

CIDOC 2020 Getty Vocabulary Program

Jonathan Ward
Sr Editor, Getty Vocabulary Program
Getty Research Institute

Getty Vocabulary Program Update CIDOC 2020

- **Current and Future Contributions / Priorities**
- **Alma Refine for catalogers**
- **OpenRefine reconciliation service advanced techniques**

Pt 1

Current and Future Contributions / Priorities

Getty Vocabulary Program Update CIDOC 2020

- **Current and Future Contributions / Priorities**
 - **Backlogged data loads**
 - **Outside consultant (Ontotext)**
 - **Systematic workflow for regular contributions**
 - **GVP overall priorities for the future**

BWR

Built Works Registry

Global IDs for Architecture
and the Built Environment



Cultural Objects
Name Authority
(CONA)®



BWR3781470

BWR ID

Wilbert Snow School Addition

BWR Name

Jeter Cook & Jepson Architects, United States, firm | Ashley, Warren Henry (1909-1968), American, architect

Creator

1994-1999

Date

Middletown, Connecticut, United States

Location

41.55

Latitude

-72.65

Longitude

schools

Work Type

BWR Data points:

- Unique ID
- Name/Title
- Creator Name(s)
- Date of work
- Location of work
- Lat/Long
- Work Type



Cultural Objects
Name Authority
(CONA)®



BWR3781470

BWR ID

Wilbert Snow School Addition

BWR Name

Jeter Cook & Jepson Architects, United States, firm | Ashley, Warren Henry (1909-1968), American, architect
Creator

1994-1999

Date

Middletown, Connecticut, United States

Location

41.55

Latitude

-72.65

Longitude

schools

Work Type

10 matching rows (4999 total)

Extensions

Show as: rows records

Show: 5 10 25 50 rows

« first < previous 1 - 1

	D[34820]	BWR Name[3482]	BWR Variant Na	Larger Entity or	Creator[34824]	ULAN Creator[34824]	ULAN Creator[34824]	ULAN Creator[34824]	ULAN Creator[34824]	ULAN Creator[34824]	ULAN Creator[34824]
48	Whalen Junior High School				Ashley, Warren	Ashley, Warren Henry Choose new match	Ashley, Warren Henry (American architect, 1909-1968)				
54	East Granby Middle and High School				Ashley, Warren	Ashley, Warren Henry Choose new match					
60	Seymour Elementary School				Ashley, Warren	Ashley, Warren Henry Choose new match					
80	Wilbert Snow Elementary School				Ashley, Warren	Ashley, Warren Henry Choose new match					
72	Center Episcopal Church				Ashley, Warren	Ashley, Warren Henry Choose new match					
94	Saybrook High School				Ashley, Warren	Ashley, Warren Henry Choose new match					
99	Waterford High School				Ashley, Warren	Ashley, Warren Henry Choose new match					
17	Fletcher Judson Elementary School				Ashley, Warren	Ashley, Warren Henry Choose new match					
29	Warren Ashley Office				Ashley, Warren	Ashley, Warren Henry Choose new match					





(4999 total)									EXCLUSIONS	▼ Individual			
Show: 5 10 25 50 rows									« first	< previous	1 - 1	next >	last »
Subsite and/or S	<input type="button" value="▼"/> City[34832]	<input type="button" value="▼"/> TGN City[34832]	<input type="button" value="▼"/> State or Provinc	<input type="button" value="▼"/> Country[34834]	<input type="button" value="▼"/> Postal Code[348	<input type="button" value="▼"/> Notes[34836]	<input type="button" value="▼"/> TGN[34837]	<input type="button" value="▼"/> Latitude[34838]					
	Middletown	Middletown	Connecticut	United States			7007159	41.55					

1 matching rows (4999 total)

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

site and/or S	City[34832]	TGN City[34832]	State or Provinc	Country[34834]	Posta
	Middletown		Connecticut	United States	

Facet ▶
Text filter
Edit cells ▶
Edit column ▶
Transpose ▶
Sort...
View ▶
Reconcile ▶ Start reconciling...
Facets Reconcile text in this column w
Actions ▶
Copy reconciliation data...
Use values as identifiers



Reconcile column "TGN City[34832]"

» Access Service API

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

- ULAN search
/ulan
- TGN search
/tgn
- AAT search
/aat
- Search all Vocab
/all

Also use relevant details from other columns:

Latest Date[34826]	<input type="checkbox"/>	
Date Completed[34827]	<input type="checkbox"/>	
Street Address or Intersection[34828]	<input type="checkbox"/>	
Neighborhood[34829]	<input type="checkbox"/>	
Borough[34830]	<input type="checkbox"/>	
Subsite and/or Site[34831]	<input type="checkbox"/>	
City[34832]	<input type="checkbox"/>	
State or Province[34833]	<input checked="" type="checkbox"/>	broaderExt
Country[34834]	<input checked="" type="checkbox"/>	broaderExt
Postal Code[34835]	<input type="checkbox"/>	
Notes[34836]	<input type="checkbox"/>	
TGN[34837]	<input type="checkbox"/>	

Reconcile against type:

Reconcile against no particular type

Auto-match candidates with high confidence

Maximum number of candidates to return ↑

1 matching rows (4999 total)

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

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

site and/or S	City[34832]	TGN City[34832]	State or Province	Country[34834]	Postal Code[348]	Notes
	Middletown	Middletown <input checked="" type="checkbox"/> Middletown (37) <input checked="" type="checkbox"/> Portland (21) <input checked="" type="checkbox"/> Create new item	edit	Connecticut	United States	

[Match this Cell](#) [Match All Identical Cells](#) [Cancel](#)

Middletown
(inhabited places)
Geographic Hierarchy:

- World
- ...
- Connecticut
- Middlesex
- **Middletown (inhabited places)**

ULAN projects:

- Belvedere Museum, Vienna – 22,000 artist name records
- Tobunken, Tokyo – 1,750 artist name records (Japanese)
- RKD – entire artist name file, with associative metadata and relationships
- MoMA – additional 600 records related to LGBTQ community

AAT:

- Chinese, Dutch translations continue; CHIN French terms
- Israel Museum Jerusalem – 2,000 Hebrew terms, objects and cultures
- Portuguese terms: USP and graduate intern Camila da Silva

 **views (visual works)** (<visual works by form of image>, <visual works by form>, ... Visual and Verbal Communication (hierarchy name))

Note: Depictions of places or things seen from a distance, with a range of sight distinct from other information forms such as diagrams, working drawings, or verbal descriptions. Generally reserved for two-dimensional media such as prints, drawings, paintings, photographs, or bas-reliefs.

Terms:

views (visual works) ([preferred](#), C, U, LC, English-P, D, U, PN)

view (visual work) (C, U, English, AD, U, SN)

風景畫 (視覺作品) (C, U, Chinese (traditional)-P, D, U, U)

fēng jǐng huà (C, U, Chinese (transliterated Hanyu Pinyin)-P, UF, U, U)

feng jing hua (C, U, Chinese (transliterated Pinyin without tones)-P, UF, U, U)

feng ching hua (C, U, Chinese (transliterated Wade-Giles)-P, UF, U, U)

gezichten (beeldmateriaal) (C, U, Dutch-P, D, U, U)

gezicht (C, U, Dutch, AD, U, U)

vistas (C, U, Spanish-P, D, U, PN)

(obras visuais) (Portuguese-P, D, U, PN)

vista (C, U, Spanish, AD, U, SN)

(obra visual) (Portuguese, AD, U, SN)

Additional Notes:

Chinese (traditional) 描繪各處地方風情或各種物品的圖像，一般是以基本平面媒材繪製而成，例如：版畫、素描、油畫、照片，或者是淺浮雕 (bas-reliefs) 等。此詞彙一般是用以指稱描繪特定視野範圍或景象的圖像，不同於圖表及施工圖等其他以呈現資訊為主的圖像或是文字描述。

Dutch Verwijst naar voorstellingen van plaatsen of dingen, doorgaans op vrijwel tweedimensionale media zoals prenten, tekeningen, schilderingen, foto's of bas-reliefs. De term wordt meestal gebruikt om een voorstelling gemaakt vanuit een bepaald gezichtsveld te onderscheiden van andere informatievormen zoals diagrammen, werktekeningen of verbale beschrijvingen.

Portuguese Representações de lugares ou coisas vistas à distância e distintas de outras formas de informação tais como diagramas, esboços ou descrições verbais.

Spanish Úsese para descripciones, incluyendo fotografías, de lugares o cosas, especialmente para distinguirlas de otra información tales como diagramas, obras dibujadas o descripciones verbales.

Future Priorities and Scope for Contributions

- Contributions based on our mission statement and steering committee
- Criteria for contributions will be published
- More assistance for contributors
 - OpenRefine
 - Workable spreadsheets
 - New contribution webform

Pt 2

Alma Refine

Alma Refine

Refine Alma BIB records using OpenRefine services

This app can be used to refine bibliographic records with services that adhere to the OpenRefine standard. References to the Getty vocabularies, WikiData, and GeoNames are included with the app.

More information



By [Josh Weisman](#)

View the app on: 

Version: v0.2.0 | [Help](#) | Last updated on 20 Nov 2020

This application allows: modals

This application uses external websites:

- <https://services.getty.edu>
- <https://wdreconcile.toolforge.org>
- <https://api.exldevnetwork.net>
- <https://lobid.org>
- <https://conciliator.herokuapp.com>

Facets

Material Type

Book (87)

Mixed material (1)

Resource Type

Book - Electronic (1)

Book - Physical (86)

Undefined (1)

Language

English (75)

French (13)

Italian (1)

Japanese (1)

Publication Year

1879 - 1973 (22)

1974 - 1991 (22)

1992 - 2006 (21)

2007 - 2018 (23)

< All Titles (1 - 50 of 88) | Mary Cassatt

Save Query

Institution

Community

Edit Record Holdings ...

Sort by : Rank Secondary Sort by : Rank Expand: Physical

1 Mearī Kasatto ten = Mary Cassatt retrospective / henshū Yokohama Bijutsukan, Kyōto Kokuritsu Kindai Bijutsukan, NHK, NHK Puromōshon.

Book (Book - Physical) text; unmediated; volume By Cassatt, Mary, ([Tōkyō] : NHK : 2016.)

マエリー・カサット展 = Mary Cassatt retrospective / 編集 横浜美術館, 京都国立近代美術館, NHK, NHKプロモーション.

Subject: Cassatt, Mary / 1844-1926 Cassatt, Mary / 1844-1926 / Exhibitions Cassatt, Mary / 1844-1926 / Criticism and interpretation and others

Creation Date: 11/20/2018 18:51:59 PST

Modification Date: 03/15/2020 03:27:56 PDT

Language: Japanese

Record number: (OCoLC)ocn954491862

Orders: 1

MMS ID: 9933859107101551

Physical (1) Electronic Digital Other details

1 Holdings | 1 of 1 items are available

Library	Location	Call Number	Accession Number	Item Availability	Related Record
GC	M	N6537.C35 A4 2016	-	1 of 1 Available	-

[Holdings](#) | [Items](#)

Edit Record Holdings ...

2 Innovative impressions : prints by Cassatt, Degas, and Pissarro / Sarah Lees, Richard R. Brettell.

Book (Book - Physical) still image; unmediated; volume By Lees, Sarah, (Tulsa, Oklahoma : Philbrook Museum of Art, ©2018)

Subject: Cassatt, Mary--1844-1926--Exhibitions. Degas, Edgar--1834-1917--Exhibitions. Pissarro, Camille,--1830-1903--Exhibitions. and others

Creation Date: 01/16/2019 11:38:57 PST

Modification Date: 05/26/2020 01:42:47 PDT

Language: English

ISBN: 9783777429786 and others

MMS ID: 9933870011001551

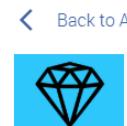
Record number: (OCoLC)1019846181

Physical (1) Electronic Digital Other details

1 Holdings | 1 of 1 items are available

Library	Location	Call Number	Accession Number	Item Availability	Related Record
GC	7G	3387-001-100	-	1 of 1 Available	-

Begin Refining



Alma Refine

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

 Set Displayed Records Select Records

Refine the 50 records on this page

Getty Vocabularies

Begin Refining



All titles ▾ Subjects ▾ Mary Cassatt

General Collections - Annex Circ

Alma Sandbox

Back to App List

Alma Refine

9924975390001551 Rare Japanese prints.

9925351190001551 Edgar Degas letters, 100\$a: Degas, Edgar,

9925691980001551 Painting in a man's world : four stories about Berthe Morisot, Mary Cassatt, Eva Gonzalès, Marie Bracqueond / 100\$a: Broeckhoven, Diane,

9925721610001551 Women impressionists /

9925884180001551 Cassatt : a retrospective / 100\$a: Cassatt, Mary,

9925897390001551 Painted faces : Mary Cassatt, Alice Neel, Karen Kilimnik /

9926326850001551 Eve's daughter/Modern woman : a mural by Mary Cassatt / 100\$a: Webster, Sally.

9926356460001551 Leaders of American impressionism: Mary Cassatt, Childe Hassam, John H. Twatchman [and] J. Alden Weir. 110\$a: Brooklyn Museum.

9927472790001551 Mary Cassatt's [sic] Children Playing With A Dog. 110\$a: Christie, Manson & Woods International Inc.

655\$a: Auction catalogs.

auction catalogs

None

Degas, Edgar

Degas, Andre

Jené, Edgar

Baran, Edgar

Hecht, Edgar

Important Impressionist artist; unlike other Impressionists, he emphasized composition and drawing, and he usually did not paint outdoors. Degas was primarily concerned with depicting movement, from horses to women in various activities, dressing, bathing, and as cabaret performers.



Getty

CIDOC - December 9, 2020

	050	_4 a N6537.C35 b A4 1996
	090	__ a N6537.C35 b A4 1996
	100	1_ a Cassatt, Mary d 1844-1926. 0 http://vocab.getty.edu/ulan/500012368
	245	10 a Cassatt : b a retrospective / c edited by Nancy Mowll Mathews.
	260	__ a [S.l.] : b H.L. Levin Associates, c c1996.
	300	__ a 357 p. : b ill. (some col.) ; c 34 cm.
	504	__ a Includes bibliographical references (p. 11-12) and index.



Pt 3

Advanced OpenRefine Techniques

Getty Vocabulary Program OpenRefine service

Open source: <http://openrefine.org>

Runs in a browser; available in multiple languages.

Data reconciliation / alignment / cleanup

For large and small-scale alignment with Getty Vocabulary data.

Uses GREL language (General Refine Expression Language)

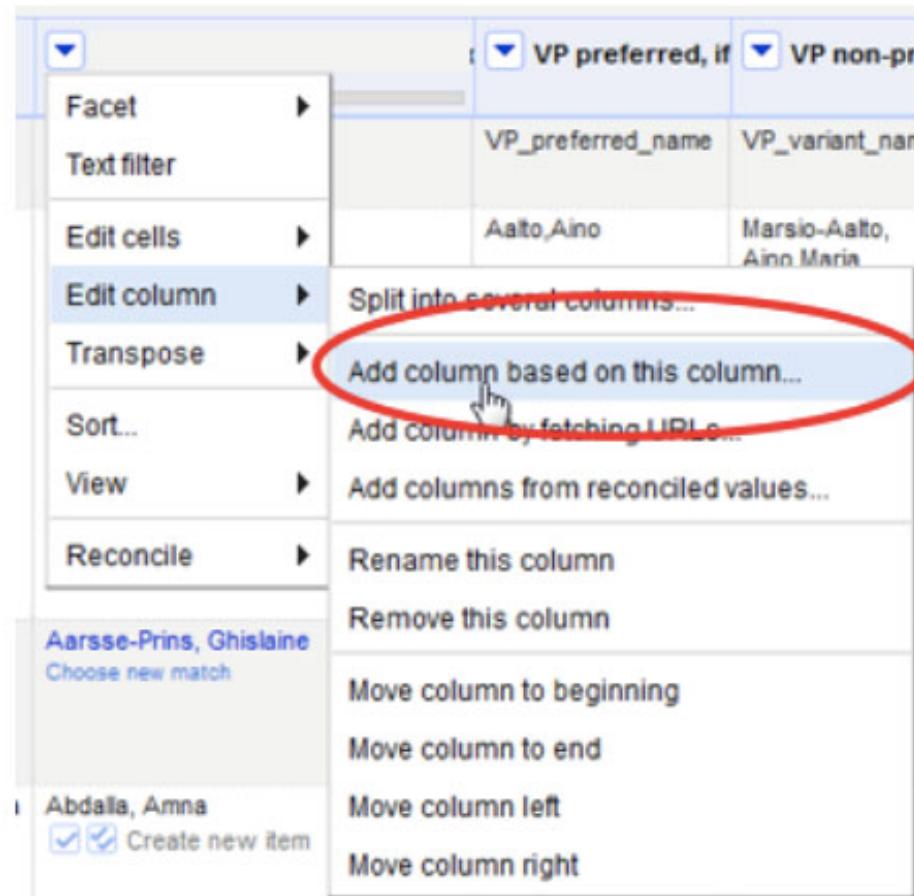
<https://github.com/OpenRefine/OpenRefine/wiki/GREL-Functions>

For specific information and tutorial:

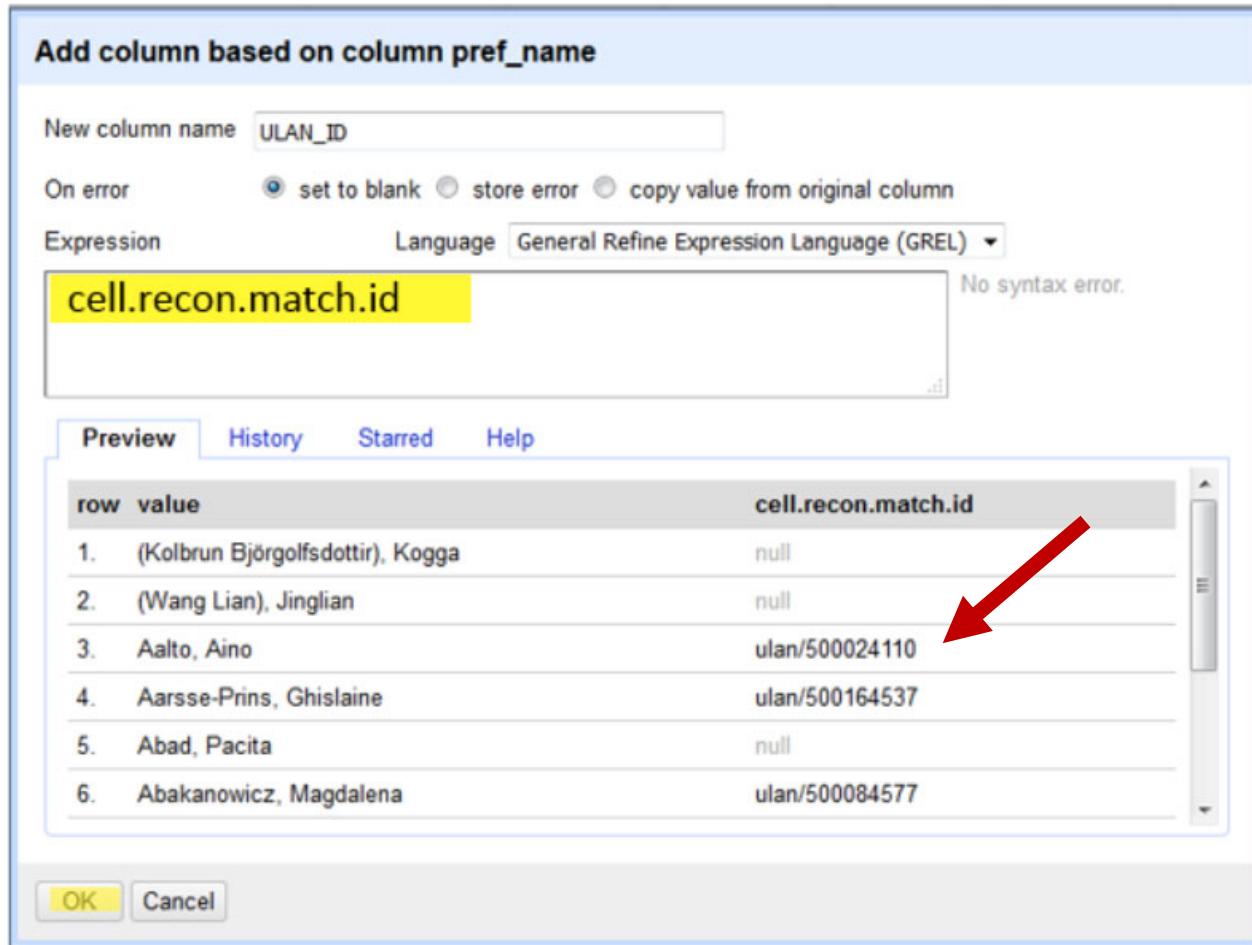
<http://www.getty.edu/research/tools/vocabularies/obtain/openrefine.html>

Getting Vocabulary IDs for Matches

Add a new column based on the “matched” column:



Name your new column, and apply a GREL expression:



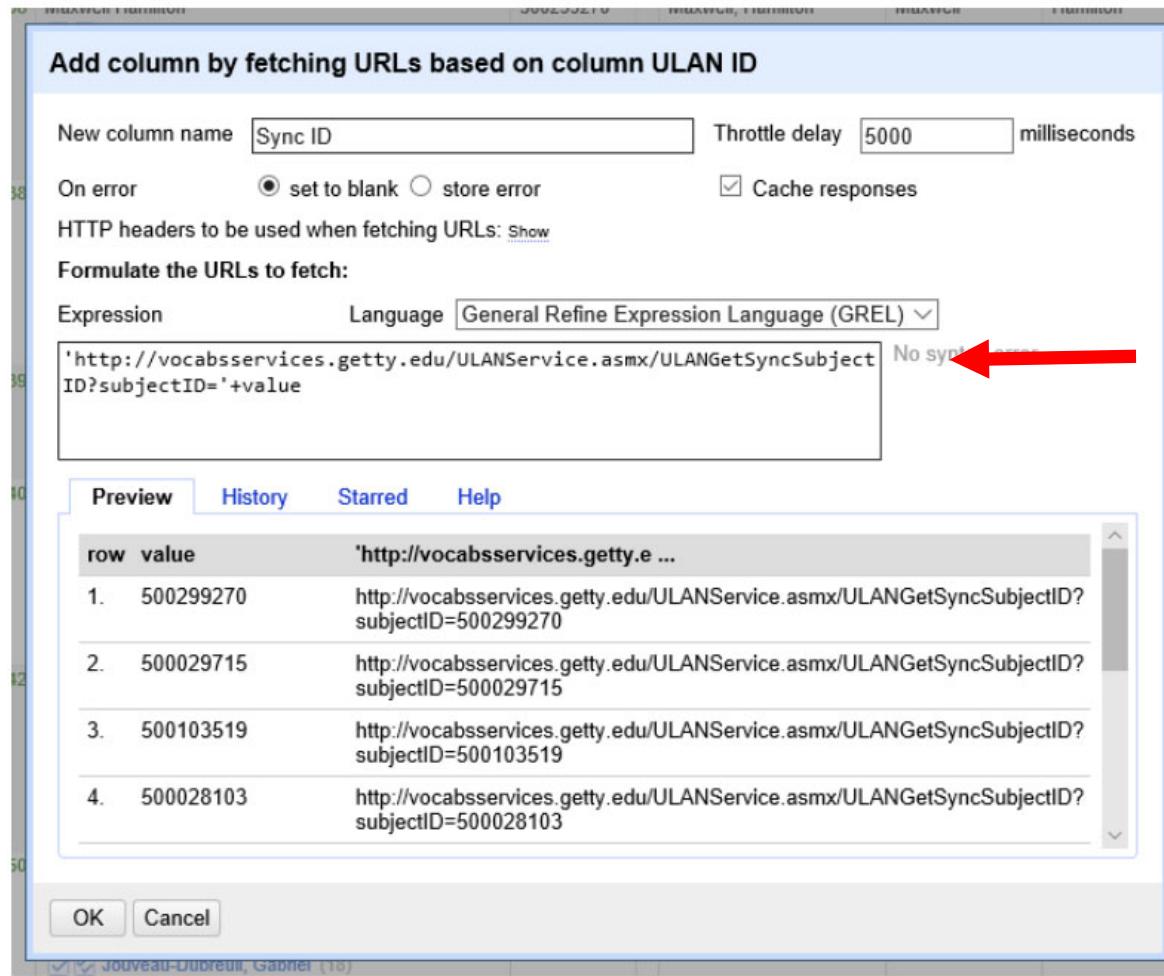
Synching Getty Vocabulary IDs Via XML Web Services using OpenRefine

Add a new column by “fetching URLs”

DISPLAYNAME	ULAN ID	Pref Term	FIRSTNAME	LASTNAME	BEGINDATE	ENDDATE
Maxwell Hamilton		Facet Text filter	H. Hamilton	Maxwell	Hamilton	0
<input checked="" type="checkbox"/> Maxwell, Hamilton (24) <input checked="" type="checkbox"/> Hamilton, Maxwell, Inc. (23) <input checked="" type="checkbox"/> Fry, Edwin Maxwell (14) <input checked="" type="checkbox"/> Bates, Maxwell (14) <input checked="" type="checkbox"/> Maxwell, John (14) <input checked="" type="checkbox"/> Create new item		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			
Charles Aubry					1811	
<input checked="" type="checkbox"/> Aubry, Charles (34) <input checked="" type="checkbox"/> Aubry, Charles (33) <input checked="" type="checkbox"/> Aubry, Emile (18) <input checked="" type="checkbox"/> Aubry, Louis (18) <input checked="" type="checkbox"/> Aubry, René Marcel (17) <input checked="" type="checkbox"/> Create new item						
Fédèle Azari	500103519	Azari,			1895	
František Vobecký	500028103	Vobec			1902	
<input checked="" type="checkbox"/> Vobecký, František (43) <input checked="" type="checkbox"/> Schmoranz, František (17) <input checked="" type="checkbox"/> Fišer, František (17) <input checked="" type="checkbox"/> Kadlec, František (17) <input checked="" type="checkbox"/> Kysela, František (17) <input checked="" type="checkbox"/> Create new item						

Define column name and input GREL expression:

<http://vocabsservices.getty.edu/ULANService.asmx/ULANGetSyncSubjectID?subjectID='+value>



Resulting XML



ID	DISPLAYNAME	ULAN ID	Sync ID
	Maxwell Hamilton	edit	500299270
	<input checked="" type="checkbox"/> Maxwell, Hamilton (24)		<?xml version="1.0" encoding="utf-8"?><Vocabulary xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="http://vocabsservices.getty.edu/Schemas/ULAN/ULANGetSyncSubjectId.xsd"><Current_Subject_ID>500299270</Current_Subject_ID></Vocabulary>
	<input checked="" type="checkbox"/> Hamilton, Maxwell, Inc. (23)		
	<input checked="" type="checkbox"/> Fry, Edwin Maxwell (14)		
	<input checked="" type="checkbox"/> Bates, Maxwell (14)		
	<input checked="" type="checkbox"/> Maxwell, John (14)		
	<input checked="" type="checkbox"/> Create new item		
	Charles Aubry		500029715
	<input checked="" type="checkbox"/> Aubry, Charles (34)		<?xml version="1.0" encoding="utf-8"?><Vocabulary xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="http://vocabsservices.getty.edu/Schemas/ULAN/ULANGetSyncSubjectId.xsd"><Current_Subject_ID>500029715</Current_Subject_ID></Vocabulary>
	<input checked="" type="checkbox"/> Aubry, Charles (33)		
	<input checked="" type="checkbox"/> Aubry, Emile (18)		
	<input checked="" type="checkbox"/> Aubry, Louis (18)		
	<input checked="" type="checkbox"/> Aubry, René Marcel (17)		
	<input checked="" type="checkbox"/> Create new item		
	Fédèle Azari		500103519
	<input checked="" type="checkbox"/> Azari, Fedele (19)		<?xml version="1.0" encoding="utf-8"?><Vocabulary xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="http://vocabsservices.getty.edu/Schemas/ULAN/ULANGetSyncSubjectId.xsd"><Current_Subject_ID>500103519</Current_Subject_ID></Vocabulary>
	<input checked="" type="checkbox"/> Āzārī Isfarāyīnī, Ḥamzah ibn 'Alī Malik (5)		
	<input checked="" type="checkbox"/> Create new item		
	František Vobecký		500028103
			<?xml version="1.0" encoding="utf-8"?>

Parse XML using Transform

ID	DISPLAYNAME	ULAN ID	Sync ID
	Maxwell Hamilton	500299270	<p>ding="utf-8"?> http://www.w3.org/2001/XMLSchema-instance" al location="http://vocabsservices.getty.edu/Schemas/ULAN/ULANGetSyncSubjectId.xsd"></p>
	<input checked="" type="checkbox"/> Maxwell, Hamilton (24)		<ul style="list-style-type: none">FacetText filterEdit cellsEdit columnTransposeSort...ViewReconcile
	<input checked="" type="checkbox"/> Hamilton, Maxwell, Inc. (23)		<ul style="list-style-type: none">Transform...Common transformsFill downBlank downSplit multi-valued cells...Join multi-valued cells...Cluster and edit...Replace
	<input checked="" type="checkbox"/> Fry, Edwin Maxwell (14)		
	<input checked="" type="checkbox"/> Bates, Maxwell (14)		
	<input checked="" type="checkbox"/> Maxwell, John (14)		
	<input checked="" type="checkbox"/> Create new item		
	Charles Aubry	500029715	<p><Vocabulary xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="http://vocabsservices.getty.edu/Schemas/ULAN/ULANGetSyncSubjectId.xsd"></p> <p><Current_Subject_ID>500029715</Current_Subject_ID></p> <p></Vocabulary></p>
	<input checked="" type="checkbox"/> Aubry, Charles (34)		
	<input checked="" type="checkbox"/> Aubry, Charles (33)		
	<input checked="" type="checkbox"/> Aubry, Emile (18)		
	<input checked="" type="checkbox"/> Aubry, Louis (18)		
	<input checked="" type="checkbox"/> Aubry, René Marcel (17)		
	<input checked="" type="checkbox"/> Create new item		
	Fédèle Azari	500103519	<p><?xml version="1.0" encoding="utf-8"?></p> <p><Vocabulary xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="http://vocabsservices.getty.edu/Schemas/ULAN/ULANGetSyncSubjectId.xsd"></p> <p><Current_Subject_ID>500103519</Current_Subject_ID></p> <p></Vocabulary></p>
	<input checked="" type="checkbox"/> Azari, Fedele (19)		
	<input checked="" type="checkbox"/> Āzārī Isfarāyīnī, Hamzah ibn 'Alī Malik (5)		
	<input checked="" type="checkbox"/> Create new item		
	František Vobecký	500028103	<p><?xml version="1.0" encoding="utf-8"?></p> <p><Vocabulary xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"</p>
	<input checked="" type="checkbox"/> Vobecký František		

Transform using GREL expression:

`value.parseXML().select('Vocabulary')[0].select('Current_Subject_ID')[0].ownText()`

Custom text transform on column Sync ID

Expression

`value.parseXml().select('Vocabulary')[0].select('Current_Subject_ID')[0].ownText()`

No syntax error.

Preview History Starred Help

row	value	value.parseXml().select("Vocab ...")
1.	<?xml version="1.0" encoding="utf-8"?> <Vocabulary xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="http://vocabsservices.getty.edu/Schemas/ULAN/ULANGetSyncSubjectId.xsd"> <Current_Subject_ID>500299270</Current_Subject_ID> </Vocabulary>	500299270
2.	<?xml version="1.0" encoding="utf-8"?> <Vocabulary xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="http://vocabsservices.getty.edu/Schemas/ULAN/ULANGetSyncSubjectId.xsd"> <Current_Subject_ID>500029715</Current_Subject_ID> </Vocabulary>	500029715

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

OK Cancel

Sync Column Becomes Current ULAN ID

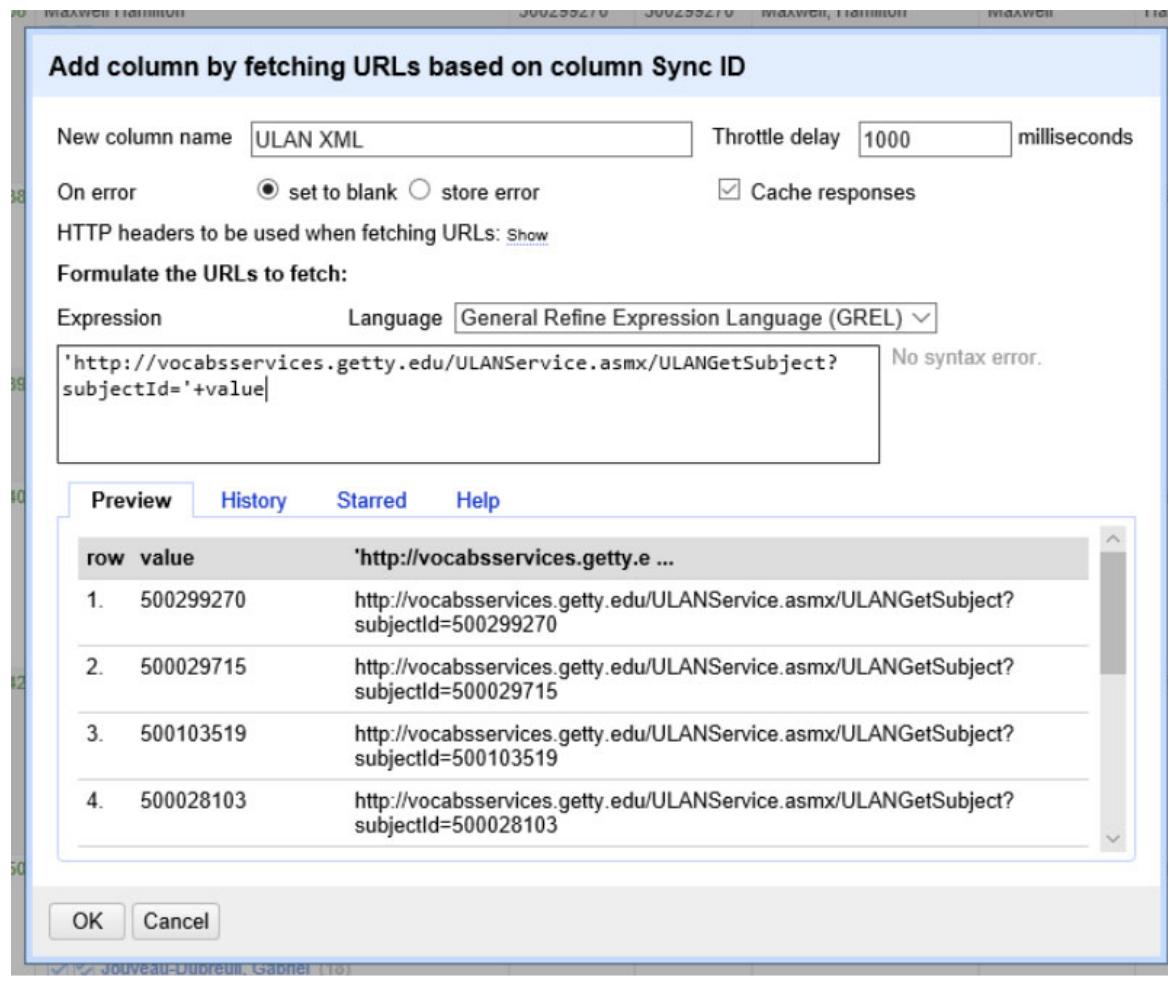


ID	DISPLAYNAME	ULAN ID	Sync ID
	Maxwell Hamilton <ul style="list-style-type: none"><input checked="" type="checkbox"/> Maxwell, Hamilton (24)<input checked="" type="checkbox"/> Hamilton, Maxwell, Inc. (23)<input checked="" type="checkbox"/> Fry, Edwin Maxwell (14)<input checked="" type="checkbox"/> Bates, Maxwell (14)<input checked="" type="checkbox"/> Maxwell, John (14)<input checked="" type="checkbox"/> Create new item	500299270	500299270
	Charles Aubry <ul style="list-style-type: none"><input checked="" type="checkbox"/> Aubry, Charles (34)<input checked="" type="checkbox"/> Aubry, Charles (33)<input checked="" type="checkbox"/> Aubry, Emile (18)<input checked="" type="checkbox"/> Aubry, Louis (18)<input checked="" type="checkbox"/> Aubry, René Marcel (17)<input checked="" type="checkbox"/> Create new item	500029715	500029715
	Fédèle Azari <ul style="list-style-type: none"><input checked="" type="checkbox"/> Azari, Fedele (19)<input checked="" type="checkbox"/> Āzārī Isfarāyīnī, Ḥamzah ibn 'Alī Malik (5)<input checked="" type="checkbox"/> Create new item	500103519	500103519
	František Vobecký <ul style="list-style-type: none"><input checked="" type="checkbox"/> Vobecký, František (43)<input checked="" type="checkbox"/> Schmoranz, František (17)<input checked="" type="checkbox"/> Fišer, František (17)<input checked="" type="checkbox"/> Kadlec, František (17)<input checked="" type="checkbox"/> Kysela, František (17)<input checked="" type="checkbox"/> Create new item	500028103	500028103

OpenRefine data enrichment example

Using full XML from Web Service

Get full XML by adding column based on synced ULAN ID, and input formula:
<http://vocabsservices.getty.edu/ULANService.asmx/ULANGetSubject?subjectId='+value>



Full XML in OpenRefine cell

ID	DISPLAYNAME	ULAN ID	Sync ID	ULAN XML
	Maxwell Hamilton <input checked="" type="checkbox"/> Maxwell, Hamilton (24) <input checked="" type="checkbox"/> Hamilton, Maxwell, Inc. (23) <input checked="" type="checkbox"/> Fry, Edwin Maxwell (14) <input checked="" type="checkbox"/> Bates, Maxwell (14) <input checked="" type="checkbox"/> Maxwell, John (14) <input checked="" type="checkbox"/> Create new item	500299270	500299270	<?xml version="1.0" encoding="utf-8"?> <Vocabulary xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="http://vocabsservices.getty.edu/Schemas/ULAN/ULANGetSubject.xsd"> <Subject Subject_ID="500299270"> <Parent_Relationships> <Preferred_Parent> <Parent_Subject_ID>500000002</Parent_Subject_ID> <Relationship_Type>Parent/Child</Relationship_Type> <Historic_Flag>Undetermined</Historic_Flag> <Parent_String>Persons, Artists [500000002]</Parent_String> <Hier_Rel_Type>Instance-BTI</Hier_Rel_Type> </Preferred_Parent> </Parent_Relationships> <Descriptive_Notes> <Descriptive_Note> <Note_Text>Comment on works: landscape</Note_Text> <Note_Language>English</Note_Language> <Note_Contributors> <Note_Contributor> <Contributor_id>2500000011/PROV</Contributor_id> </Note_Contributor> </Note_Contributors> <Note_Sources> <Note_Source> <Source> <Source_ID>2100043316/Provenance Index Databases, Authority file (1985-)</Source_ID> </Source> </Note_Source> </Note_Sources> </Descriptive_Note> </Descriptive_Notes> <Record_Type>Person</Record_Type> <Merged_Status>Merged</Merged_Status> <Terms> <Preferred_Term> <Term_Text>Maxwell, Hamilton</Term_Text> <Display_Name>Index</Display_Name> <Historic_Flag>N/A</Historic_Flag> <Vernacular>Undetermined</Vernacular> <Term_ID>1500668205</Term_ID> <Term_Languages> <Term_Language>

Parse Preferred Biography from OpenRefine cell using:

```
value.parseXml().select('Vocabulary')[0].select('Subject')[0].select('Biographies')[0].select('Preferred_Biography')[0].select('Biography_Text')[0].ownText()
```

The screenshot shows the 'Add column based on column ULAN XML' dialog in OpenRefine. The 'New column name' is 'Pref Bio'. The 'On error' option is set to 'set to blank'. The 'Expression' language is 'General Refine Expression Language (GREL)'. The expression itself is:

```
value.parseXml().select('Vocabulary')[0].select('Subject')[0].select('Biographies')[0].select('Preferred_Biography')[0].select('Biography_Text')[0].ownText()
```

The 'Preview' tab is selected, showing the following table:

row	value	
1.	<?xml version="1.0" encoding="utf-8"?> <Vocabulary xmlns:xsi="http://www.w3.org/2001/XMLSchema- instance" xsi:noNamespaceSchemaLocation="http://vocab: <Subject Subject_ID="500299270"> <Parent_Relationships> <Preferred_Parent> <Parent_Subject_ID>500000002</Parent_Subject	Scottish painter and architect, 1830-1923

At the bottom of the dialog are 'OK' and 'Cancel' buttons.

Preferred Biography in OpenRefine cells:

ID	DISPLAYNAME	ULAN ID	Sync ID	Pref Bio
	Maxwell Hamilton <input checked="" type="checkbox"/> Maxwell, Hamilton (24) <input checked="" type="checkbox"/> Hamilton, Maxwell, Inc. (23) <input checked="" type="checkbox"/> Fry, Edwin Maxwell (14) <input checked="" type="checkbox"/> Bates, Maxwell (14) <input checked="" type="checkbox"/> Maxwell, John (14) <input checked="" type="checkbox"/> Create new item	500299270	500299270	Scottish painter and architect, 1830-1923
	Charles Aubry <input checked="" type="checkbox"/> Aubry, Charles (34) <input checked="" type="checkbox"/> Aubry, Charles (33) <input checked="" type="checkbox"/> Aubry, Emile (18) <input checked="" type="checkbox"/> Aubry, Louis (18) <input checked="" type="checkbox"/> Aubry, René Marcel (17) <input checked="" type="checkbox"/> Create new item	500029715	500029715	French photographer and designer, 1811-1877
	Fédèle Azari <input checked="" type="checkbox"/> Azari, Fedele (19) <input checked="" type="checkbox"/> Āzārī Isfarāyīnī, Ḥamzah ibn 'Alī Mālik (5) <input checked="" type="checkbox"/> Create new item	500103519	500103519	Italian photographer, 1895-1930
	František Vobecký <input checked="" type="checkbox"/> Vobecký, František (43) <input checked="" type="checkbox"/> Schmoranz, František (17) <input checked="" type="checkbox"/> Fišer, František (17) <input checked="" type="checkbox"/> Kadlec, František (17) <input checked="" type="checkbox"/> Kysela, František (17) <input checked="" type="checkbox"/> Create new item	500028103	500028103	Czech photographer, 1902-1990
	Eugen Wiškovský <input checked="" type="checkbox"/> Wiškovský, Eugen (18) <input checked="" type="checkbox"/> Kahler, Eugen von (17) <input checked="" type="checkbox"/> Rzibek, Eugen (17) <input checked="" type="checkbox"/> Nevan, Eugen (17) <input checked="" type="checkbox"/> Osswald, Eugen (13) <input checked="" type="checkbox"/> Create new item	500031206	500031206	Czech photographer, 1888-1964

Parsing Variant Terms from Getty Vocabulary XML

Add column, and using full XML, parse all variant names for an artist/person:

```
join(forEach(value.parseXml().select('Vocabulary')[0].select('Subject')[0].select('Terms')[0].select('Non-Preferred_Term'),term,term.select('Term_Text')[0].ownText()),", ")
```

Add column based on column ULAN XML

New column name

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

Expression Language

```
join(forEach(value.parseXml().select('Vocabulary')[0].select('Subject')[0].select('Terms')[0].select('Non-Preferred_Term'),term,term.select('Term_Text')[0].ownText()),", ")
```

No syntax error.

Preview History Starred Help

row	value	join(forEach(value.parseXml()....
1.	<?xml version="1.0" encoding="utf-8"?> <Vocabulary xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="http://vocabss: <Subject Subject_ID="500299270"> <Parent_Relationships> <Preferred_Parent> <Parent_Subject_ID>500000002</Parent_Subject	Hamilton Maxwell, H. Maxwell, Maxwell

OK Cancel

Variant names, comma-separated, in OpenRefine cells:

ID	DISPLAYNAME	ULAN ID	Sync ID	Variants	Pref Bio	Pref Term
	Maxwell Hamilton <input checked="" type="checkbox"/> Maxwell, Hamilton (24) <input checked="" type="checkbox"/> Hamilton, Maxwell, Inc. (23) <input checked="" type="checkbox"/> Fry, Edwin Maxwell (14) <input checked="" type="checkbox"/> Bates, Maxwell (14) <input checked="" type="checkbox"/> Maxwell, John (14) <input checked="" type="checkbox"/> Create new item	500299270	500299270	Hamilton Maxwell, H. Maxwell, Maxwell	Scottish painter and architect, 1830-1923	Maxwell, Hamilton
	Charles Aubry <input checked="" type="checkbox"/> Aubry, Charles (34) <input checked="" type="checkbox"/> Aubry, Charles (33) <input checked="" type="checkbox"/> Aubry, Emile (18) <input checked="" type="checkbox"/> Aubry, Louis (18) <input checked="" type="checkbox"/> Aubry, René Marcel (17) <input checked="" type="checkbox"/> Create new item	500029715	500029715	Charles Aubry, Aubry, Charles-Hippolyte	French photographer and designer, 1811-1877	Aubry, Charles
	Fédèle Azari <input checked="" type="checkbox"/> Azari, Fedele (19) <input checked="" type="checkbox"/> Āzārī Isfārāyīnī, Hamzah ibn 'Alī Mālik (5) <input checked="" type="checkbox"/> Create new item	500103519	500103519	Fedele Azari	Italian photographer, 1895-1930	Azari, Fedele
	František Vobecký <input checked="" type="checkbox"/> Vobecký, František (43) <input checked="" type="checkbox"/> Schmoranz, František (17) <input checked="" type="checkbox"/> Fišer, František (17) <input checked="" type="checkbox"/> Kadlec, František (17) <input checked="" type="checkbox"/> Kysela, František	500028103	500028103	Vobecký, Frantisek, František Vobecký, Vobecký, Frantisek	Czech photographer, 1902-1990	Vobecký, František

Getty

CIDOC - December 9, 2020

Using OpenRefine with the SPARQL Endpoint – and TGN

Get preferred term by adding column by “fetching URLs” based on TGN ID, and input formula:

[http://vocab.getty.edu/sparql.json?query=select+\(str\(?t\)+as+?term\)+{tgn:'+escape\(value,'url'\)+'+gvp:prefLabelGVP/gvp:term+?t}'](http://vocab.getty.edu/sparql.json?query=select+(str(?t)+as+?term)+{tgn:'+escape(value,'url')+'+gvp:prefLabelGVP/gvp:term+?t})

Add column by fetching URLs based on column TGN ID

New column name: Pref Term Throttle delay: 1000 milliseconds

On error: set to blank store error Cache responses

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

Formulate the URLs to fetch:

Expression: `'http://vocab.getty.edu/sparql.json?query=select+(str(?t)+as+?term)+{tgn:'+escape(value,'url')+'+gvp:prefLabelGVP/gvp:term+?t}'` Language: General Refine Expression Language (GREL) No syntax error.

Preview History Starred Help

row	value
1.	7001371
2.	7000354
3.	7013051
4.	1079604

OK Cancel

Transform resulting JSON: value.parseJson().results.bindings[0].term.value

Custom text transform on column Pref Term

Expression Language

No syntax error.

Preview History Starred Help

row	value	value.parseJson().results.bind ...
1.	{ "head": { "vars": ["term"] }, "results": { "bindings": [{ "term": { "type": "literal", "value": "Yerushalayim" } }] }	Yerushalayim

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

OK Cancel

Get coordinates by adding column by “fetching URLs” based on TGN ID, and input formula:

'[http://vocab.getty.edu/sparql.json?query=select+?lat+?long+{tgn:'+escape\(value,'url'\)+'+foaf:place+?place+wgs:lat+?lat.+?place+wgs:long+?}'](http://vocab.getty.edu/sparql.json?query=select+?lat+?long+{tgn:'+escape(value,'url')+'+foaf:place+?place+wgs:lat+?lat.+?place+wgs:long+?}')'

The screenshot shows a dialog box titled "Add column by fetching URLs based on column TGN ID". The "New column name" is set to "LatLong", "Throttle delay" is 1000 milliseconds, and "On error" is set to "set to blank". The "Cache responses" checkbox is checked. The "HTTP headers to be used when fetching URLs" section has a "Show" link. Below it, the "Formulate the URLs to fetch:" section includes an "Expression" field containing the SPARQL query: '[http://vocab.getty.edu/sparql.json?query=select+?lat+?long+{tgn:'+escape\(value,'url'\)+'+foaf:place+?place+wgs:lat+?lat.+?place+wgs:long+?}'](http://vocab.getty.edu/sparql.json?query=select+?lat+?long+{tgn:'+escape(value,'url')+'+foaf:place+?place+wgs:lat+?lat.+?place+wgs:long+?}')'. A note says "No syntax error." Below the expression is a preview table with four rows, each showing a TGN ID and its corresponding URL. At the bottom are "OK" and "Cancel" buttons.

row	value	URL
1.	7001371	http://vocab.getty.edu/sparql.json?query=select+?lat+?long+{tgn:7001371+foaf:place+?place+wgs:lat+?lat.+?place+wgs:long+?long}
2.	7000354	http://vocab.getty.edu/sparql.json?query=select+?lat+?long+{tgn:7000354+foaf:place+?place+wgs:lat+?lat.+?place+wgs:long+?long}
3.	7013051	http://vocab.getty.edu/sparql.json?query=select+?lat+?long+{tgn:7013051+foaf:place+?place+wgs:lat+?lat.+?place+wgs:long+?long}
4.	1079604	http://vocab.getty.edu/sparql.json?query=select+?lat+?long+{tgn:1079604+foaf:place+?place+wgs:lat+?lat.+?place+wgs:long+?long}

Parse resulting JSON:
`value.parseJSON().results.bindings[0].lat.value+','+value.parseJSON().results.bindings[0].long.value`

Custom text transform on column LatLong

Expression Language General Refine Expression Language (GREL) ▾

```
value.parseJson().results.bindings[0].lat.value+','+value.parseJson().results.bindings[0].long.value
```

No syntax error.

Preview History Starred Help

row	value	value.parseJson().results.bind ...
1.	{ "head": { "vars": ["lat", "long"] }, "results": { "bindings": [{ "lat": { "datatype": "http://www.w3.org/2001/XMLSchema#decimal", "type": "literal", "value": "31.7667" } } } }	31.7667, 35.2333

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

OK Cancel

Resulting data enrichment:

All	TGN ID	Pref Term	LatLong
1.	7001371	Yerushalayim	31.7667, 35.2333
2.	7000354	Marrakech	31.633, -8.009
3.	7013051	Montréal	45.5, -73.6
4.	1079604	Qiryat Tiv'on	32.7167, 35.1333
5.	7010273	Sankt-Peterburg	59.8833, 30.25
6.	7001369	Tel Aviv-Yafo	32.0833, 34.7667
7.	1079614	Rishon LeZiyyon	31.95, 34.8
8.	7007567	New York	40.7, -74
9.	7023856	Shanghai	31.222222, 121.458056
10.	7001370	Ramat Gan	32.0833, 34.8167
11.	1079547	Kinneret	32.7167, 35.5667
12.	1079641	Umm el Fahm	32.5167, 35.15
13.	7001371	Yerushalayim	31.7667, 35.2333
14.	7001362	Hefa	32.8167, 34.9833
15.	7012178	Dushanbe	38.6333, 68.85
16.	7004443	Düsseldorf	51.216667, 6.766667
17.	7012333	Mulhouse	47.738082, 7.347932
18.	7001371	Yerushalayim	31.7667, 35.2333
19.	7001369	Tel Aviv-Yafo	32.0833, 34.7667
20.	7001369	Tel Aviv-Yafo	32.0833, 34.7667
21.	7001369	Tel Aviv-Yafo	32.0833, 34.7667
22.	1079466	Holon	32.0167, 34.7667

Getty Vocabulary Program OpenRefine service

<https://www.getty.edu/research/tools/vocabularies/obtain/openrefine.html>

Thank you.

Jonathan Ward
Sr Editor, Getty Vocabulary Program
Getty Research Institute

Getty

CIDOC - December 9, 2020