

OpenRefine

Bearbeitung von Daten der Forschungsdokumentation mit OpenRefine (3.3-beta) und Wikidata

OpenRefine - Start & Installation

- OpenRefine ist eine plattformunabhängige JAVA-Applikation zur Bearbeitung tabellarischer Daten.
- Download und Installation via www.openrefine.org
- Standardinstallation:
 - Download der Paketdatei
 - Entpacken
 - Starten der Datei `refine`
- Nach Aufruf der Datei `refine` startet OpenRefine in einem Browserfenster mit der lokalen Adresse <http://127.0.0.1:3333>

Neuanlage eines OpenRefine-Projektes

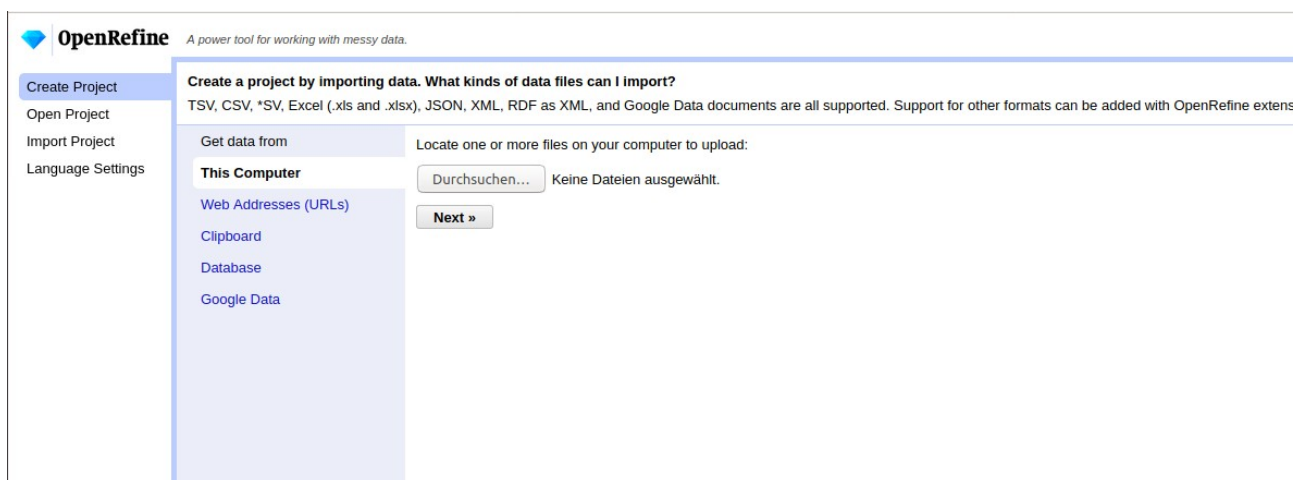


Abbildung 1: Create Project

- „Create Project“ (vgl. Abbildung 1) Auswahl der Datenimportmöglichkeiten nach OpenRefine
 - Dateiupload: Spreadsheet-Dokumente (TSV, CSV; XLS, ODT)
 - JSON, XML
 - Copy & Paste über Zwischenablage von Dateiinhalten in oben genannten Formaten

OpenRefine-Workshop

- Direkter Datenbankzugriff
- Nach Auswahl der Datenquelle wird ein Preview (vgl. Abbildung 2) angezeigt, um festzustellen, ob OpenRefine die Quelldatei korrekt importieren kann, bzw. ob diese wohl formatiert ist. Diverse Einstellungsmöglichkeiten je nach Importformat sind noch möglich (bspw. ob die erste Zeile Überschriften enthält)
- Angabe des „Projektname“ und ggf. Angabe von Tags zur Suche und Beschreibung von OpenRefine Projekten

The screenshot shows the OpenRefine interface. At the top, there's a navigation bar with 'Start Over' and 'Configure Parsing Options'. Below that is a table with 14 rows of partner data. The columns include PARTNER_ID, Stammdatensatz_ID, PARTNER_DE, PARTNER_EN, ACRONYM, STRASSE, POSTLEITZAHL, ORT, LAND_ID, TELEFON, FAX, EMAIL, URL, ANMERKUNGEN, BEARBEITER, and AENDERUNG. Below the table is a 'Parse data as' panel with options for 'Excel files', 'JSON files', 'Line-based text files', 'CSV / TSV / separator-based files', and 'Fixed-width field text files'. There are also checkboxes for 'Ignore first', 'Parse next', 'Discard initial', and 'Load at most' with corresponding values for lines, columns, rows, and data.

1.	9635	9635	Landwirtschaftliche Universität Tirana	Agricultural University of Tirana	UBT	Koder Kamez	1029	Tirana	8			www.ubt.edu.al	10		Anna	Fri Aug 31 00:00 CEST 2018
2.	10368	10158	"Fan S. Noli" University, Korçe, Albania	"Fan S. Noli" University, Korçe, Albania					8				10		Anna	Fri Jul 13 00:00 CEST 2018
3.	9704	9704	Polytechnische Universität Tirana	Polytechnic University of Tirana	PUoT			Tirana	8				10		Anna	Fri Aug 31 00:00 CEST 2018
4.	10158	10158	Fan Noli Universität	Fan Noli University	UNKO	Shetitorgji, "Rrindasi"	7001	Korça	8			www.unkorce.edu.al	10		Anna	Fri Jun 29 00:00 CEST 2018
5.	11895	11895	Universität Tirana	University of Tirana	UT	Deshmoret e Kombit	1000	Tirana	8				10		Anna	Fri Sep 28 00:00 CEST 2018
6.	3276		General Directorate of Forests and Pastures, Forest Protection	General Directorate of Forests and Pastures, Forest Protection		Rruga Sami Frasheri 4		Tirana	8					5708		Tue Nov 09 13 CET 2004
7.	10862		Menaxhim Agro-rrejtësor e Ekonomik	Agro-Environment and Economic Management Center	AEEM	Zef Zuberi, N5	1016	Tirana	8			http://www.karavista-guest-house-tourism.org/index.php/en/j98nby			6497	Thu Oct 01 15:00 CEST 2015
8.	6850		Institut Algérien du Pétrole	Algerian Institute for Petrol	IALPET	Avenue du 10 Novembre	35000	Boumerdes	12				100		Marian	Wed Oct 22 12:00 CEST 2003
9.	8702	8702	Nationale Universität La Plata	National University of La Plata	UNLP		1925	La Plata	32				10		Anna	Fri Aug 31 00:00 CEST 2018
10.	8700	8700	Nationale Universität Córdoba	National University of Córdoba	UNC		X5000GYA	Córdoba	32				10		Anna	Fri Aug 31 00:00 CEST 2018
11.	12138	12138	Universität Buenos Aires	University of Buenos Aires	UBA	Viamonte 430/444	C1111AAI	Buenos Aires	32			www.uba.ar	10		Anna	Fri Aug 31 00:00 CEST 2018
12.	12044	12044	Nationale Universität Cuyo	National University of Cuyo	UNCuyo		CP M5502JMA	Ciudad de Mendoza	32				10		Anna	Fri Aug 31 00:00 CEST 2018
13.	8415	8415	Nationale Universität Rio Cuarto	National University of Rio Cuarto	UNRC		X5804BYA	Rio Cuarto	32				10		Anna	Fri Aug 31 00:00 CEST 2018
14.	12045	12045	Nationale Universität Rosario	National University of	UNR		S2000CGK	Rosaria	32				10		Anna	Fri Aug 31 00:00 CEST 2018

Abbildung 2: Preview des Dateimports

Facettieren, Sortieren, Filtern

Das Bearbeiten der Spalteinhalte erfolgt immer durch Auswahl der Funktionsmöglichkeiten nach Anklicken des „Drop-Down“-Symbols links von jeder Spaltenüberschrift

- Der erste Eintrag im Funktionsmenü ist immer „Facet“ - je nach Inhaltstyp der Spalte (Text, Zahl, Datumswert) sind unterschiedliche Facettierungsmöglichkeiten vorgesehen (vgl. Abbildung 3)

OpenRefine-Workshop

id_U6 U8_2018_ANNA3110 xlsx [Permalink](#)

		PARTNER_ID	Stammdatensatz	PARTNER_DE	located in the ad	located in the ad	wikidata_id	locate
★	1.	9635	9635			Tirana	Q1738772	
★	2.	10368	10158			Korçë	Q1637768	
★	3.	9704	9704			Tirana	Q821042	
★	4.	10158	10158			Korçë	Q1637768	
★	5.	11895	11895			Tirana	Q428520	
★	6.	3276						

Abbildung 3: Facettierung auswählen

- Nach auswählen der Facette erscheint im linken Teilframe ein Menü mit der entsprechend gruppierten Facettenauswahl (vgl. Abbildung 4)
 - Diese Auswahl kann nach Wert oder Anzahl der Vorkommnisse sortiert werden.
 - Auswahl durch Anklicken der Einträge.
 - Auswahl mehrere Einträge durch Klick auf „include“ neben dem Auswahltext
 - Für inverse Auswahl die entsprechenden Ausschlusskriterien auswählen und auf „invert“ klicken
 - Zurücksetzen der Facettierung erfolgt mit „reset“

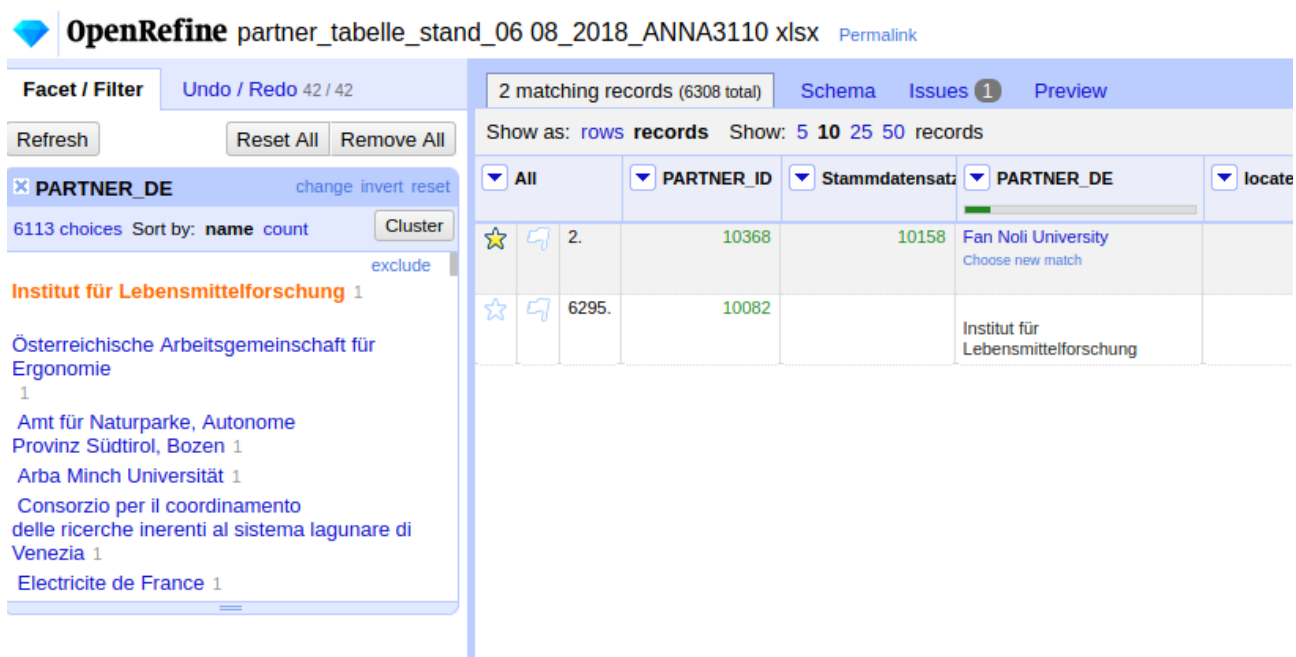


Abbildung 4: Facetten-Auswahlbox

Info: Facettierungseinstellungen werden nicht im Projekt gespeichert und sind nach jedem Neustart von OpenRefine erneut auszuwählen – dies kann insbesondere bei aufwändigen Facettierungen und Filterungen problematisch sein. Facettierungsschritte bei entsprechender Komplexität daher idealerweise dokumentieren!

- Um den aktuellen Bearbeitungsstand inklusiver Facettierung unmittelbar für das Projekt zu dokumentieren bietet OpenRefine die „Permalink-Funktionalität“ an – rechts neben dem Projektnamen in der Kopfzeile steht der Link „Permalink“ (vgl. Abbildung 5) diese Adresse beinhaltet auch sämtliche eingestellte Facettierungsoptionen und erlaubt somit eine Sicherung dieser Bearbeitungsstände.



Abbildung 5: Permalink für das OpenRefine-Projekt zur Sicherung von Facetten und Filteroptionen

- Der Textfilter (vgl. Abbildung 6) erlaubt das Suchen nach bestimmten Textelementen. Dabei ist auch die Eingabe Regulärer Ausdrücke möglich. (Checkbox regular expression)

OpenRefine-Workshop

OpenRefine partner_tabelle_stand_06 08_2018_ANNA3110.xlsx [Permalink](#)

Facet / Filter Undo / Redo 42 / 42 6308 records Schema Issues 2 Preview

Refresh Reset All Remove All

PARTNER_DE invert reset

case sensitive regular expression

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

	All	PARTNER_ID	Stammdatensatz	PARTNER_DE
★	1.	9635	9635	Tirana
★	2.	10368	10158	
★	3.	9704	9704	Tirana
★	4.	10158	10158	

Facet
Text filter
Edit cells
Edit column
Transpose
Sort...

Abbildung 6: Textfilter

- Im Bereich „Customized facets“ (vgl. Abbildung 1) befinden sich weitere granulare Facettierungsoptionen:
 - Word facet splittet die vorhandenen Strings in einzelne Wörter auf und lässt nach diese Facettieren.
 - Duplicates facet erlaubt ein relativ einfaches Überprüfen nach doppelten Werten
 - Facettierung anhand unterschiedlicher Null- oder Fehlerwerte.

atching records (6308 total) Schema Issues 2 Preview

rows records Show: 5 10 25 50 records

	PARTNER_ID	Stammdatensatz	PARTNER_DE	located in the ad	located in the ad	wikidata_id	located at street	Ring
	9635	9635				Q1738772		
	10368	10158				Q1637768		
	9704	9704				Q821042	112338	
	10158	10158				Q1637768		
	11895	11895						
	8702	8702						
0.	8700	8700	National University of Córdoba					
1.	12138	12138	University of Buenos Aires				28196	
2.	12044	12044	National University of Cuyo				28218	
3.	8415	8415	Nationale UNIVERSITÄT Río Cuarto					

Facet
Text facet
Text filter
Numeric facet
Edit cells
Timeline facet
Edit column
Scatterplot facet
Transpose
Custom text facet...
Sort...
Custom Numeric Facet...
View
Customized facets
Word facet
Reconcile
National University of La Plata (90)
Create new item
Search for match
Duplicates facet
Numeric log facet
1-bounded numeric log facet
Text length facet
Log of text length facet
Unicode char-code facet
Facet by error
Facet by null
Facet by empty string
Facet by blank (null or empty string)

Abbildung 7: Benutzerdefinierte Facetten

Bearbeiten von Zellen und Spalten

- Mit „Edit Cells“ und „Edit Columns“ stehen zwei Menüeinträge zur Verfügung, die das Bearbeiten der Inhalte bzw. das Anlegen neuer Spalten ermöglichen.

Info: In OpenRefine können standardmäßig neue Spalten angelegt werden, allerdings keine neuen Reihen bzw. Datensätze angelegt werden.

Edit Cells

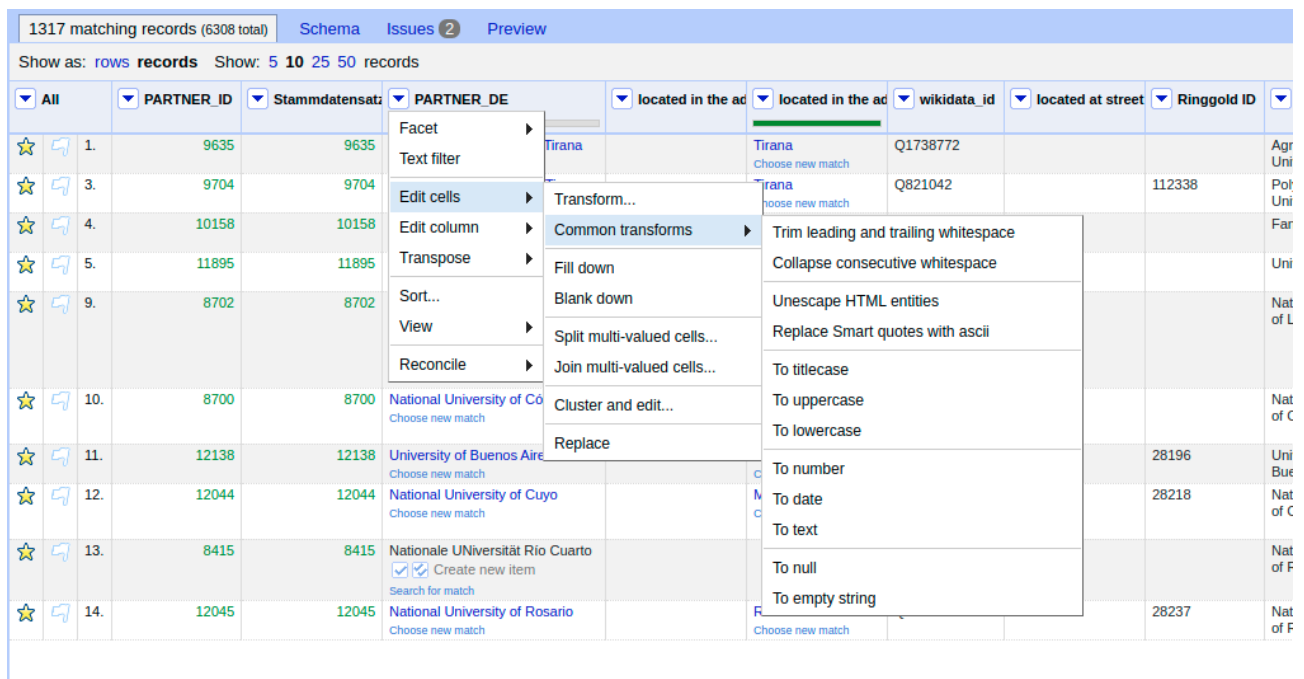


Abbildung 8: Edit Cells Menü

Der Menü-Eintrag „Edit Cells“ (vgl. Abbildung 8) erlaubt das Verändern und Bearbeiten der Zellinhalte der ausgewählten Spalte.

- Transform: Im Transform-Bereich (vgl. Abbildung 9) können mit Hilfe der Skriptsprachen GREL, Clojure oder Python Zellinhalte transformiert werden.
 - Das Vorschauenfenster zeigt dabei das zu erwartende Ergebnis im Vergleich zum aktuellen Zellenwert an.
 - Eine „History“ speichert bereits verwendete Transformationen ab und lässt diese erneut ausführen. Benutzte Transformationen können auch markiert werden und sind dann im Bereich „Starred“ abrufbar.

OpenRefine-Workshop

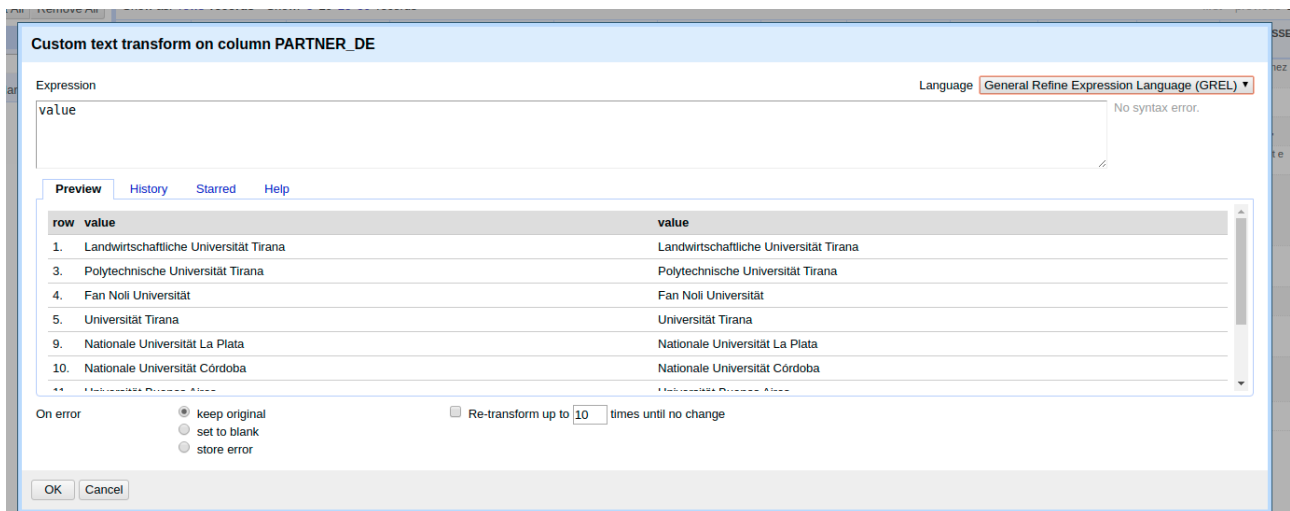


Abbildung 9: Transform Fenster

- „Commons Transforms“ bietet eine Reihe von häufig und standardisierten Benutzungsschritten an wie zB. Entfernung von Leerzeichen am Beginn und Ende der Zelle, Umwandlung in Groß-/Kleinbuchstaben, Typ-Umwandlung (Text/Zahl/Datum).
- Die Operation „Fill down“ füllt leere Zellen mit dem zuletzt vorhandenen Wert aus. „Blank down“ leert die Zellen.
- „Split multi-valued cells“ und „Join multi-valued cells“ erlaubt das Aufteilen (bzw. Zusammenführen) von Zellen mit mehrfachen Wertinhalten (die bspw. durch eindeutige Zeichen getrennt sind) auf. Werte werden dabei innerhalb der Spalte in neue Reihen aufgesplittet. Die Zusammengehörigkeit der Werte in den neuen Reihen zu dem ursprünglichen Datensatz bleibt im Record-Modus erhalten. (vgl. Row-/Record-Mode)
- „Cluster and edit“ - erlaubt auf sehr einfache und anschauliche Art und Weise das Bereinigung und Zusammenführen von ähnlichen Einträgen innerhalb einer Spalte. (vgl. Abbildung 11)
- „Replace“ erlaubt das einfache Suchen und Ersetzen von Strings innerhalb einer Spalte.

OpenRefine-Workshop

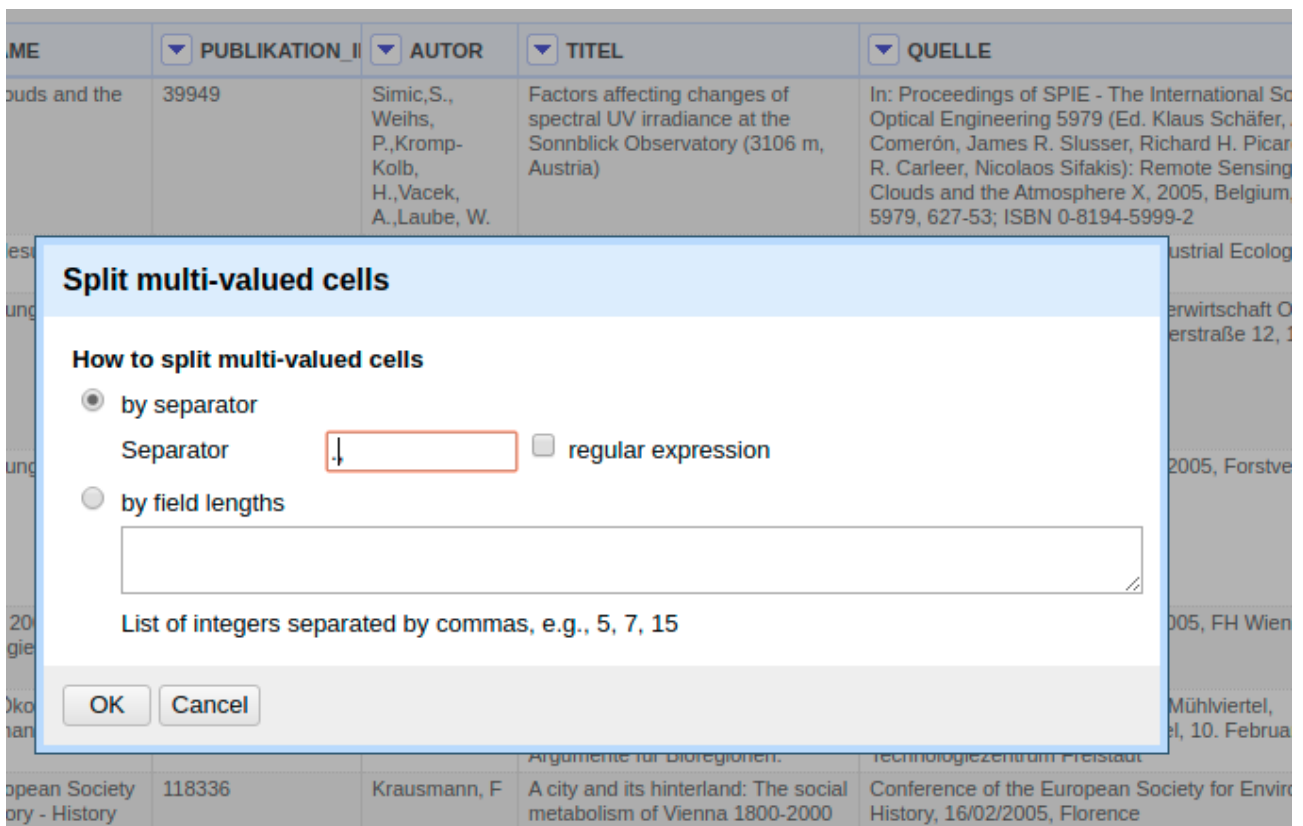


Abbildung 10: Split multi-valued cells

OpenRefine-Workshop

Cluster & Edit column "AUTOR"

This feature helps you find groups of different cell values that might be alternative representations of the same thing. For example, the two strings "New York" and "new york" are very likely to refer to the same concept and just have capitalization differences, and "Gödel" and "Godel" probably refer to the same person. [Find out more...](#)

Method: **key collision** Keying Function: **fingerprint** 387 clusters found

Cluster Size	Row Count	Values in Cluster	Merge?	New Cell Value
8	26	<ul style="list-style-type: none">Berghofer, E. (14 rows)Berghofer E. (3 rows)Berghofer, E. (3 rows)E. Berghofer (2 rows)BERGHOFER E (1 rows)BERGHOFER E. (1 rows)BERGHOFER, E (1 rows)Berghofer E (1 rows)	<input type="checkbox"/>	Berghofer, E.
6	26	<ul style="list-style-type: none">Kromp-Kolb, H. (14 rows)Kromp-Kolb, H. (4 rows)Kromp-Kolb, H. (3 rows)Kromp-Kolb, H (3 rows)H. Kromp-Kolb (1 rows)Kromp-Kolb H (1 rows)	<input type="checkbox"/>	Kromp-Kolb, H.
5	29	<ul style="list-style-type: none">Haltrich, D (12 rows)Haltrich, D. (12 rows)D. Haltrich (2 rows)Haltrich D (2 rows)Haltrich D. (1 rows)	<input type="checkbox"/>	Haltrich, D
5	24	<ul style="list-style-type: none">Schmutz, S (12 rows)Schmutz, S. (6 rows)	<input type="checkbox"/>	Schmutz, S

Select All Unselect All Export Clusters Merge Selected & Re-Cluster Merge Selected & Close Close

Abbildung 11: Cluster and edit

Edit Column

- „Split into several columns“ funktioniert analog wie „split multi-valued cells“ nur werden die Inhalte nicht in neue Reihen transformiert sondern in neue Spalten.
- „Join columns“ führt Spalten in eine zusammen.
- „Add column*“ erlaubt das Hinzufügen neuer Spalten basierend auf vorhandenen oder zu erstellenden Werten. Dabei öffnet sich das „Transform-Fenster“ (vgl. Abbildung 9) um mit den in OpenRefine integrierten Skriptsprachen GREL, Clojure, Python Bearbeitungen durchzuführen.
- „Rename“, „Remove“, „Move“ erlauben das Entfernen, Umbenennen oder Verschieben einzelner Spalten innerhalb des Projektes.

Row-/Record-Mode

Open Refine unterscheidet in der Bearbeitung und Darstellung der Datensätze zwischen dem Rows- und dem Record-Mode. Dies ermöglicht für bestimmte Werte innerhalb eines Datensatzes (zB für AutorInnen oder Schlagwörter bei einem bibliographischen Datensatz) mehrere Werte eindeutig als eigenständige Reihe zu erhalten, aber dennoch die Gesamtheit des Datensatzes zu wahren. In der

OpenRefine-Workshop

Menüleiste gibt es den Eintrag „Show as rows records“ der das Umschalten zwischen den beiden Darstellungs- bzw. Bearbeitungsmodi ermöglicht. (vgl. Abbildung 12 und Abbildung 13)

All	KONFERENZ_D	KONFERENZ_O	KONFERENZ_NAME	PUBLIKATION_ID	AUTOR	TITEL	QUELLE	WEITERE
☆	19.09.2005 - 21.09.2005	Bruges, Belgium	Remote Sensing of Clouds and the Atmosphere X	39949	Simic, S	Factors affecting changes of spectral UV irradiance at the Sonnblick Observatory (3106 m, Austria)	In: Proceedings of SPIE - The International Society for Optical Engineering 5979 (Ed. Klaus Schäfer, Adolfo T. Comerón, James R. Slusser, Richard H. Picard, Michel R. Carleer, Nicolaos Sifakis); Remote Sensing of Clouds and the Atmosphere X, 2005, Belgium, Vol. 5979, 627-53, ISBN 0-8194-5999-2	
☆					Weihls, P			
☆					Kromp-Kolb, H			
☆					Vacek, A			
☆					Laube, W.			

Abbildung 12: Datensatz im Record Mode

All	KONFERENZ_D	KONFERENZ_O	KONFERENZ_NAME	PUBLIKATION_ID	AUTOR	TITEL	QUELLE
☆	1. 19.09.2005 - 21.09.2005	Bruges, Belgium	Remote Sensing of Clouds and the Atmosphere X	39949	Simic, S	Factors affecting changes of spectral UV irradiance at the Sonnblick Observatory (3106 m, Austria)	In: Proceedings of SPIE - The International Society for Optical Engineering 5979 (Ed. James R. Slusser, Richard H. Picard, Michel R. Carleer, Nicolaos Sifakis); Remote Sensing of Clouds and the Atmosphere X, 2005, Belgium, Vol. 5979, 627-53, ISBN 0-8194-5999-2
☆	2.				Weihls, P		
☆	3.				Kromp-Kolb, H		
☆	4.				Vacek, A		
☆	5.				Laube, W.		

Abbildung 13: Datensatz im Raw-Mode

Reconciliation

Ein Schwerpunkt in der Nutzung und vor allem auch in der weiteren Entwicklung von OpenRefine liegt in der Verwendung als Upload-Tool für die Linked Open Data Plattform [wikidata](https://www.wikidata.org/). Um das Editieren der Plattform zu ermöglichen, ist es notwendig die entsprechenden verlinkten Entitäten zu den jeweils in den Zellen vorhandenen Werten zu ermitteln. Dieser Abgleichvorgang wird Reconciliation genannt. Dabei wird der vorhandene Zellenwert gegen die entfernte Datenbank abgeglichen, wobei auch Werte aus anderen Spalten zur näheren Beschreibung zusätzlich verwendet werden können. Die in der Datenbank gefundenen Ergebnisse werden hinsichtlich ihrer Konfidenz gewertet und entsprechend in einer Liste angegeben. OpenRefine-User können das Matching bestätigen oder entfernen.

Neben dem standardmäßig integrierten Reconciliationservice für Wikidata bieten immer mehr weitere Linked Data Plattformen die Möglichkeit Reconciliation in OpenRefine durchzuführen:

- Gemeinsame Normdatei (GND) via [lobid.org](https://lobid.org/gnd/api#openrefine): <https://lobid.org/gnd/api#openrefine>
- VIAF, ORCID, Open Library: <http://refine.codefork.com/>

Ablauf des Reconciliation-Vorganges

- Klick auf Bearbeitungs-Drop-Down in der ausgewählten Spalte: Reconcile – Starte reconciling.
- Das Reconciliation-Fenster (vgl. Abbildung 14) öffnet sich, Auswahl des zu nutzenden Services (zB. Wikidata)
- Bei Wikidata-Reconciliation wird in einem ersten Check bereits abgeprüft, um welche Klasse an Entitäten es sich womöglich handeln kann. Bei zutreffendem Ergebnis kann dies bereits als erste Einschränkung in der Menge der abzugleichenden Daten auf der Fremddatenbank verwendet werden. (zB für AutorInnen die Einschränkung auf Entitäten die als „human (Q5)“ in Wikidata verlinkt sind)
- Liegen beispielsweise weitere sinnvolle Werte im OpenRefine Projekt vor (bspw. Fremd-ID wie GND oder ORCID, sowie biographische Daten wie Geburtsdatum etc.), so können diese als weitere präzisierende Elemente eingesetzt werden. (Natürlich in der Hoffnung, dass diese verwendeten Werte auch in der Wikidata-Entität vorhanden sind)
- Start Reconciling.

Reconcile column "AUTOR"

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

- human (Q5)
- scholarly article (Q13442814)
- cellular component (Q5058355)
- family name (Q101352)

Also use relevant details from other columns:

Column	Include?	As Property
KONFERENZ_DATUM	<input type="checkbox"/>	<input type="text"/>
KONFERENZ_ORT	<input type="checkbox"/>	<input type="text"/>
KONFERENZ_NAME	<input type="checkbox"/>	<input type="text"/>
PUBLIKATION_ID	<input type="checkbox"/>	<input type="text"/>
TITEL	<input type="checkbox"/>	<input type="text"/>
QUELLE	<input type="checkbox"/>	<input type="text"/>
WEITERE_QUELLENANGABEN	<input type="checkbox"/>	<input type="text"/>
Datum Alt	<input type="checkbox"/>	<input type="text"/>
KONFERENZ_DATUM2	<input type="checkbox"/>	<input type="text"/>
KONFERENZ_Kontrolle	<input type="checkbox"/>	<input type="text"/>

Reconcile against type:
 Reconcile against no particular type
 Auto-match candidates with high confidence
 Maximum number of candidates to return

Abbildung 14: Wikidata Reconciliation

OpenRefine-Workshop

- Nach erfolgreichem Abgleich (vgl. Abbildung 15) werden bei „eindeutigem Match“, d.h. wenn genau ein potentielles Ergebnis erhalten wird, automatisch gematcht. Bei einer Liste an möglichen Matches wird diese als solche in der Zelle angezeigt, in der Reihenfolge der potentiellen Übereinstimmung mit Angabe eines Konfidenzwertes.
- Als letzter Eintrag in der Zelle wird „search for match“ angezeigt, wo manuell durch Variation des vorliegenden Suchbegriffes eine erneute einzelne Reconciliation durchgeführt werden kann.

The screenshot shows the OpenRefine interface with a table of 10 matching rows. The first row is highlighted, showing a match for the author 'Snezana B. Simić' with a confidence score of 100.00. The table columns are: KONFERENZ_ID, KONFERENZ_O, KONFERENZ_NAME, PUBLIKATION_ID, AUTOR, TITEL, and QUELLE. The 'AUTOR' column for the first row shows a list of suggested authors, including 'Snezana B. Simić' (100.00) and 'Weiths, P.' (86.00). The 'QUELLE' column contains the citation: 'In: Proceedings of SPIE - The International Society for Optical Engineering 5979 (Ed. Klaus Schäfer, Adolfo T. Comerón, James R. Slusser, Richard H. Picard, Michel R. Carlier, Nikolaos Sifakis); Remote Sensing of Clouds and the Atmosphere X, 2005, Belgium, Vol. 5979, 627-53; ISBN 0-8194-5999-2'.

Abbildung 15: Reconciliation Ergebnis

GREL

Mit der OpenRefine Expression Language (GREL) steht eine einfache Skriptsprache parallel zu Python oder Clojure zur Verfügung um automatisierte Transformationen in Spalten und Zellen durchführen zu können.

Variablen in OpenRefine

Um in OpenRefine mit den Skriptsprachen agieren zu können, ist es notwendig, die eingebauten Variablen zur Adressierung der unterschiedlichen Zellwerte zu kennen.

(<https://github.com/OpenRefine/OpenRefine/wiki/Variables>)

value	Die Variable value beinhaltet den Zellenwert der jeweils ausgewählten Spalte
cells[„Spaltenname“].value	Durch Angabe des Spaltennamens können Werte aus anderen Spalten als der ausgewählten verwendet werden. (zB. für den

OpenRefine-Workshop

	Einsatz in Stringverbindungen mit Abhängigkeit)
row	row liefert verschiedene reihenabhängige Werte
<ul style="list-style-type: none"> • row.index • row.cells • row.columnNames • row.starred • row.flagged • row.record 	<ul style="list-style-type: none"> • Indexnummer der aktuellen Reihenfolge • ident mit cells. • Reihen mit Star-Flag • Reihen mit Flag-Flag • ident mit record
record	record liefert verschiedene datensatzabhängige Werte
<ul style="list-style-type: none"> • record.index • record.cells • record.fromRowIndex • record.toRowIndex 	<ul style="list-style-type: none"> • Indexnummer des aktuellen Datensatzes • Zellwerte der in der Spalte enthaltenen rows • row.index der ersten Row im Record • row.index der letzten Row im Record
cell.value	ident mit value
cell.recon	Objekt mit Werten aus dem Reconciliation-Vorgang
<ul style="list-style-type: none"> • cell.recon.match (.id .name .type) 	Liefert ID, Name oder Typ des gematchten Recon-Objektes zurück
<ul style="list-style-type: none"> • cell.recon.matched 	True (boolean) wenn Zelle gematcht.
<ul style="list-style-type: none"> • cell.recon.judgment 	Liefert „matched“, „new“, oder „none“ (string)
<ul style="list-style-type: none"> • cell.recon.best 	wie cell.recon.match nur für das besten Recon-Ergebnis
<ul style="list-style-type: none"> • cell.recon.candidates (.id .name .type) 	Werte der ersten drei Recon-Ergebnisse als Objekt

Links

- OpenRefine (<https://www.openrefine.org>)
- GitHub (<https://github.com/OpenRefine>)
- Dokumentation (<https://github.com/OpenRefine/OpenRefine/wiki>)
- Stackoverflow (<https://stackoverflow.com/questions/tagged/openrefine>) – unter dem Tag [OpenRefine] sammeln sich nicht nur Fragen zu Problemen, sondern es werden vielmehr kleine Tutorials zu praktischen Situationen beim Arbeiten mit OpenRefine gesammelt und anschaulich erklärt.
- Wikidata:OpenRefine (<https://www.wikidata.org/wiki/Wikidata:Tools/OpenRefine>)
Informationen zur Nutzung von OpenRefine mit und für Wikidata.
- GND-Reconciliation-Service (<https://lobid.org/gnd/api#openrefine>)
- VIAF, LoC, ORCID Reconciliation-Services (<http://refine.codefork.com/>)