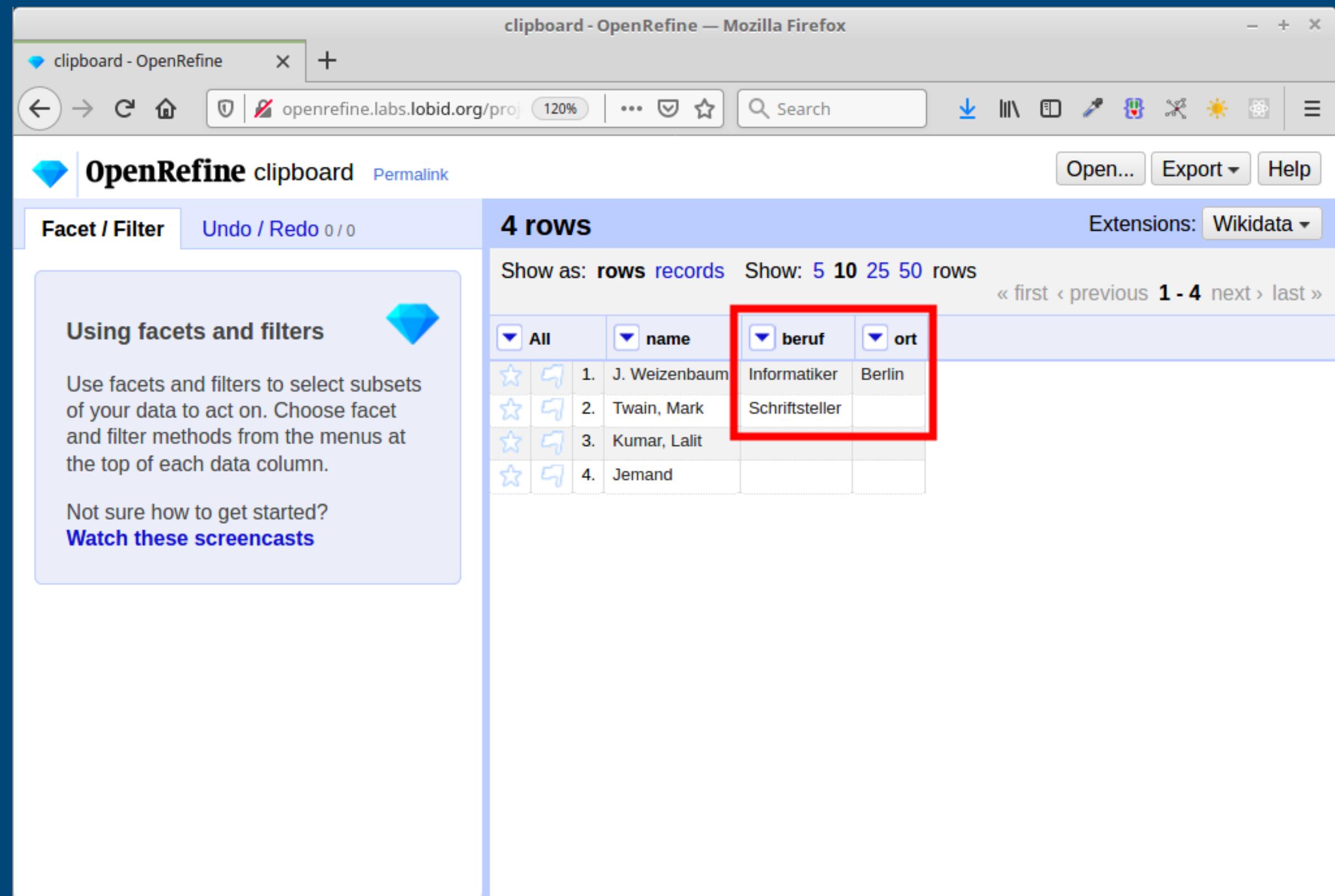
All name beruf ort

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

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](#)



		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 Permalink

Reconcile column "name"

» Access Service API

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

- Normdatenressource AuthorityResource
- Individualisierte Person DifferentiatedPerson
- Werk Work
- Körperschaft CorporateBody

Also use relevant details from other columns:

	Column Include? As Property
beruf	<input checked="" type="checkbox"/> Beruf
Beruf oder Beschäftigung	professionOrOccupation
Beruf oder Beschäftigung (Literal)	professionOrOccupationAsLiteral
Berufliche Beziehung	professionalRelationship

Select an item from the list:

- Beruf oder Beschäftigung professionOrOccupation
- Beruf oder Beschäftigung (Literal) professionOrOccupationAsLiteral
- Berufliche Beziehung professionalRelationship

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 Reconciling Cancel

The screenshot shows the OpenRefine reconciliation dialog for the 'name' column. On the left, there's a list of entity types: Normdatenressource (AuthorityResource), Individualisierte Person (DifferentiatedPerson), Werk (Work), and Körperschaft (CorporateBody). Below this, there are three radio button options: 'Reconcile against type: Person' (selected), 'Reconcile against no particular type', and 'Auto-match candidates with high confidence' (checked). At the bottom, there's a 'Maximum number of candidates to return' input field with a spinner, and buttons for 'Add Standard Service...', 'Start Reconciling', and 'Cancel'. On the right, there's a section titled 'Also use relevant details from other columns:' with a table. The table has two columns: 'Column Include? As Property' and 'Select an item from the list'. The first row shows 'beruf' with a checked checkbox next to 'Beruf'. Below this is a list of items: 'Beruf oder Beschäftigung' (professionOrOccupation), 'Beruf oder Beschäftigung (Literal)' (professionOrOccupationAsLiteral), and 'Berufliche Beziehung' (professionalRelationship). The entire 'Column Include? As Property' section is highlighted with a red box.

# Gezielte Suche nach Treffern

clipboard - OpenRefine - Mozilla Firefox

clipboard - OpenRefine

localhost:3333/project?project=232754877

Search

Open... Export Help

OpenRefine clipboard Permalink

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

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

2. Twain, Mark  
Schriftsteller

Search for match

Search for "J. Weizenbaum"

J. Weizenbaum

Select an item from the list:

Weizenbaum, Joseph 118820591  
1923-2008 | Informatiker

WEIZENBAUM, J. 187478929

Match New Item Don't Reconcile

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

Red arrow pointing to the 'Search for match' button in the reconciliation dialog.

# 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

The screenshot shows the OpenRefine interface with a blue header bar containing the OpenRefine logo and the tagline "A power tool for working with messy data.". A sidebar on the left lists "Create Project", "Open Project", "Import Project", and "Language Settings". The main content area displays a step in the "Create a project by importing data" wizard. It asks "What kinds of data files can I import?" and lists supported formats: TSV, CSV, \*SV, Excel (.xls and .xlsx), JSON, XML, RDF as XML, and Google Data documents with OpenRefine extensions. Below this, there's a section titled "Get data from" with options "This Computer" and "Web Addresses (URLs)". Under "Web Addresses (URLs)", a URL "https://blog.lobid.org/data/persons.csv" is entered. Buttons for "Add Another URL" and "Next »" are visible. A sidebar on the right lists "Clipboard", "Database", and "Google Data".

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

OpenRefine - Mozilla Firefox

OpenRefine A power tool for working with messy data.

Create Project    « 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.	Klopfer, Paul	1876	1967	personal
4.	Frank, Ellen	1904	1999	personal
5.	Lange, Annemarie	1907	1976	personal
6.	Hoff, August	1892	1971	personal
7.	Sebök, 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	1912	1975	personal

Parse data as: CSV / TSV / separator-based files    Character encoding:    Update Preview

Line-based text files    Columns are separated by:  commas (CSV)     Ignore first 0 line(s) at beginning of file

Fixed-width field text files     tabs (TSV)     Parse next 1 line(s) as column headers

PC-Axis text files     custom: ,     Discard initial 0 row(s) of data

JSON files    Escape special characters with \     Load at most 0 row(s) of data

MARC files     Use character " to enclose cells containing column separators

JSON-LD files     Column names (comma separated):     Parse cell text     Store blank rows

RDF/N3 files

RDF/N-Triples files

RDF/Turtle files

Version 3.3 [dbc732c]

Preferences    Help    About

persons csv - OpenRefine - Mozilla Firefox

persons csv - OpenRefine +

localhost:3333/project?project=2565457314865 ... Search

Open... Export Help

OpenRefine persons csv Permalink

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	<input type="checkbox"/>
DateOfDeath	<input type="checkbox"/>
Type	<input type="checkbox"/>

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. Göhl, Ernst 1907 1944 personal

A red box highlights the 'Individualisierte Person' option under 'Reconcile each cell to an entity of one of these types'. A red arrow points to this highlighted option.

# Project anlegen und Namen abgleichen

[openrefine.labs.lobid.org](http://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)

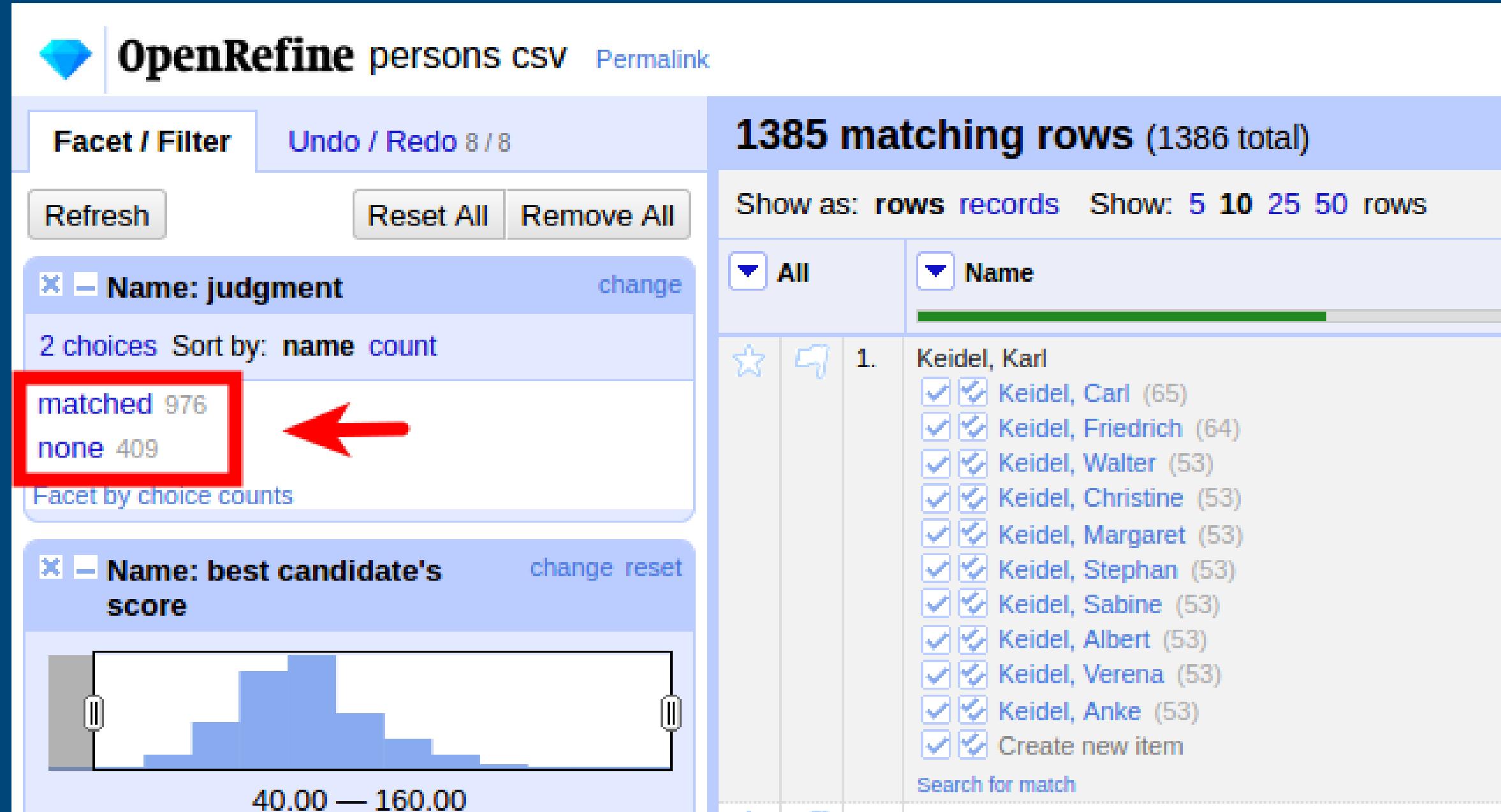
Show as: rows records Show: 5 10 25 50 rows

All Name

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 - Mozilla Firefox

persons csv - OpenRefine

localhost:3333/project?project=2565457314865

Search

Open... Export Help

OpenRefine persons csv Permalink

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

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

Extensions: Wikidata

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

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

3. Klopfer, Paul Hugo

Choose new match

1876 1967 personal

4. Frank, Ellen

Frank, Ellen (59)

Frank, Ellen (59)

Ellen, Ellen (44)

Ellen, R. F. (39)

Frank (33)

DateOfDeath Type

Facet Text filter

Edit cells Transform... (highlighted with red box and arrow)

Edit column Common transforms

Transpose

Sort...

View

Reconcile

Fill down

Blank down

Split multi-valued cells...

Join multi-valued cells...

Cluster and edit...

Replace

javascript:{}

# 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

persons csv - OpenRefine - Mozilla Firefox

persons csv - OpenRefine +  
localhost:3333/project?project=2565457314865 ... Search  
Open... Export Help

OpenRefine persons csv Permalink

Facet / Filter Refresh  
Name: judgn  
2 choices Sort by matched 911 none 475 Facet by choice co  
Name: best c

Custom text transform on column DateOfBirth

Expression Language General Refine Expression Language (GREL)  
value + "\*" No syntax error.

Preview History Starred Help

row	value	value + "*"
1.	1902	1902*
2.	1909	1909*
3.	1876	1876*
4.	1904	1904*
5.	1907	1907*
6.	1892	1892*
7.	1901	1901*

On error  keep original  set to blank  store error Re-transform up to 10 times until no change

OK Cancel

Frank, Ellen (59)  
Frank, Ellen (59)  
Ellen, Ellen (44)

The screenshot shows the OpenRefine interface with a modal dialog titled "Custom text transform on column DateOfBirth". In the "Expression" field, the value "value + '\*' is entered, which is highlighted with a red box. The "Language" dropdown is set to "General Refine Expression Language (GREL)". Below the expression, a preview table shows the transformation of several dates: 1902, 1909, 1876, 1904, 1907, 1892, and 1901, each followed by an asterisk. The "keep original" radio button is selected under "On error". The background of the main OpenRefine interface shows facets for "Name: judgn" and "Name: best c".

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

Daten bereit für neuen,  
verbesserten Abgleich

The screenshot shows the OpenRefine interface for a 'persons.csv' dataset. The main panel displays 1386 rows of data with columns for Name, DateOfBirth, DateOfDeath, and Type. A context menu is open over the first row under the 'Name' column header. The menu path 'Name' -> 'Reconcile' is highlighted with a red circle labeled '1'. The 'Actions' submenu is expanded, showing options like 'Match each cell to its best candidate' and 'Clear reconciliation data'. The 'Clear reconciliation data' option is highlighted with a red box and a red arrow labeled '2' pointing to it.

OpenRefine persons.csv Permalink

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

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

Extensions: Wikidata

All Name DateOfBirth DateOfDeath Type

1. Facet Text filter Edit cells Edit column Transpose Sort... View Reconcile Start reconciling...

2. Beckmann, Hannes Beckmann, Hanne Beckmann, Hanne Beckmann, ... (50) Create new item

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

Vorherigen Abgleich zurücksetzen

persons csv - OpenRefine - Mozilla Firefox

persons csv - OpenRefine +

localhost:3333/project?project=2565457314865

Reconcile column "Name"

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
<b>Sterbedatum</b>	dateOfDeath
Östlichster Längengrad	eastermostLongitude
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

javascript:{}

Frank, EICH (55) Elke, Elke (44)

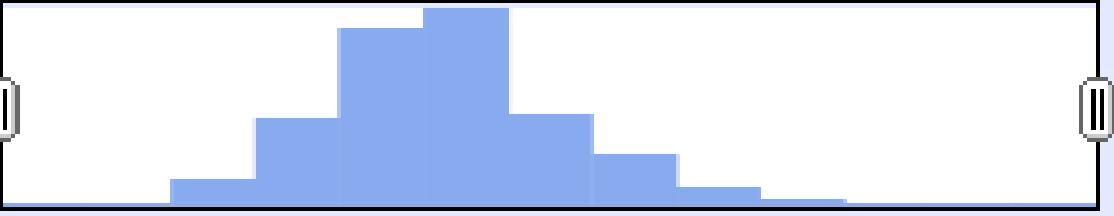
 | OpenRefine persons csv [Permalink](#)

**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  
  
40.00 — 170.00

**1386 rows**

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

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

>95% automatisch abgeglichen

# 2. Beispiel: Beruf & Parteizugehörigkeit

The screenshot shows the OpenRefine interface with a blue sidebar on the left containing options: Create Project (selected), Open Project, Import Project, and Language Settings. The main area displays a step-by-step guide for creating a project by importing data. The current step is "Get data from Web Addresses (URLs)". It instructs the user to enter one or more web addresses (URLs) pointing to data to download. A text input field contains the URL <https://blog.lobid.org/data/politicians.csv>. Below the input field are two buttons: "Add Another URL" and "Next >". To the right of the input field, there is a list of alternative data sources: "Clipboard", "Database", and "Google Data".

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

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

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

Add Another URL    Next >

Web Addresses (URLs)

Clipboard

Database

Google Data

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 « Start Over Configure Project name politicians csv Tags Create Project »

Open Project Import Project Language Settings

**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 " character to enclose cells containing

	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

Version 3.3 [cf578cf]

Preferences Help About

politicians csv - OpenRefine - Mozilla Firefox

politicians csv - OpenRefine +

localhost:3333/project?project=2116118209 ... Search

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
parliamentary_group	<input type="checkbox"/>
GND_Kennung	<input type="checkbox"/>
Wikidata_ID	<input type="checkbox"/>

Reconcile against type:

Reconcile against no particular type

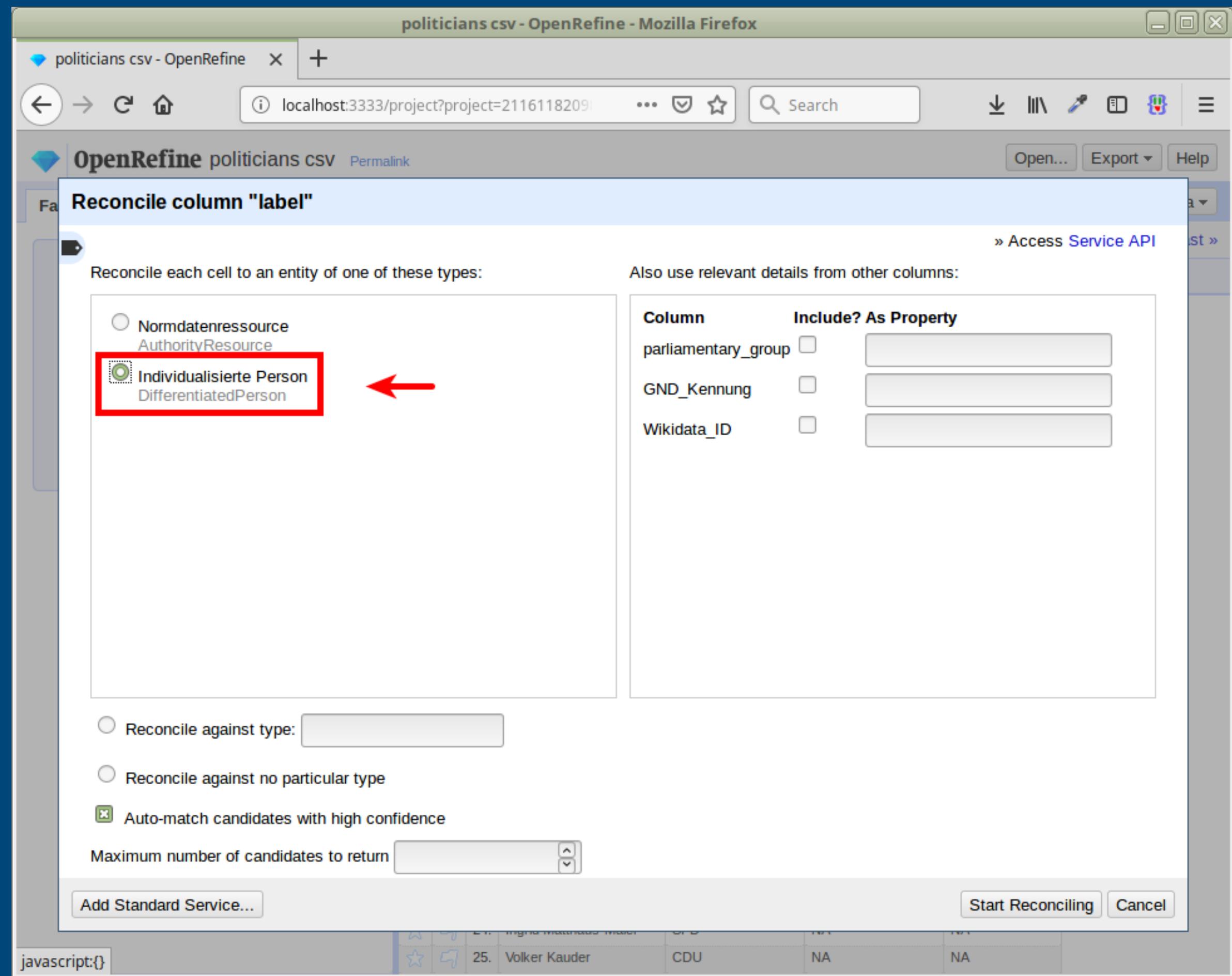
Auto-match candidates with high confidence

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](http://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 Rossmannith	
	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Rossmannith, Kurt (85)	
	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Rossmannith, Kurt (82)	
	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Rossmannith, Alois (57)	
	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Rossmannith, Peter (57)	
	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Rossmannith, Sabine (57)	
	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Rossmannith, Christian (57)	
	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Rossmannith, Sandra (56)	
	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Rossmannith, 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 Identifier mappen

# Welche Abkürzungen kommen in den Daten vor?

The screenshot shows the OpenRefine interface with a dataset titled "politicians csv" containing 358 rows. The left sidebar displays facets for "parliamentary\_group" and "label". A red box highlights the "parliamentary\_group" facet, which lists categories like CDU, FDP, fraktionslos, GRUENE, LINKE, NA, PDS, and SPD, each with a count. A red circle highlights the "parliamentary\_g" column header in the main data grid. A red arrow points to the "Text facet" option in the context menu that appears when right-clicking on this header. The menu also includes options like Text filter, Numeric facet, Timeline facet, Scatterplot facet, Custom text facet..., Custom Numeric Facet..., Customized facets, and Reconcile.

parliamentary_group	label	Facet
CDU 118	Rexrodt, Günter	Text facet
FDP 50	Hans-Eberhard Urbaniak	
fraktionslos 1	Urbaniak, Regina (53)	
GRUENE 30	Urbaniak, Jan (52)	
LINKE 7	Urbaniak, Beata (52)	
NA 31	Urbaniak, Tamina (52)	
PDS 11	Urbaniak, Rafał (52)	
SPD 110	Urbaniak, Petra (52)	
	Urbaniak, Gertrud (51)	
	Urbaniak, Thomas (51)	
	Urbaniak, James (51)	
	Urbaniak, Michał (51)	
	Create new item	
	Search for match	
	Schwanhold, Ernst	
	Choose new match	
	Kurt Rossmannith	
	Rossmannith, Kurt (79)	
	Rossmannith, Kurt (78)	
	Rossmannith, Jonas (53)	
	Rossmannith, Frieda (53)	
	Roßmannith, Birgit (53)	
	Rossmannith, Sandra (53)	
	Rossmannith, Peter (52)	
	Rossmannith, Alois (52)	
	Rossmannith, Sabine (52)	
	Roßmannith, Richard (52)	
	Create new item	
	Search for match	

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

The screenshot shows the lobid gnd search interface. A red box labeled 1 highlights the search bar containing the query "sozialdemokratische partei deutschlands". The search results table displays four entries. The fourth entry, which is highlighted with a red box labeled 3, is "Sozialdemokratische Partei Deutschlands" (Körperschaft, Deutschland, 2022139-3). To the right of the results is a sidebar titled "Ergebnisse eingrenzen: Entitätstyp" with several filter options. One filter, "GND-Sachgruppe" (Parteien, Politische Organisationen), is highlighted with a red box labeled 2.

665 Treffer, zeige 1 bis 10:				
Sozialdemokratische Partei Deutschlands. Partevorstand	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)

# Gesuchte Mappings

- SPD → [https://d-nb.info/gnd/???](https://d-nb.info/gnd/)
- CDU → [https://d-nb.info/gnd/???](https://d-nb.info/gnd/)
- GRUENE → [https://d-nb.info/gnd/???](https://d-nb.info/gnd/)
- FDP → [https://d-nb.info/gnd/???](https://d-nb.info/gnd/)
- LINKE → [https://d-nb.info/gnd/???](https://d-nb.info/gnd/)
- PDS → [https://d-nb.info/gnd/???](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:

The screenshot shows the OpenRefine interface with the following details:

- Facet / Filter** tab is selected.
- parliamentary\_group** facet is active, with 8 choices: fraktionslos (1), GRUENE (30), LINKE (7), NA (31), PDS (11), SPD (110). A red box highlights the "edit" button next to the GRUENE choice.
- Undo / Redo** button shows 3 / 3.
- Refresh**, **Reset All**, and **Remove All** buttons are present.
- 358 rows** are shown in the main pane.
- Show as:** rows records (selected) and **Show:** 5 10 25 50.
- All** and **label** facets are displayed.
- The main list shows:
  - 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** input field is at the bottom.

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

The screenshot shows the OpenRefine interface for a project titled "politicians csv". The top navigation bar includes "OpenRefine", the project name, and "Permalink". Below the navigation is a toolbar with "Facet / Filter" (selected), "Undo / Redo 8 / 8", and two buttons: "Extract..." (highlighted with a red arrow) and "Apply...". A "Filter:" input field is present. The main workspace displays a list of 318 matching rows (358 total). The columns are labeled "All" and "label". The first four rows are listed:

All	label	Row Number	Value
<input type="star"/>	<input type="minus"/>	1.	Rexrodt, Günter Choose new match
<input type="star"/>	<input type="minus"/>	2.	Urbaniak, Hans-Eberhard Choose new match
<input type="star"/>	<input type="minus"/>	3.	Schwanhold, Ernst Choose new match
<input type="star"/>	<input type="minus"/>	4.	Kurt Rossmannith <input checked="" type="checkbox"/> Rossmannith, Kurt (87) <input checked="" type="checkbox"/> Rossmannith, Kurt (85) <input type="checkbox"/> Rossmannith (800)

# 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 interface for a project titled "politicians csv". The main area displays a list of 358 rows, with the first few entries being:

		label
1.		Rexrodt, Günter Choose new match
2.		Hans-Eberhard Urbaniak <input checked="" type="checkbox"/> Urbaniak, Regina (53) <input checked="" type="checkbox"/> Urbaniak, Jan (52) <input checked="" type="checkbox"/> Urbaniak, Beata (52) <input checked="" type="checkbox"/> Urbaniak, Tamina (52) <input checked="" type="checkbox"/> Urbaniak, Rafal (52) <input checked="" type="checkbox"/> Urbaniak, Petra (52) <input checked="" type="checkbox"/> Urbaniak, Gertrud (51) <input checked="" type="checkbox"/> Urbaniak, Thomas (51) <input checked="" type="checkbox"/> Urbaniak, James (51) <input checked="" type="checkbox"/> Urbaniak, Michal (51) <input checked="" type="checkbox"/> Create new item Search for match
3.		Schwanhold, Ernst Choose new match

The reconciliation interface on the left shows the following steps:

0. Create project
1. Reconcile cells in column label to type DifferentiatedPerson

A red arrow points from the text "358 rows" to the reconciliation results table. A red box highlights the "Apply..." button in the toolbar above the reconciliation interface.

politicians csv - OpenRefine - Mozilla Firefox

politicians csv - OpenRefine +

localhost:3333/project?project=2116118209 ... Search

Open... Export Help

OpenRefine politicians csv Permalink

Facet / Filter

Filter:

0. Create p

1. Reconcile Differen

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"
}
```

Wikidata next > last »

ata\_ID

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 X +

localhost:3333/project?project=2116118209 ... Search

Open... Export Help

OpenRefine politicians csv Permalink

Facet / Filter Undo / Redo 3 / 7

Refresh Reset All Remove All

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

358 rows Extensions: Wikidata

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

All label parliamentary\_g affiliation GND\_Kennung

Facet Text filter

Edit cells Edit column Transpose

Sort... View

Reconcile Start reconciling...

Create new item Search for match

Schwanhold, Ernst Actions

Choose new match Match each cell to its best candidate

Kurt Rossmannith Copy reconciliation data...

Rossmannith, Kurt Create a new item for each cell

Rossmannith, Kurt Create one new item for similar cells

Rossmannith, Jonas (53) Match all filtered cells to...

Rossmannith, Frieda (53)

Rossmannith, Birgit (53)

Rossmannith, Sandra (53)

Rossmannith, Peter (52)

Rossmannith, Alois (52)

Rossmannith, Sabine (52)

Rossmannith, Richard (52)

Create new item Discard reconciliation judgments

Search for match Clear reconciliation data

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

Süssmuth, Rita NA NA NA

1 2

The screenshot shows the OpenRefine interface for a 'politicians csv' project. On the left, there are two facets: 'label: judgment' (with 2 choices, 185 matched, 173 none) and 'label: best candidate's score' (with 357 Numeric, 0 Non-numeric, 1 Blank). A red circle labeled '1' points to the 'Remove All' button in the facet bar. On the right, the main workspace displays 358 rows of data with columns for 'label', 'parliamentary\_g', 'affiliation', and 'GND\_Kennung'. A context menu is open over the 'label' column, with a red circle labeled '2' pointing to the 'Clear reconciliation data' option under the 'Actions' section. The menu also includes options like 'Start reconciling...', 'Create new item', 'Search for match', and 'Discard reconciliation judgments'.

politicians csv - OpenRefine - Mozilla Firefox

politicians csv - OpenRefine +

localhost:3333/project?project=2116118209 ... Search

Open... Export Help

**OpenRefine politicians csv Permalink**

**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	<input type="checkbox"/>
affiliation	<input checked="" type="checkbox"/> af

Select an item from the list:

- Beziehung, Bekanntschaft, Freundschaft acquaintanceOrFriendship
- Affiliation** affiliation
- Affiliation (Literal) affiliationAsLiteral
- Biografische oder historische Angaben biographicalOrHistoricalInformation
- 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

The screenshot shows the 'Reconcile column "label"' dialog in OpenRefine. On the left, under 'Reconcile each cell to an entity of one of these types:', the 'Individualisierte Person' option is selected. On the right, under 'Also use relevant details from other columns:', the 'affiliation' column is selected with the 'Include? As Property' checkbox checked. Below this, a list of service types is shown, with 'Affiliation' highlighted by a red box and a red arrow pointing to it. Other options include 'Affiliation (Literal)', 'Biografische oder historische Angaben', 'In Beziehung stehende Körperschaft', and 'In Beziehung stehendes Geografikum'. At the bottom, there are options for reconciling against type, no particular type, and auto-match candidates with high confidence, along with a 'Maximum number of candidates to return' input field.

 | 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 Rossmannith
	<input checked="" type="checkbox"/>	Rossmannith, Kurt (85)
	<input checked="" type="checkbox"/>	Rossmannith, Kurt (82)
	<input checked="" type="checkbox"/>	Rocker, Kurt (64)
	<input checked="" type="checkbox"/>	Pfeiffer, Kurt (63)
	<input checked="" type="checkbox"/>	Thürk, Kurt (63)
	<input checked="" type="checkbox"/>	Schmücker, Kurt (63)
	<input checked="" type="checkbox"/>	Malangré, Kurt (63)
	<input checked="" type="checkbox"/>	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

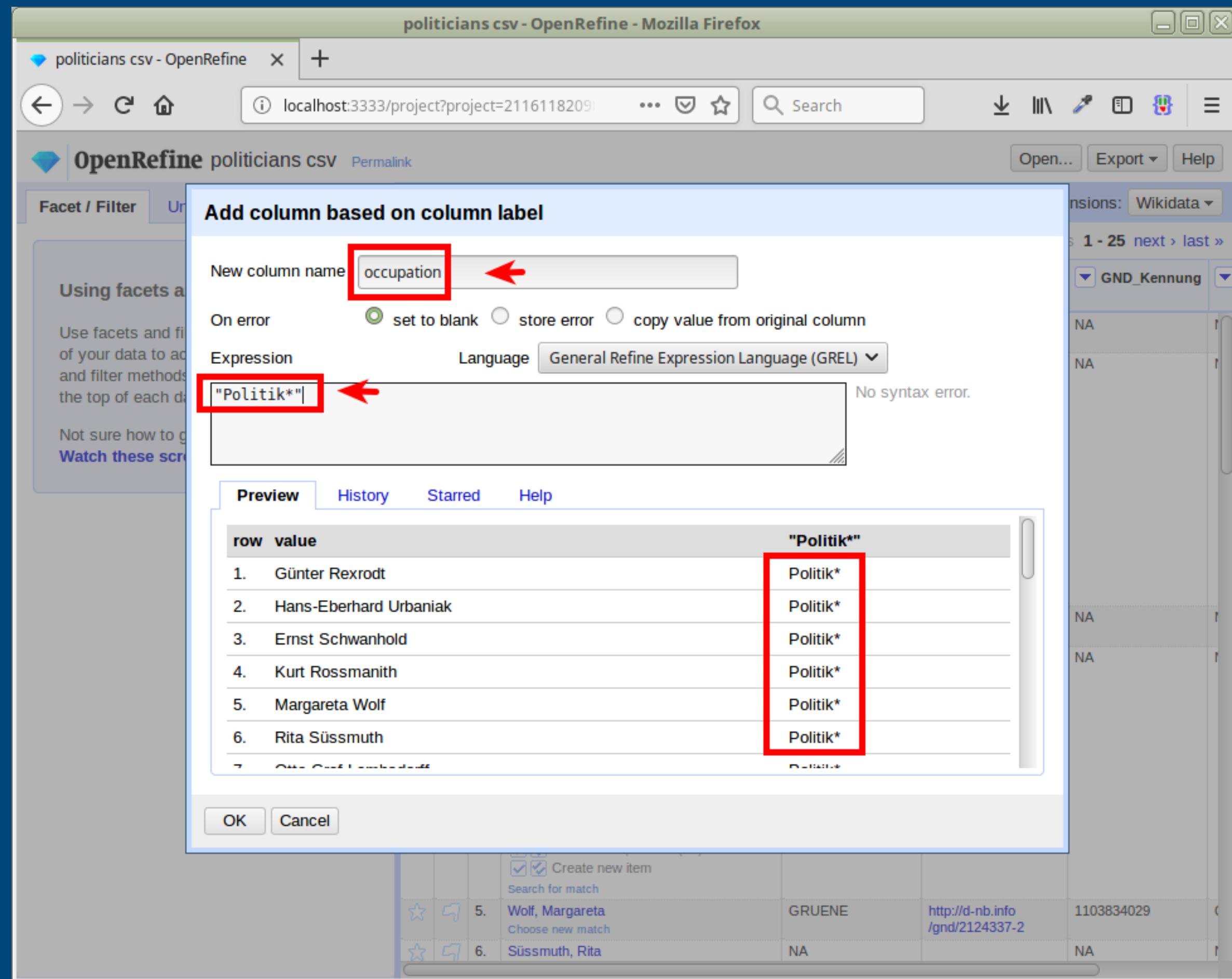
**label: judgment** change  
2 choices Sort by: name count  
matched 264  
none 94  
Facet by choice counts

**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.	Facet		
2.	Text filter		
3.	Edit cells		
4.	Edit column	Split into several columns...	b.info/g
	Transpose	Join columns...	b.info/g
	Sort...	Add column based on this column...	
	View	Add column by fetching URLs...	
	Reconcile	Add columns from reconciled values...	
	Schmäcker, Kurt	Rename this column	
	Malangré, Kurt (6)	Remove this column	
	Kiesinger, Kurt Ge		



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 ... Search

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	<input checked="" type="checkbox"/> Beruf oder Beschäftigung
parliamentary_group	<input type="checkbox"/>
affiliation	<input checked="" type="checkbox"/> af

Select an item from the list:

- Beziehung, Bekanntschaft, Freundschaft acquaintanceOrFriendship
- Affiliation affiliation
- Affiliation (Literal) affiliationAsLiteral
- Biografische oder historische Angaben biographicalOrHistoricalInformation
- 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

The screenshot shows the 'Reconcile column "label"' dialog in OpenRefine. On the left, under 'Reconcile each cell to an entity of one of these types:', the 'Individualisierte Person' (DifferentiatedPerson) option is selected. On the right, under 'Also use relevant details from other columns:', there is a table with two rows. The first row, 'occupation', has a checked checkbox in the 'Include? As Property' column and the value 'Beruf oder Beschäftigung'. The second row, 'affiliation', also has a checked checkbox and the value 'af'. Both of these rows are highlighted with a red border. Below the table, a list of entities is shown, including 'Beziehung, Bekanntschaft, Freundschaft', 'Affiliation', 'Affiliation (Literal)', 'Biografische oder historische Angaben', 'In Beziehung stehende Körperschaft', and 'In Beziehung stehendes Geografikum'. At the bottom of the dialog, there are three radio button options: 'Reconcile against type:' (with a text input field), 'Reconcile against no particular type', and 'Auto-match candidates with high confidence' (which is checked). There is also a 'Maximum number of candidates to return' input field with a spinner, and buttons for 'Add Standard Service...', 'Start Reconciling', and 'Cancel'.

 **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 Rossmannith
	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Rossmannith, Kurt (87)
	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Rossmannith, Kurt (85)
	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Rocker, Kurt (69)
	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Pfeiffer, Kurt (68)
	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Thürk, Kurt (68)
	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Schmücker, Kurt (68)
	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Malangré, Kurt (68)
	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> 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

8 records

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

All	id	name	Beruf oder Beschäftigung	Wirkungsort
1.	1	Haddaway	Komponist 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	Arzt Choose new match	Zürich Choose new match
			Mathematiker Choose new match	
			Naturwissenschaftler Choose new match	
3.	3	Kuckart, Judith Choose new match	Schriftstellerin Choose new match	Berlin Choose new match
			Regisseurin Choose new match	Zürich Choose new match
			Tänzerin Choose new match	
			Choreografin Choose new match	

Open... Export Help

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

Extensions: Wikidata

8 records

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

All	id	name	Beruf oder Beschäft	Wirkungsort	D
1.	1	Haddaway Choose new match	Komponist Choose new match	Köln Choose new match	<a href="http://commons.wikimedia.org/wiki/Special:FilePath/Zurich-coat%20of%20arms.svg">http://commons.wikimedia.org/wiki/Special:FilePath/Zurich-coat%20of%20arms.svg</a>
			Sänger Choose new match	Kitzbühel Choose new match	<a href="http://commons.wikimedia.org/wiki/Special:FilePath/Zurich-coat%20of%20arms.svg">http://commons.wikimedia.org/wiki/Special:FilePath/Zurich-coat%20of%20arms.svg</a>
			Musiker Choose new match		
2.	2	Geßner, Johannes Choose new match	Arzt Choose new match	Zürich Choose new match	<a href="http://commons.wikimedia.org/wiki/Special:FilePath/Zurich-coat%20of%20arms.svg">http://commons.wikimedia.org/wiki/Special:FilePath/Zurich-coat%20of%20arms.svg</a>
			Mathematiker Choose new match		
			Naturwissenschaftler Choose new match		
3.	3	Kuckart, Judith Choose new match	Schriftstellerin Choose new match	Berlin Choose new match	<a href="https://commons.wikimedia.org/wiki/Special:FilePath/Cityscape_Berlin.jpg">https://commons.wikimedia.org/wiki/Special:FilePath/Cityscape_Berlin.jpg</a>
			Regisseurin Choose new match	Zürich Choose new match	<a href="http://commons.wikimedia.org/wiki/Special:FilePath/Zurich-coat%20of%20arms.svg">http://commons.wikimedia.org/wiki/Special:FilePath/Zurich-coat%20of%20arms.svg</a>
			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

Extensions: Wikidata

Edit Wikidata schema  
Manage Wikidata account  
Import schema  
Export schema  
Upload edits to Wikidata  
Export to QuickStatements

# Flexibler Export mit Templating

nk

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.	Rossmannith, 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	<a href="https://d-nb.info/gnd/5010217-5">https://d-nb.info/gnd/5010217-5</a> NA

Open... Export ▾ Help

OpenRefine project archive to file

Tab-separated value

Comma-separated value

HTML table

Excel (.xls)

Excel 2007+ (.xlsx)

ODF spreadsheet

Custom tabular exporter...

SOI Exporter

Templating...

OpenRefine project archive to Google Drive...

Google Sheets

Wikidata edits...

QuickStatements file

Wikidata schema

Facet / Filter

Refresh

label: judge

2 choices Sort by

matched 252

none 66

Facet by choice count

label: best

Numeric  
357Error  
0

## Templating Export

Prefix

1

Row Template

2

```
{  
  "rows" : [  
    {  
      "label" : {{jsonize(cells["label"].value)}},  
      "occupation" : {{jsonize(cells["occupation"].value)}},  
      "parliamentary_group" : {{jsonize(cells["parlamentar"])}},  
      "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

Export Cancel

```
{  
  "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 Rossmanith",  
      "occupation" : "Politik*",  
      "parliamentary_group" : "CDU",  
      "affiliation" : "https://d-nb.info/gnd/7230-8",  
      "GND_Kennung" : "NA",  
      "Wikidata_ID" : "NA"  
    }  
  ]  
}
```

Extensions: Wikidata

- 10 next &gt; last &gt;

D\_Kennung

D_Kennung	Wikidata_ID
NA	NA
4029	Q18
NA	NA
NA	NA
NA	NA

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

W3C Entity Reconciliation Cor +<https://github.com/reconciliation-api>

Search



≡



Search or jump to...

Pull requests Issues Marketplace Explore



## W3C Entity Reconciliation Community Group

Writing specs and tooling to improve data matching on the web

<https://www.w3.org/community/recon...> [public-reconciliation@w3c.org](mailto:public-reconciliation@w3c.org)[Repositories 4](#)[Packages](#)[People 28](#)[Teams](#)[Projects](#)[Settings](#) Find a repository...

Type: All

Language: All

Customize pins

[New](#)[testbench](#)

Web application to try out reconciliation services interactively

JavaScript

MIT

3

0

7

0

Updated 12 days ago

[census](#)

Inventory of the reconciliation ecosystem and its neighbourhood

Ruby

3

3

0

0

Updated on Oct 26

[specs](#)

Specifications of the reconciliation API

reconciliation

reconciliation-service

reconciliation-api

HTML

2

11

23

2

Updated on Oct 24

## Top languages

HTML JavaScript Ruby

## People

28 &gt;

[Invite someone](#)

# 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 [load all endpoints via JSONP](#).

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

# 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		<b>The myths of security : what the computer security industry doesn't want you to know / John Viega with introduction by Elia Zafrani</b>	<a href="#">Edit Record</a> <a href="#">Holdings</a> <a href="#">...</a>
		Book By Viega, John. (Beijing ; Cambridge ; Sebastopol, CA : O'Reilly, ©2009.)	Language: English MMS ID: 99229211100121
		Subject: Internet industry—Security measures. Computer security. Data protection. and others	ISBN: 9780596523022 and others
		Modification Date: 18/08/2020 13:13	Record number: (OCOlc)263978539 
		Creation Date: 03/02/2020 09:30	
		Edition: 1st ed.	
<a href="#">Physical (2)</a> 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

**Mapping Editor**

**BK**      **RVK**      **not saved**

**56.81 Wohnungsbau** = **ZH 5000 - ZH 5960 Wohnungsbau**

Dennis Tobola

**Concordances**      **Mapping Search**      **Mapping Navigator**

**Mappings:** C L W      **Recommendations:** CR CC CO

From	To	Creator	Date
<b>C Concordance Registry</b>			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	
<b>L Local</b>			No Mappings
<b>W Wikidata-Mappings</b>			No Mappings
<b>CR colo-conc Recommendations</b>			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		

**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

**RVK Regensburg**

Classification 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 Bauingenieurwesen  
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/c000000011/\\_source](http://localhost:9200/loc/work/c000000011/_source)

# 1. Einzeldatensätze als JSON

z.B. [http://localhost:9200/loc/work/c000000011/\\_source](http://localhost:9200/loc/work/c000000011/_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	<a href="#">/works/c000000011</a>
		2.	Text filter	<a href="#">/works/c000000020</a>
		3.	<a href="#">Edit cells</a>	<a href="#">/works/c000000026</a>
		4.	<a href="#">Edit column</a>	<a href="#">Split into several columns...</a>
		5.	<a href="#">Transpose</a>	<a href="#">Add column based on this column...</a>

[Edit cells](#)

[Edit column](#)

[Transpose](#)

[Sort...](#)

[View](#)

[Reconcile](#)

[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](#)

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/c000000011>

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/c000000011/_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/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 **json**

Throttle delay

5000

milliseconds

On error

 set to blank  store error Cache responsesHTTP headers to be used when fetching URLs: [Show](#)**Formulate the URLs to fetch:**

Expression

Language

General Refine Expression Language (GREL) **"http://localhost:9200/loc/work/" + split(value, '/')[-1] + "/\_source"**

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

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

son



1. <http://id.loc.gov/resources/works/c000000011>

Facet

Text filter

Edit cells

**Edit column**

Transpose

Sort...

View

Reconcile

Renaming

Remove this column

Move column to beginning

Move column to end

Move column left

Move column right

localhost:3000/context.json", "id": "http://id.loc.gov/resources/v...  
"encodingLevel": [{"type": "EncodingLevel", "code": "5"}], "time", "@value": "2004-11-19T10:19:25"}, "creationDate", "@value": "1993-10-26"}, "descriptionConventions": [{"type": "DescriptionConventions", "value": "lcrb"}, "type": "DescriptionConventions"}], "generationPro...  
f-s  
sif  
'sc  
tio  
cto  
J.  
era  
icA  
nc  
'Fr  
ct"]  
es  
"m  
/11  
n f  
ers  
in .  
124

[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**  set to blank  store error  copy value from original column[Expression](#)[Language](#)General Refine Expression Language (GREL) [▼](#)

No syntax error.

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

row	value	
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"}]}}, "label": "Letters from an American farmer. French. 1784	value.parseJson().label

clipboard - OpenRefine X +

127.0.0.1:3333/project?project=1657999837942

Search

Open... Export ▾ Help

Clipboard

Facet / Filter Undo / 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 records Show: 5 10 25 50 rows

All	uri	author	label
1.	<a href="http://id.loc.gov/resources/works/c000000011">http://id.loc.gov/resources/works/c000000011</a>	St. John de Crèvecoeur, J. Hector, 1735-1813.	Letters from an American farmer. French. 1784
2.	<a href="http://id.loc.gov/resources/works/c000000020">http://id.loc.gov/resources/works/c000000020</a>	Howells, William Dean, 1837-1920.	Italian journeys /
3.	<a href="http://id.loc.gov/resources/works/c000000026">http://id.loc.gov/resources/works/c000000026</a>	Ryan, George, active 1811.	Report of the trial of George Ryan : before the S May term 1811, for highway robbery.
4.	<a href="http://id.loc.gov/resources/works/c000000028">http://id.loc.gov/resources/works/c000000028</a>	Miscellaneous Pamphlet Collection (Library of Congress)	The devil upon crutches in England; or, Night sc
5.	<a href="http://id.loc.gov/resources/works/c000000100">http://id.loc.gov/resources/works/c000000100</a>	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 data 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 rows

Show as: [rows](#) [records](#)

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

[All](#)[title](#)[uri](#)

		1.	<a href="#">Letters from an American farmer</a>
		2.	<a href="#">Italian journeys</a>
		3.	<a href="#">Report of the trial of George Ryan</a>
		4.	<a href="#">The devil upon crutches in England</a>
		5.	<a href="#">Educational interests of Colorado</a>

<a href="http://id.loc.gov/resources/works/11#Work240-15">http://id.loc.gov/resources/works/11#Work240-15</a>
<a href="http://id.loc.gov/resources/works/c000000020">http://id.loc.gov/resources/works/c000000020</a>
<a href="http://id.loc.gov/resources/works/c000000026">http://id.loc.gov/resources/works/c000000026</a>
<a href="http://id.loc.gov/resources/works/c000000028">http://id.loc.gov/resources/works/c000000028</a>
<a href="http://id.loc.gov/resources/works/c0000000100">http://id.loc.gov/resources/works/c0000000100</a>

# 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



W3C Community Group  
Draft Report



# 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.





ReSpec

W3C Community Group  
Draft Report

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

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"  
      }  
    ]  
  }  
}
```

specs/latest at master · recon



https://github.com/reconciliation-api/specs/tree/n...



Search



Search or jump to...



Pull requests

Issues

Marketplace

Explore



reconciliation-api / specs

Public

Unwatch

16

Star

19

Fork

3

Code

Issues 29

Pull requests 3

Discussions

Actions

Projects

Security

Insights

...

master

specs / latest /

Go to file

Add file

...



wetneb and tfmorris Make 'query' field in reconciliation queries optional. (#53) ...

✓ c348b86 20 days ago

History

..

examples

Make 'query' field in reconciliation queries optional. (#53)

20 days ago

schemas

Make 'query' field in reconciliation queries optional. (#53)

20 days ago

index.html

Make 'query' field in reconciliation queries optional. (#53)

20 days ago



# 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:

Reconcile

Type:

- Normdatenressource  
AuthorityResource
- Körperschaft  
CorporateBody
- Konferenz oder Veranstaltung  
ConferenceOrEvent
- Schlagwort  
SubjectHeading
- Werk  
Work
- Individualisierte Person  
DifferentiatedPerson
- Familie  
Family
- Custom:
- Reconcile against no particular type

**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](#)

**Properties:**

Beruf oder Beschäftigung	Politik*	Delete
Affiliation	https://d-nb.info/gnd/2022139-3	Delete

Brandt, Willy
107.3416

Normdatenressource (AuthorityResource),  
 Individualisierte Person (DifferentiatedPerson)

Brandt, Hugo
73.07417

Normdatenressource (AuthorityResource),  
 Individualisierte Person (DifferentiatedPerson)

Brandt, Willy
71.96527

Normdatenressource (AuthorityResource),  
 Individualisierte Person (DifferentiatedPerson)

Brandt, Willy
71.96527

Normdatenressource (AuthorityResource),  
 Individualisierte Person (DifferentiatedPerson)

Brandt, Willy
71.90846

Normdatenressource (AuthorityResource),  
 Individualisierte Person (DifferentiatedPerson)

Prölß, Willy
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