Data Science 1 - SS 2020

19.05.2020

start um 13:15 Uhr

Dr. Karsten Tolle

Team Steckbriefe (Abgabe über OLAT)

Wird nicht benotet! Es geht darum zu verstehen, wer was macht, um:

- Die Teambildung zu unterstützen.
- Zu verstehen, wer Hilfe beim Definieren der Ziele und Finden von Datenquellen braucht.
- Die Mentoring-Treffen vorzubereiten.

Jeder der ein Projekt bearbeiten möchte (also CPs erhalten will), sollte eine Datei hochladen (oder erstellen).

1. Falls ihr noch keinen Partner/Team habt, ladet bitte eine leere .txt-Datei mit dem Namen: ohneTeam.txt hoch.
2. Wer einen Partner/Team hat:
   a) Wählt für Euer Team einen kurzen Namen (ein Wort)!
   b) Einer vom Team (ihr entscheidet wer) lädt einen Steckbrief des Teams hoch (entsprechend der Vorlage unten). Die Datei soll dabei wie folgt benannt werden: <Team-Name>_Steckbrief.txt
   c) Alle anderen vom Team laden eine leere Datei hoch mit dem Namen: <Team-Name>.txt

… aktuell 14 Teams (stand 18.05. 9:00 Uhr) eingetragen …

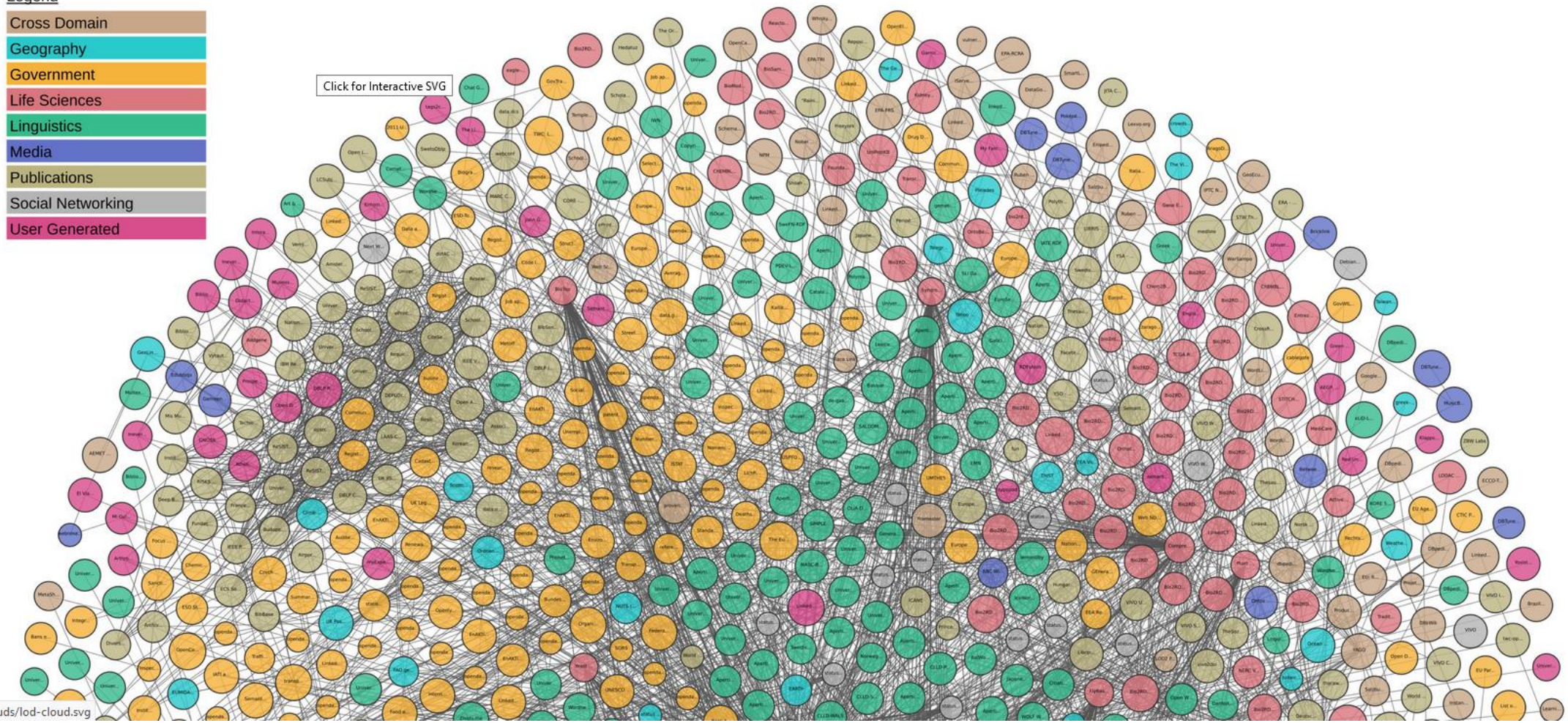
Bitte Teams bis zum 22.05. (Freitag) eintragen!

The Linked Open Data Cloud

Legend

- Cross Domain
- Geography
- Government
- Life Sciences
- Linguistics
- Media
- Publications
- Social Networking
- User Generated

Click for Interactive SVG



Publicly Available Datasets

- The University of California at Irvine Machine Learning Repository (UCI) hosts famous data
  - http://archive.ics.uci.edu/ml/
- Many governments host open data
  - https://www.govdata.de/
  - http://data.gov.uk
  - http://data.gov
  - https://offenedaten.frankfurt.de
- Kaggle Datasets
  - www.kaggle.com/
- Else
  - Linked Open Data Cloud: https://lod-cloud.net/
  - Open Knowlegde Foundation Deutschland: https://okfn.de/
  - Nomisma.org: http://nomisma.org/datasets
  - Wikidata: https://www.wikidata.org
  - NASA: http://data.nasa.gov/
  - AWS: http://aws.amazon.com/publicdatasets/

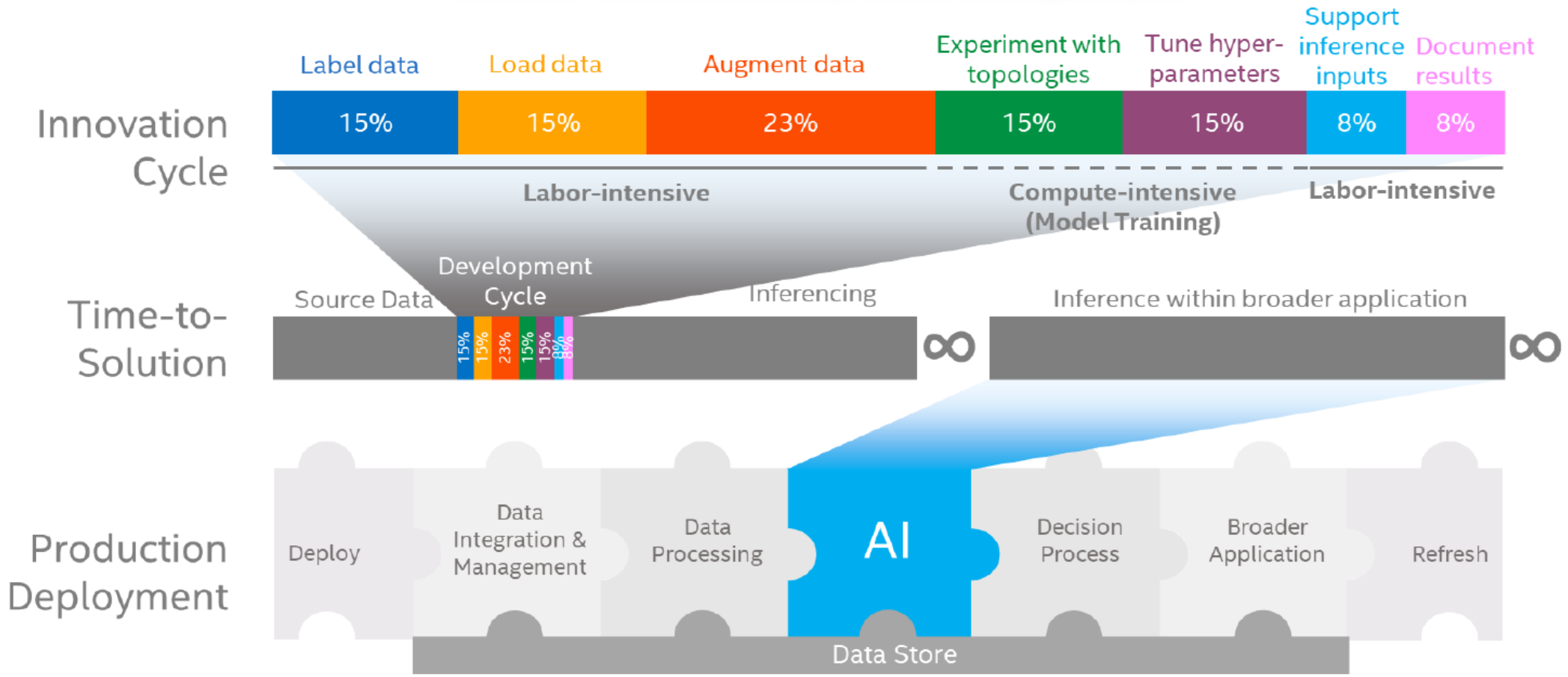
Unsupervised, Supervised, Semi-Supervised and Reinforcement Learning

Regression
Classification
Clustering
Decision Trees
Data Generation
Image Processing
Speech Processing
Natural Language Processing
Recommender Systems
Adversarial Networks
Reinforcement Learning

… https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/

- Supervised: Input and Output (labels - aka *ground truth*)
- Unsupervised: Input is given NO output!
- Semi-Supervised: mixture of the two above
- Reinforcement Learning: no output needed, but rewards are given – typically agents taking actions (https://en.wikipedia.org/wiki/Reinforcement_learning)

DEEP LEARNING IN PRACTICE



Time-to-solution is more significant than time-to-train

https://csc.uni-frankfurt.de/wiki/doku.php?id=public:events#intel_ai_workshopoctober_2019

Clustering

Clustering

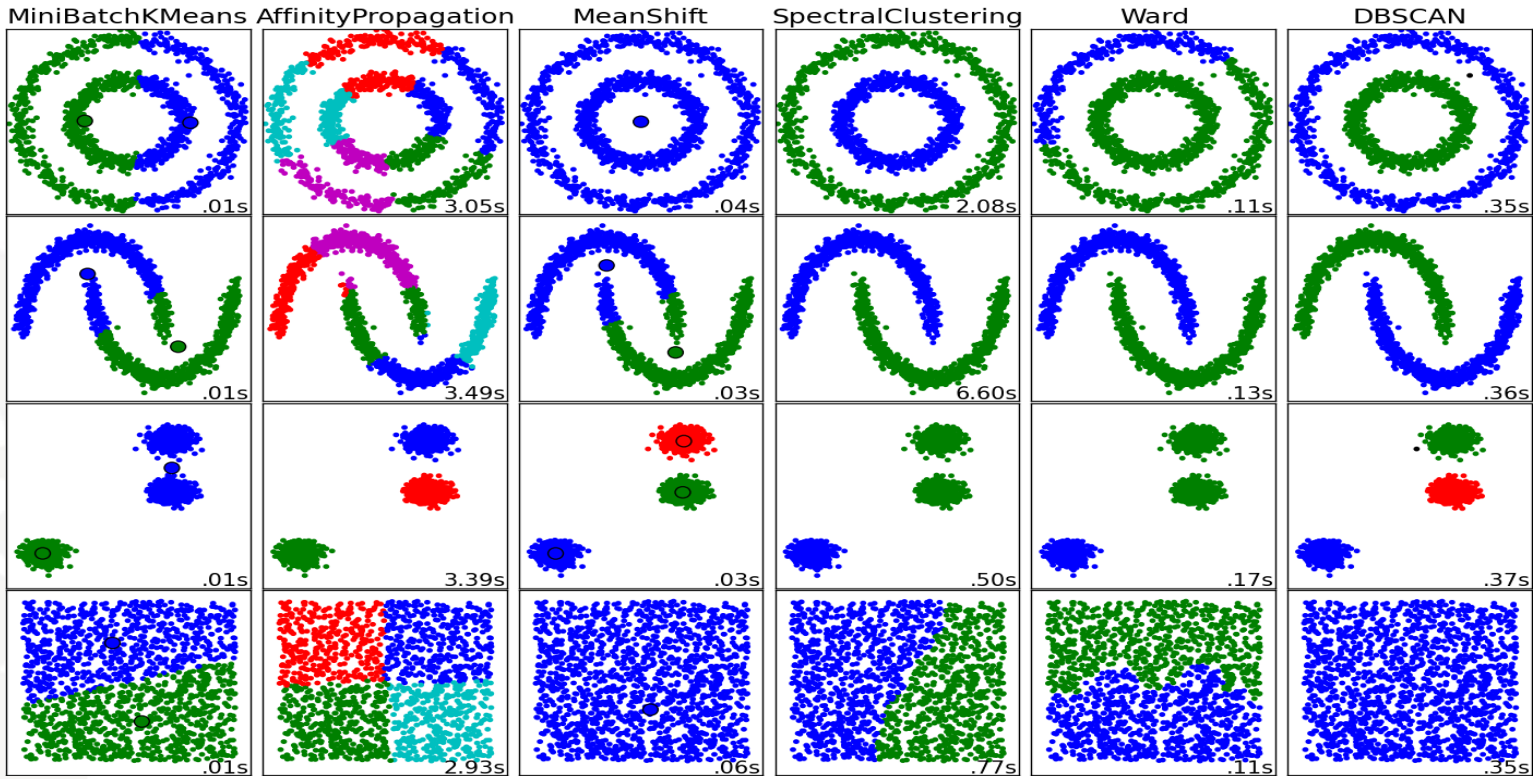
- One of Unsupervised Machine Learning Approaches
- Methods to identify possible groupings in a data-set
- Analyze the groupings automatically
- **cluster**: set of all data records that share a sub-set of the overall properties
- Infinitely many different possible clusterings (ways to cluster the data)
- Induced by the **similarity function** (some of which create the same output clusters)

Clustering Algortihms

- k-means clustering
- Hierarchical clustering
- Spectral clustering
- Density based clustering
- Graph-based clustering
- Fuzzy clustering
- Search-based clustering

Comparison of clustering algortihms

- The following image is from the Scikit-learn (https://scikit-learn.org/stable/) documentation and it demonstrates how the different algorithms compare to each other when executed on the same data:

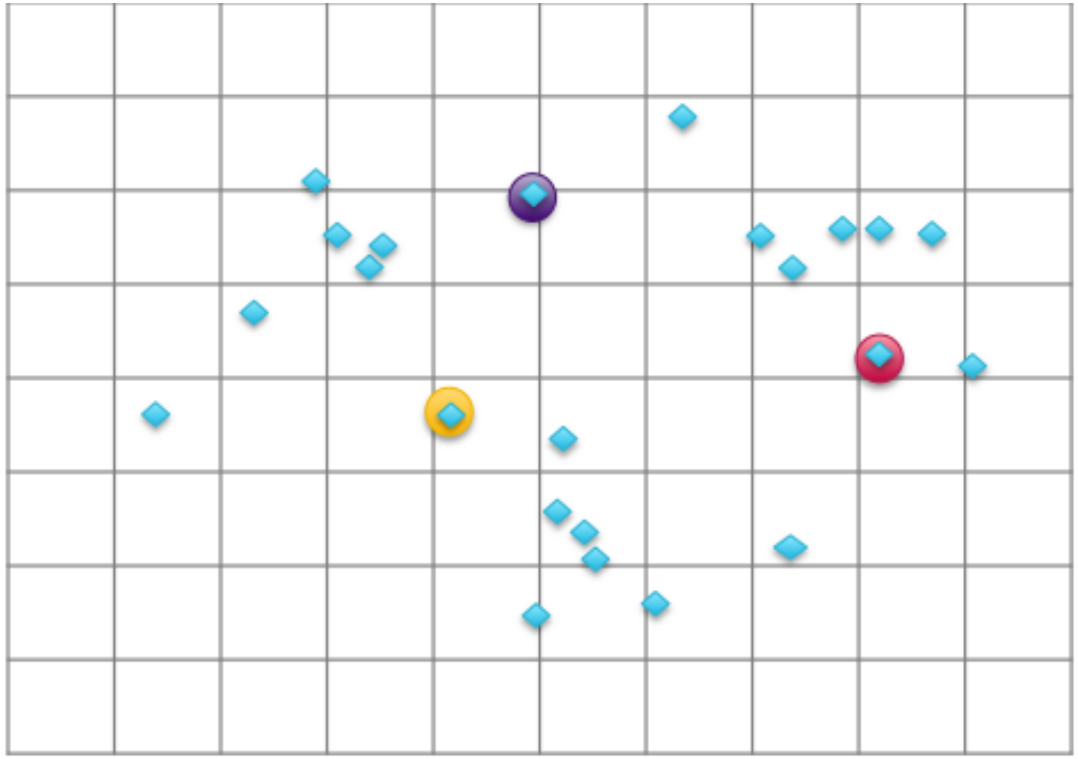


k-means

- Simple and efficient; best known partitional clustering algorithm
- Sensitive to choice of seed clusters (initial centroids)
- For every point in the dataset
  - Measure distance to each centroid
  - Assign point to centroid with lowest distance
- For every cluster
  - Calculate the mean position of all points
- Repeat until convergence criteria is met
  - No points change clusters
  - Centroid changes within a short distance
  - Fixed number of iterations

ref: Cloudera's academia teaching materials

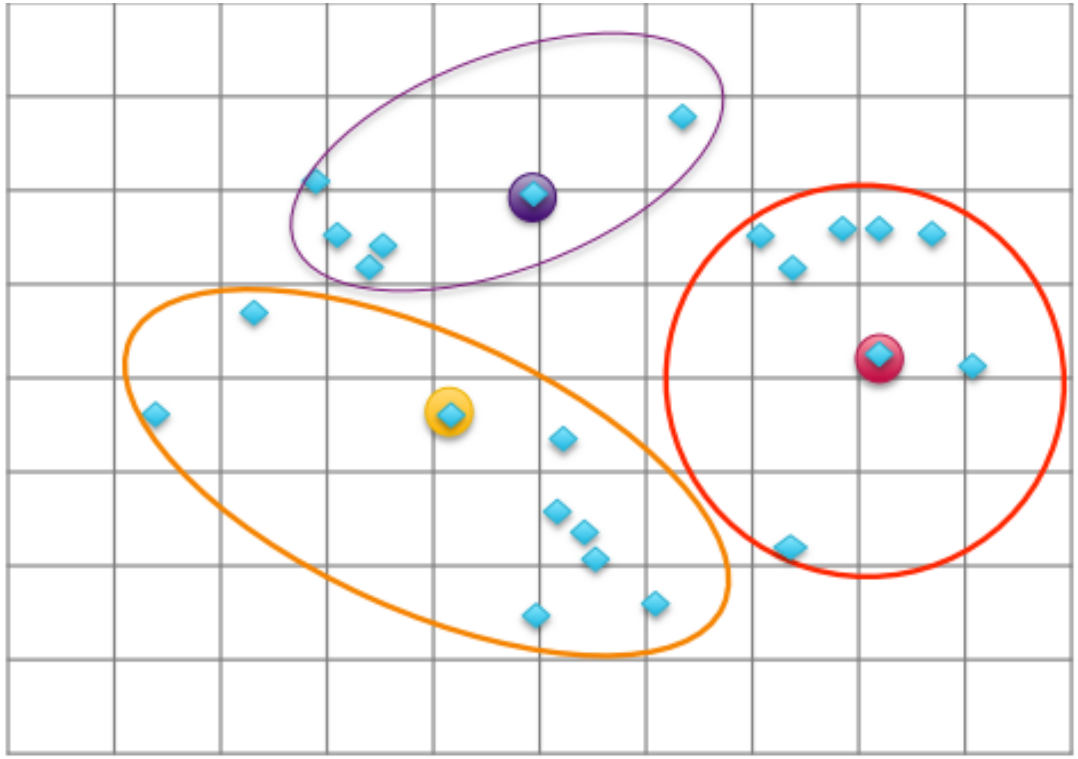
k-means demonstration



1. Choose *K* random points as starting centers

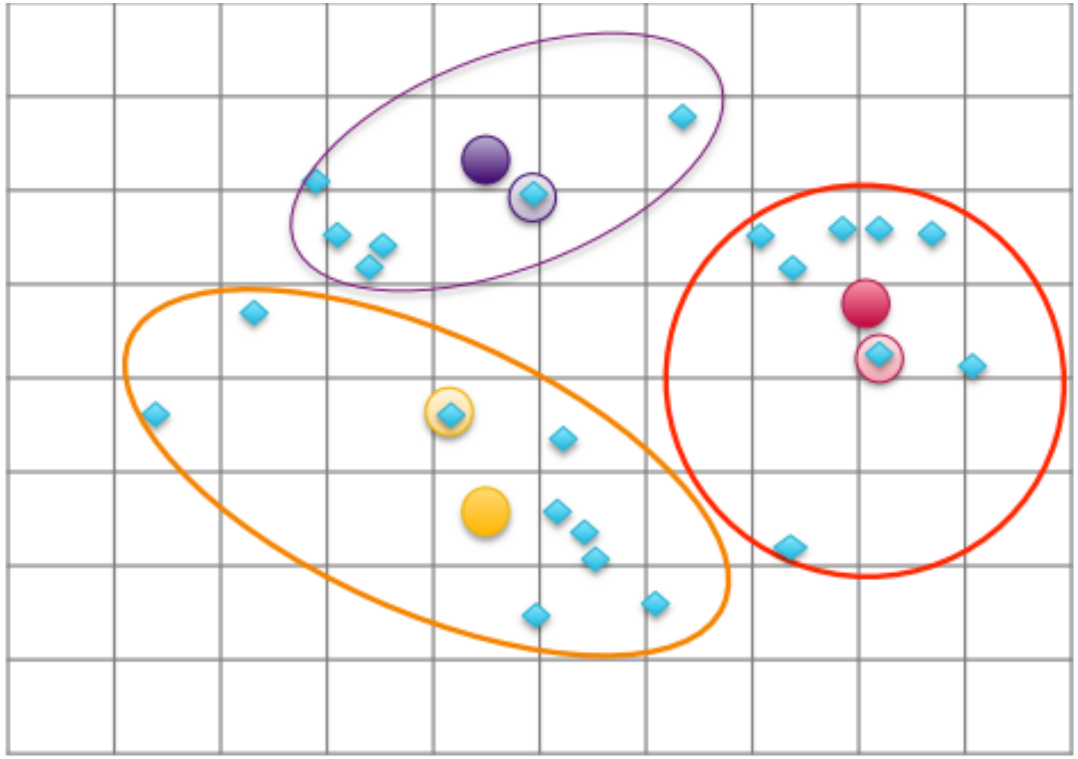
ref: Cloudera's academia teaching materials

k-means demonstration



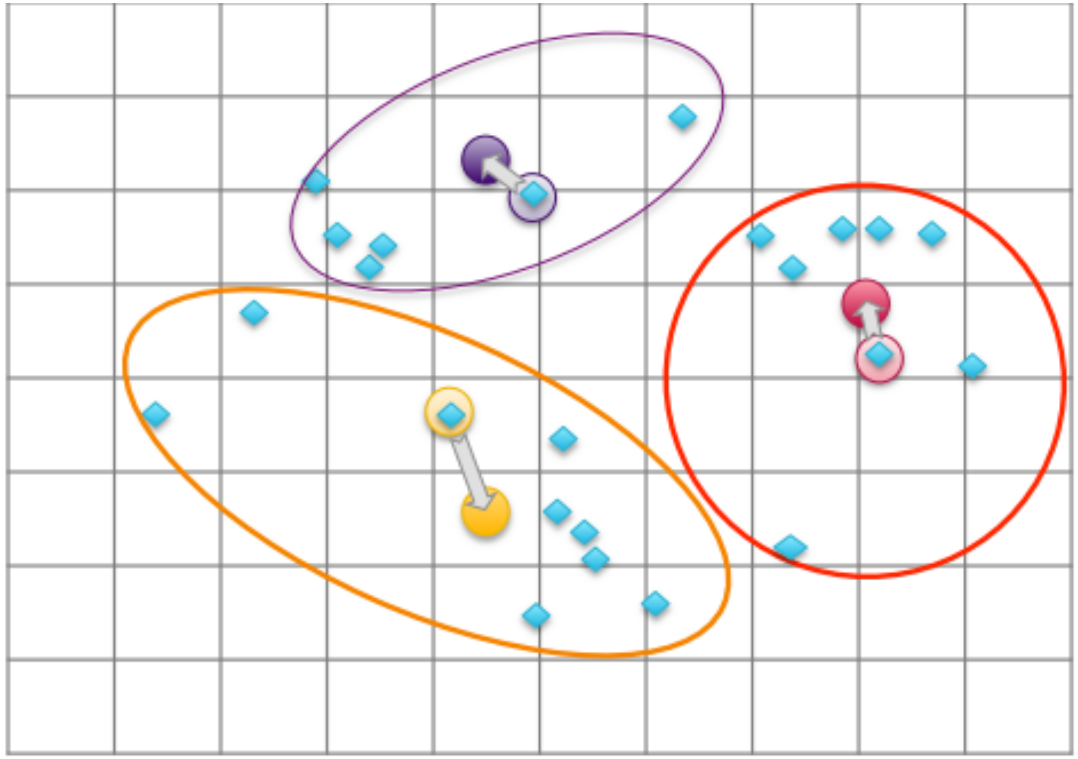
1. Choose *K* random points as starting centers
2. Find all points closest to each center

k-means demonstration



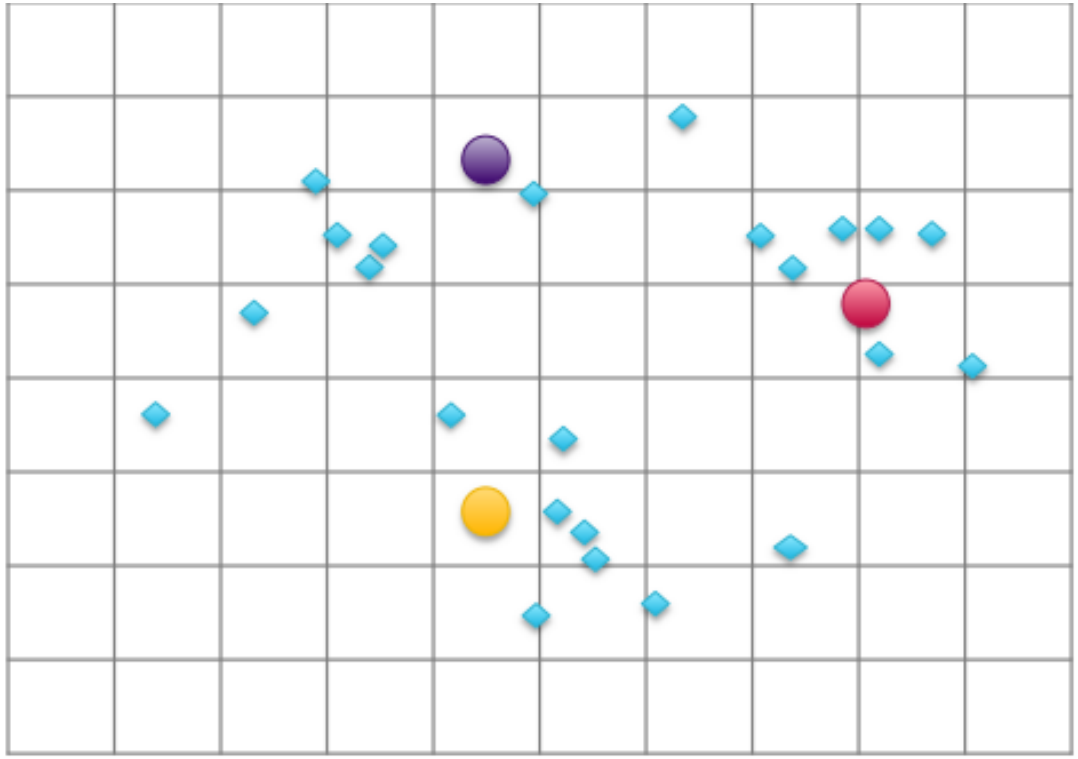
1. Choose *K* random points as starting centers
2. Find all points closest to each center
3. Find the center (mean) of each cluster

k-means demonstration



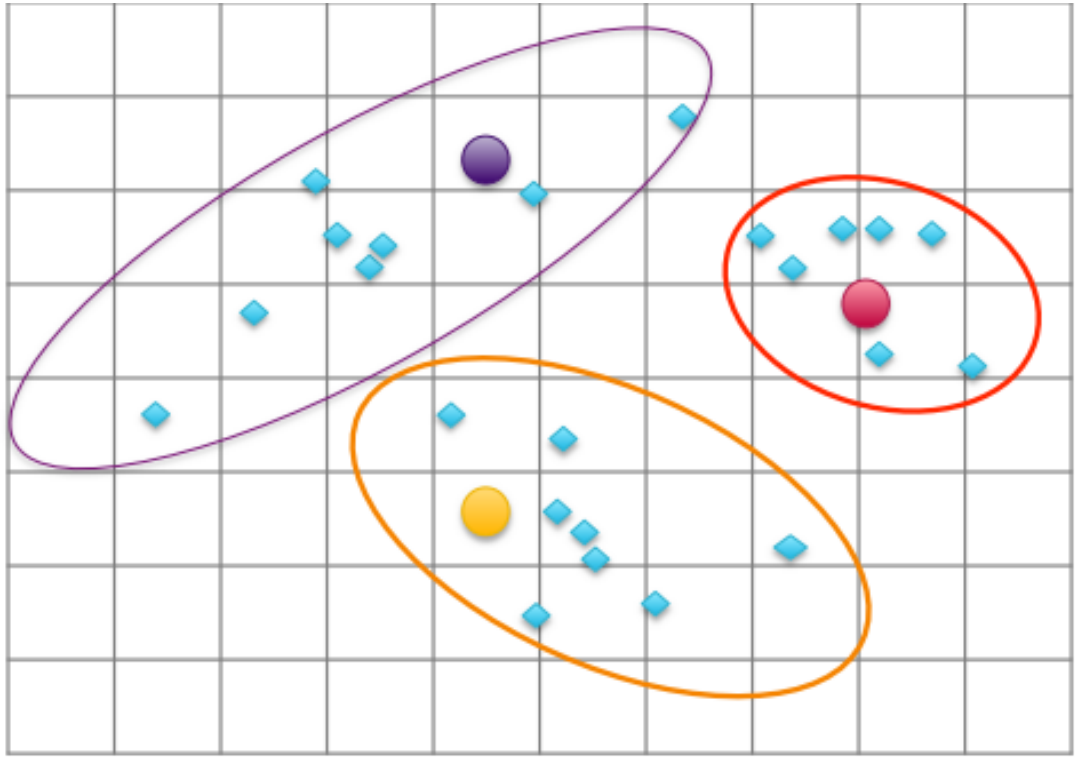
1. Choose *K* random points as starting centers
2. Find all points closest to each center
3. Find the center (mean) of each cluster
4. If the centers changed, iterate again

k-means demonstration



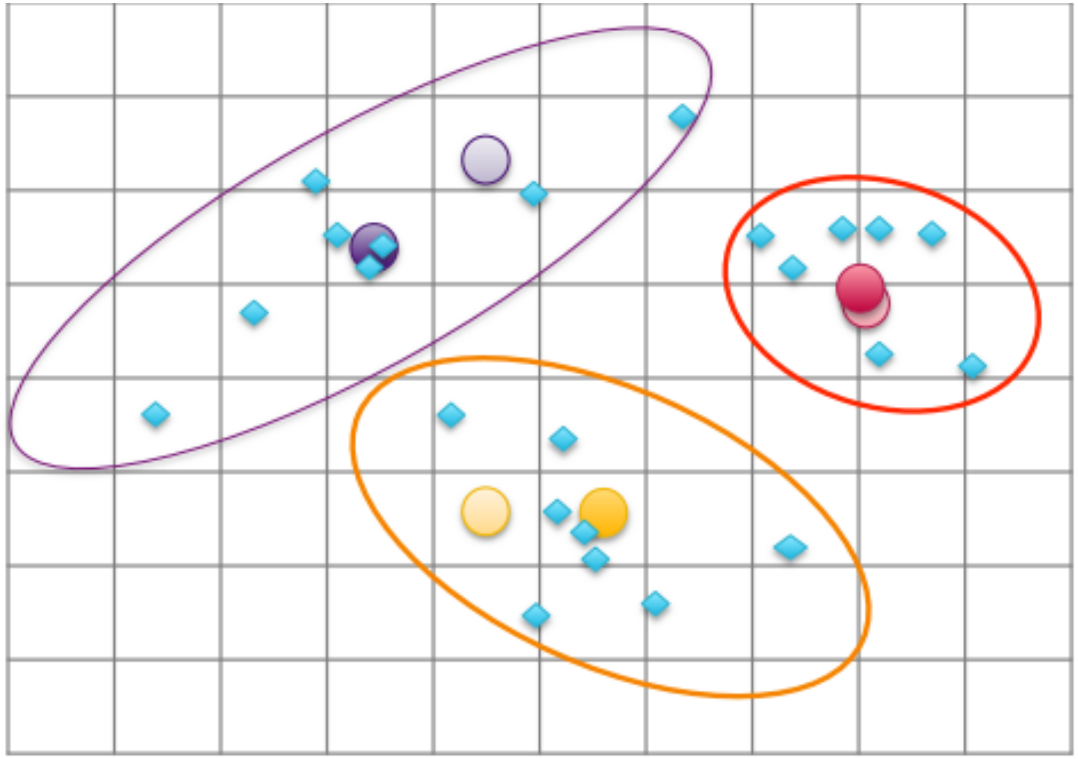
1. Choose *K* random points as starting centers
2. Find all points closest to each center
3. Find the center (mean) of each cluster
4. If the centers changed, iterate again

k-means demonstration



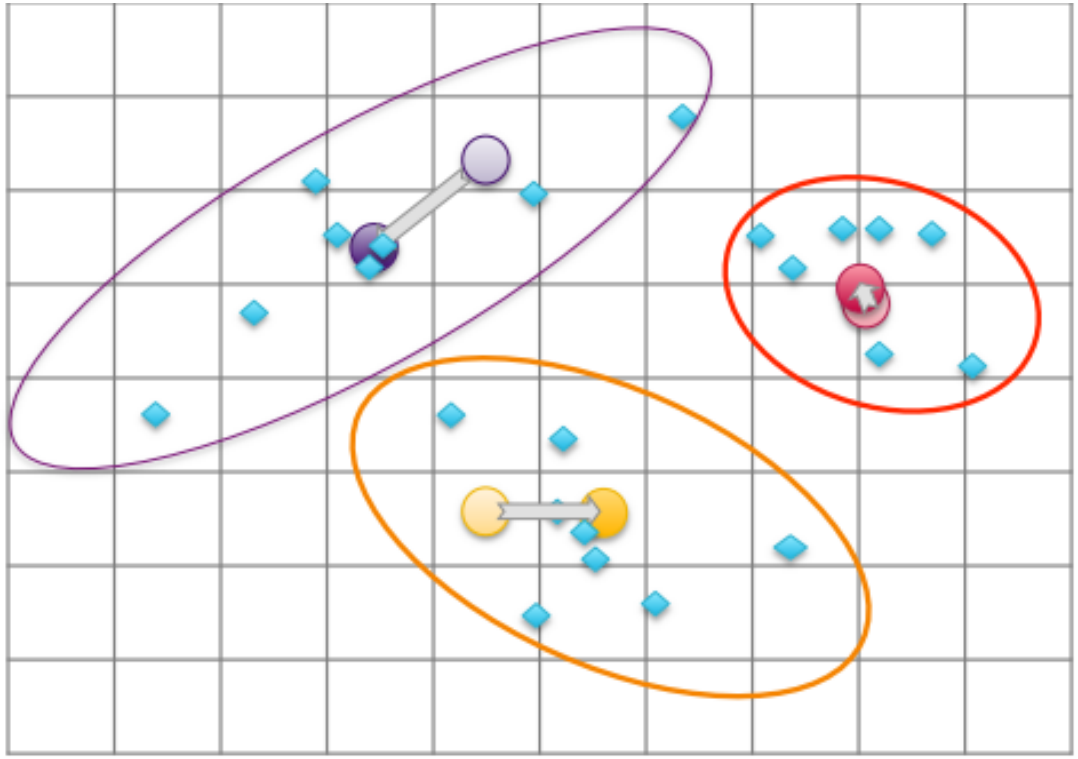
1. Choose *K* random points as starting centers
2. Find all points closest to each center
3. Find the center (mean) of each cluster
4. If the centers changed, iterate again

k-means demonstration



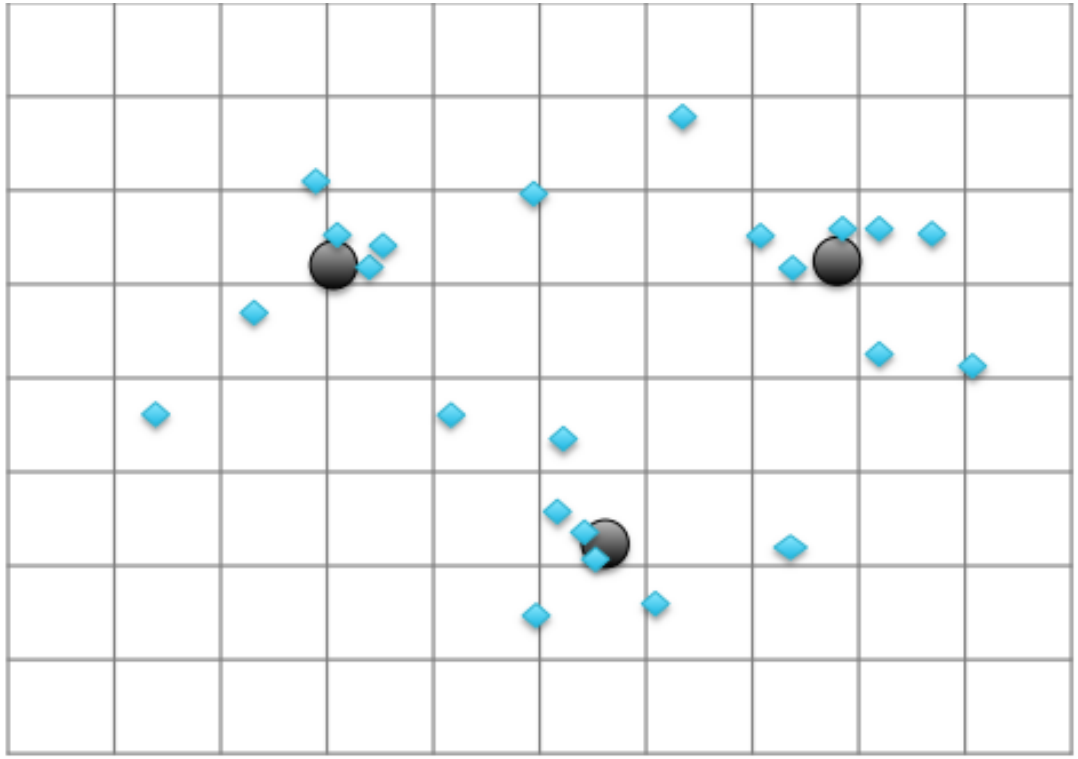
1. Choose *K* random points as starting centers
2. Find all points closest to each center
3. Find the center (mean) of each cluster
4. If the centers changed, iterate again

k-means demonstration



1. Choose *K* random points as starting centers
2. Find all points closest to each center
3. Find the center (mean) of each cluster
4. If the centers changed, iterate again

k-means demonstration



1. Choose *K* random points as starting centers
2. Find all points closest to each center
3. Find the center (mean) of each cluster
4. If the centers changed, iterate again

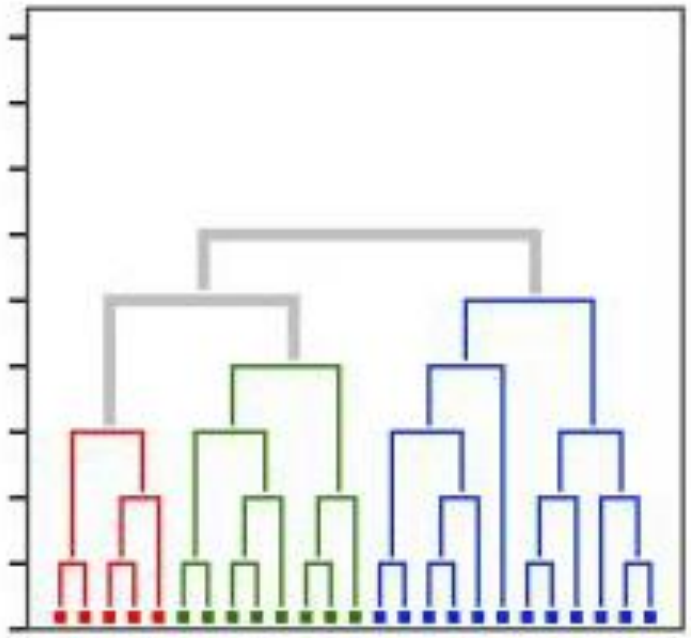
   ...
5. Done!

… how to find the right number of clusters?

https://en.wikipedia.org/wiki/Determining_the_number_of_clusters_in_a_data_set

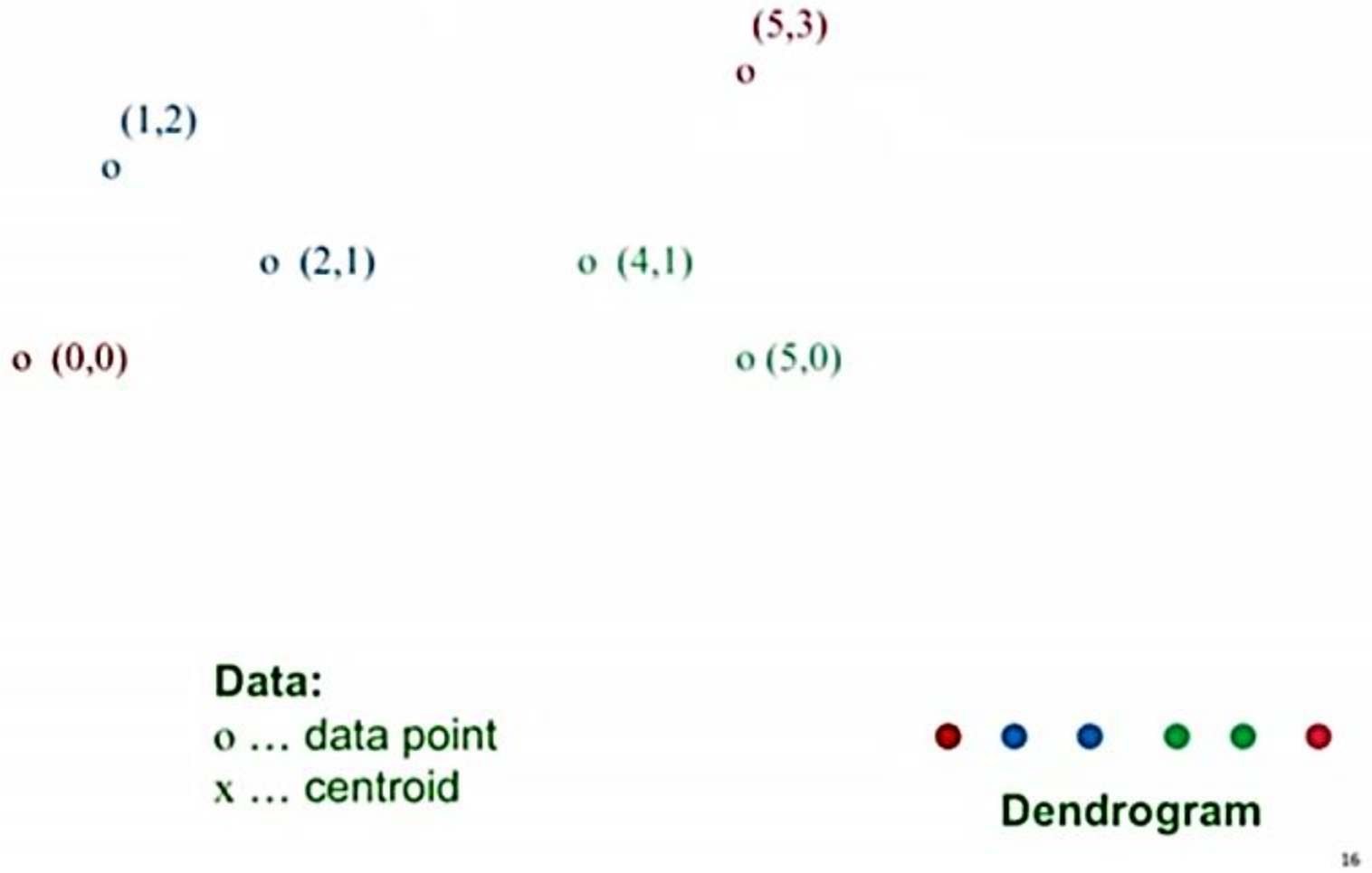
Hierarchical clustering demonstration

- Easy to understand
- A dendrogram can be used to depict this hierarchical clustering process
- Bottom-up (=agglomerative) method: iterate over each data record
- Compare two data records using a similarity function, and join most similar pair

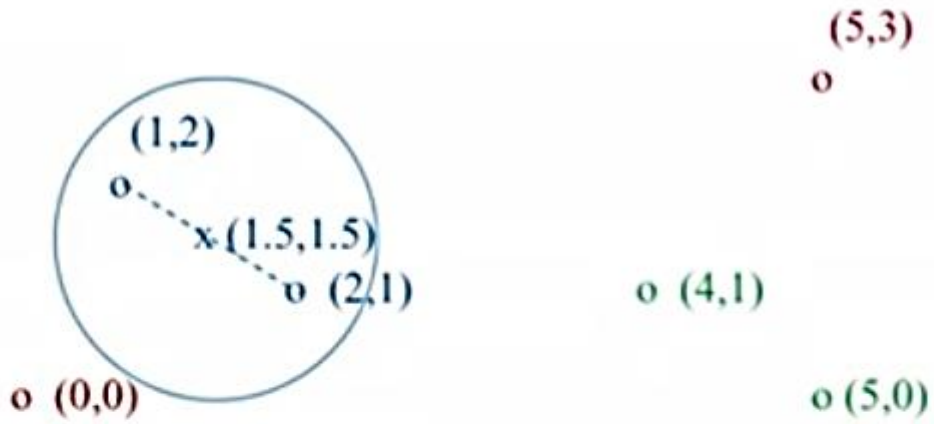


dendrogram

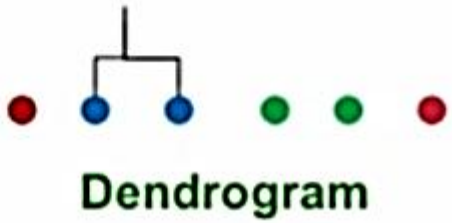
Hierarchical clustering demonstration



Hierarchical clustering demonstration



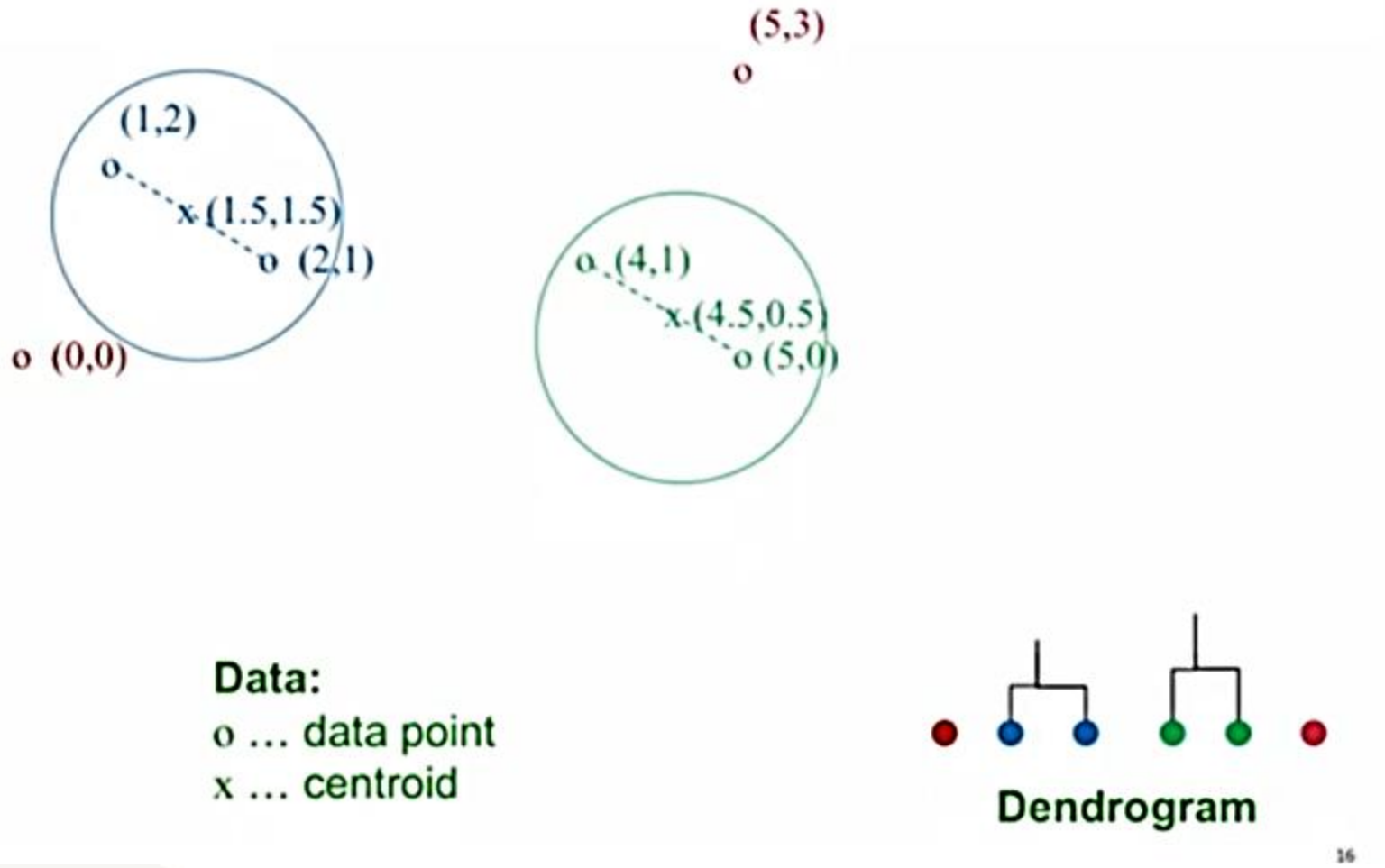
Data:
o … data point
x … centroid



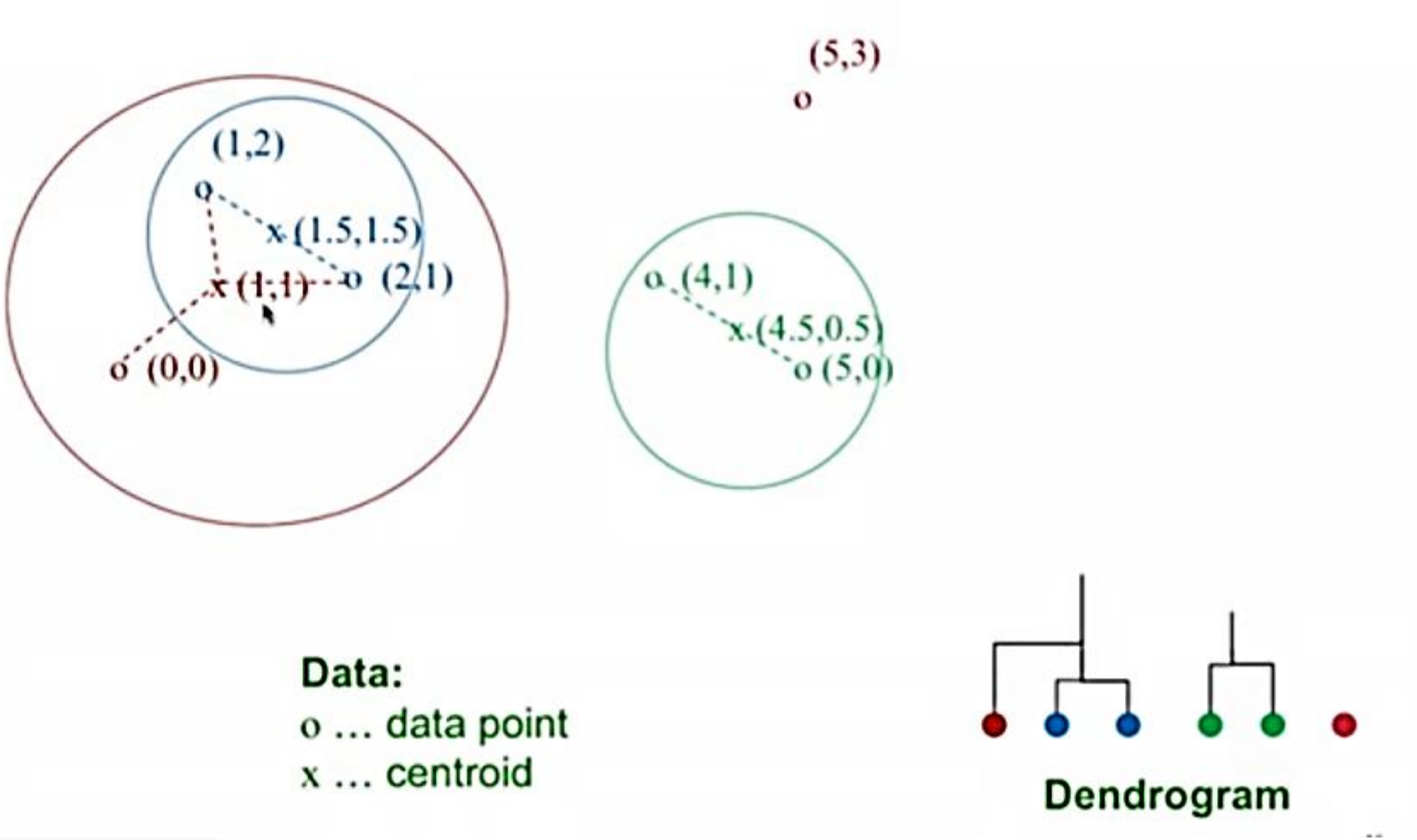
Dendrogram

16

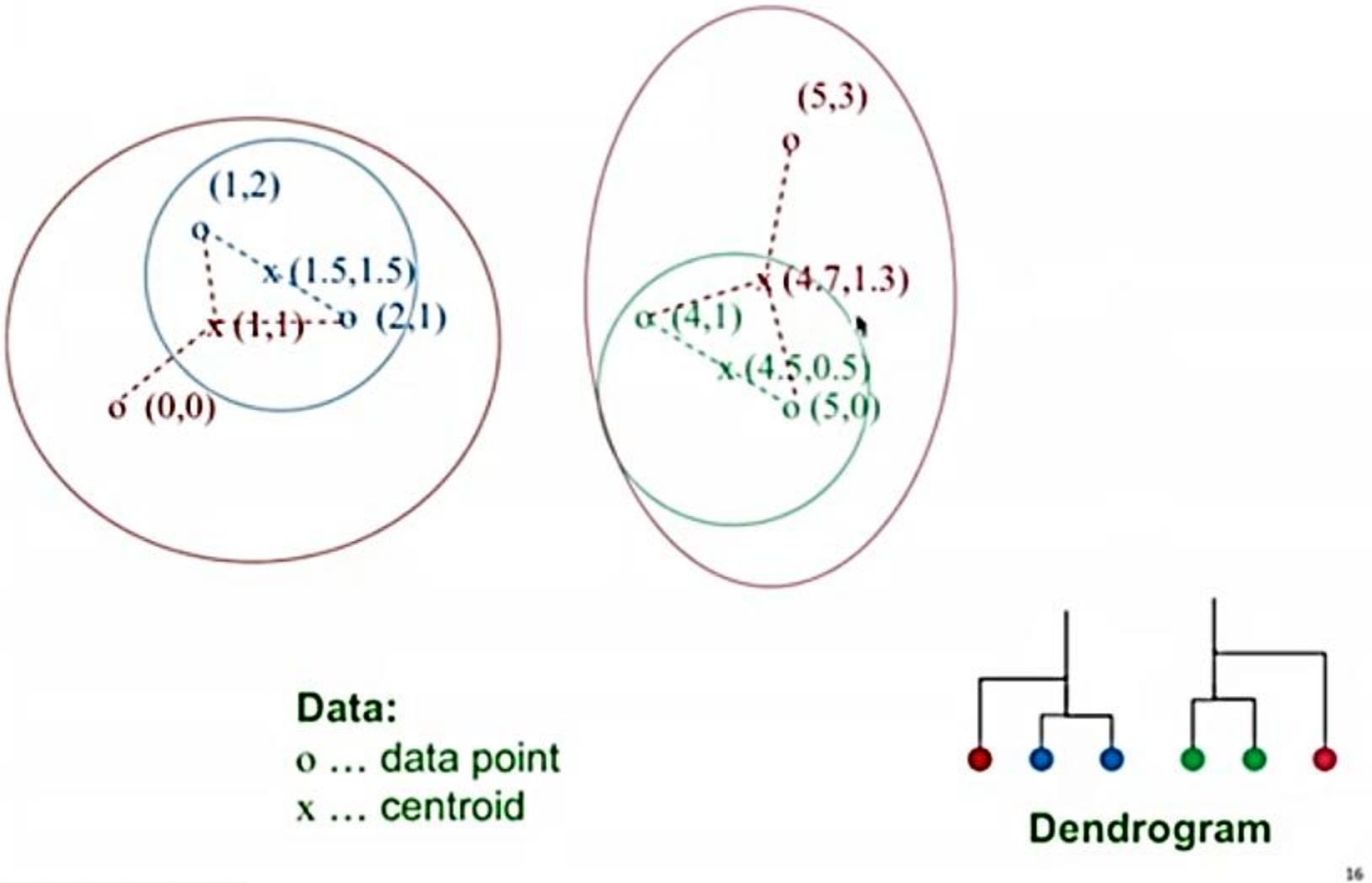
Hierarchical clustering demonstration



Hierarchical clustering demonstration



Hierarchical clustering demonstration



Unterschied zwischen: relationalen Daten und LOD

Frage 1: Wie merge ich zwei Datenbank-Tabellen aus verschiedenen Datenbanken, welche den gleichen Entitätstypen repräsentieren und die gleichen Attribute haben?
Z.B. Prägeherr (Person)

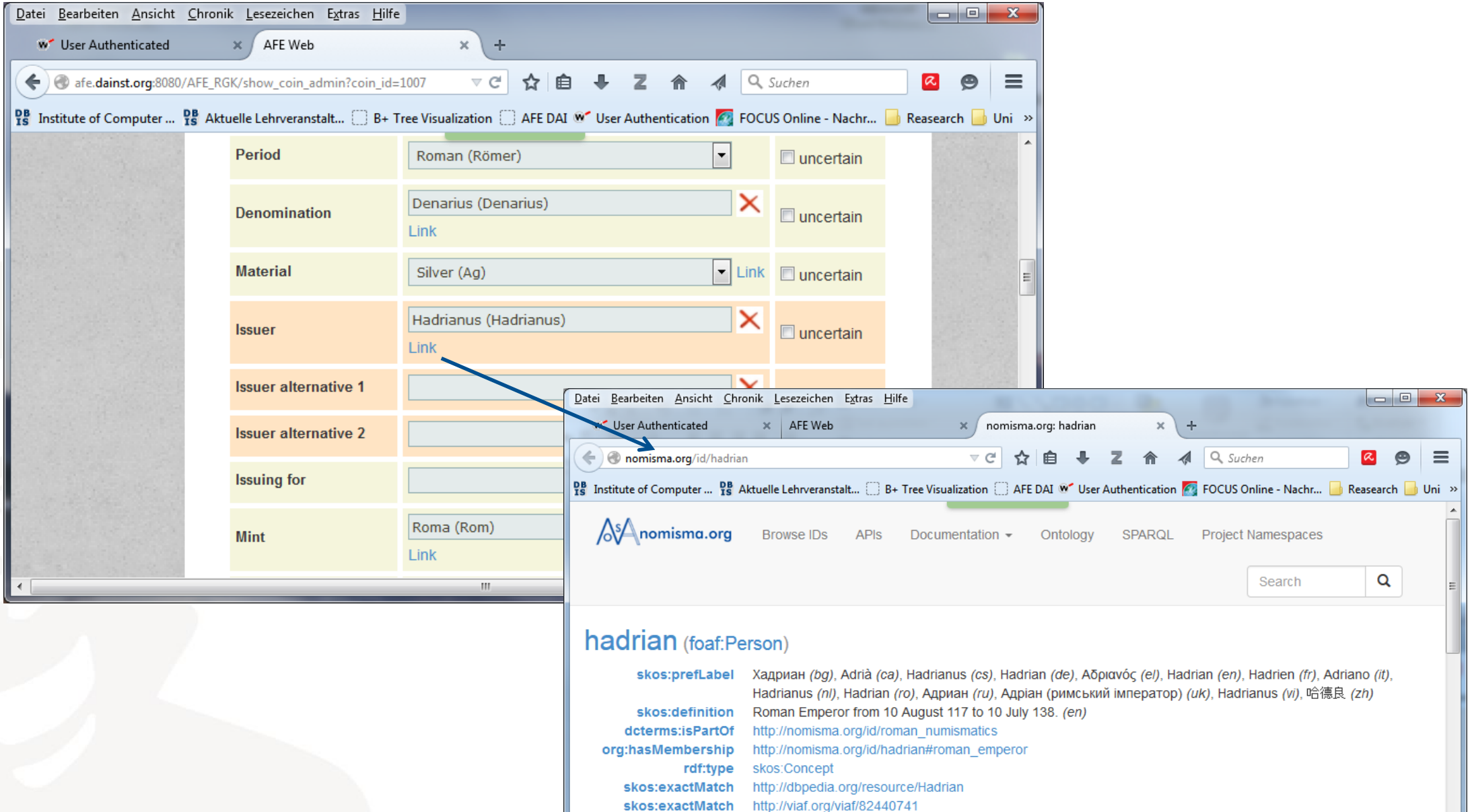
ID	Name	active_from	active_to
1	Domitianus II	271	271

ID	Name	active_from	active_to
14	Domitianus II.	270	271

Frage 2: Wie wäre das mit RDF/LOD?

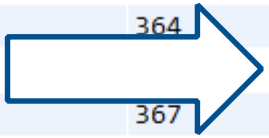
26

Zero Star – existierende LOD nutzen

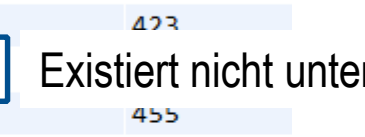


Zero Star – existierende LOD nutzen

id	name	Nomisma	active from	active to
158	Magnentius	magnentius	350	353
161	Iulianus II.	julian_the_apostate	355	363
162	Valentinianus I.	valentinian_i	364	375
163	Valens	valens		
165	Gratianus	gratian	367	383
167	Valentinianus II.	valentinian_ii	375	392
168	Valentinianische Dyna...		364	392
169	Magnus Maximus	magnus_maximus	383	388
170	Honorius	honorius	393	423
171	Theodosianische Dynastie			
173	Valentinianus III.	valentinian_iii	425	455
174	Iustinianus I.	justinian_i	527	565
182	Macrinus	macrinus	217	218
183	Flavisch		69	96
187	Maximinus II.	maximinus_daia	305	313
189	Licinius I.	licinius	308	324
194	Constantius I.	constantius_chlorus	293	306
200	Domitianus II.	domitianus	271	271
202	Aemilianus	aemilianus	253	253
203	Allectus	allectus	293	296
205	Anastasius I.	anastasius	461	518
207	Antehmius	anthemius	467	472
208	Arcadius	arcadius	383	408
209	Avitus	avitus	455	456

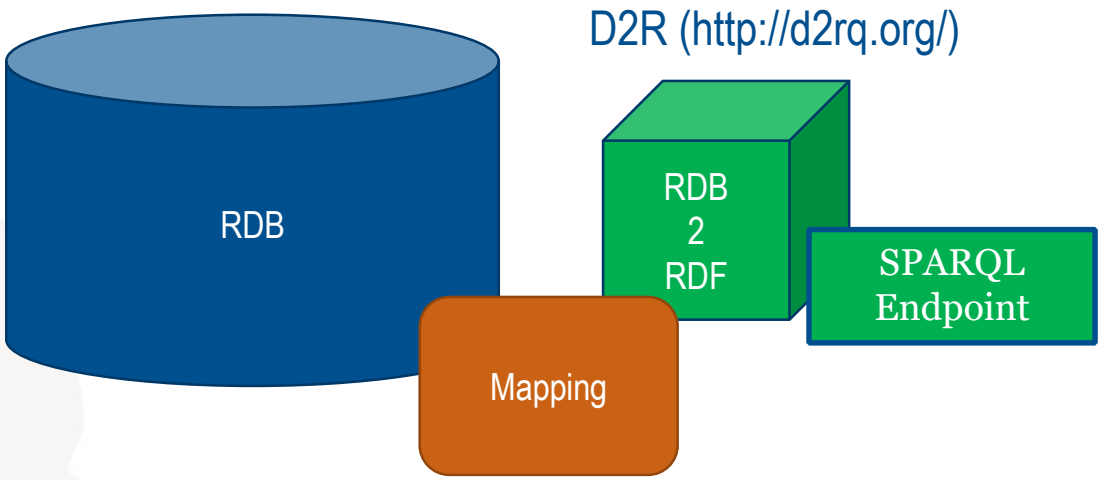


http://nomisma.org/id/valens



Existiert nicht unter Nomisma.org

Von einer SQL-Datenbank zu 4-Star LOD (oder mehr)



RDB 2 RDF: http://www.w3.org/2001/sw/rdb2rdf/wiki/Implementations

D2R (http://d2rq.org/)

Description of http://afe.dainst.org:8080/d2rq/resource/AFE_coin_1678

Resource URI: http://afe.dainst.org:8080/d2rq/resource/AFE_coin_1678

Home | All coinfind

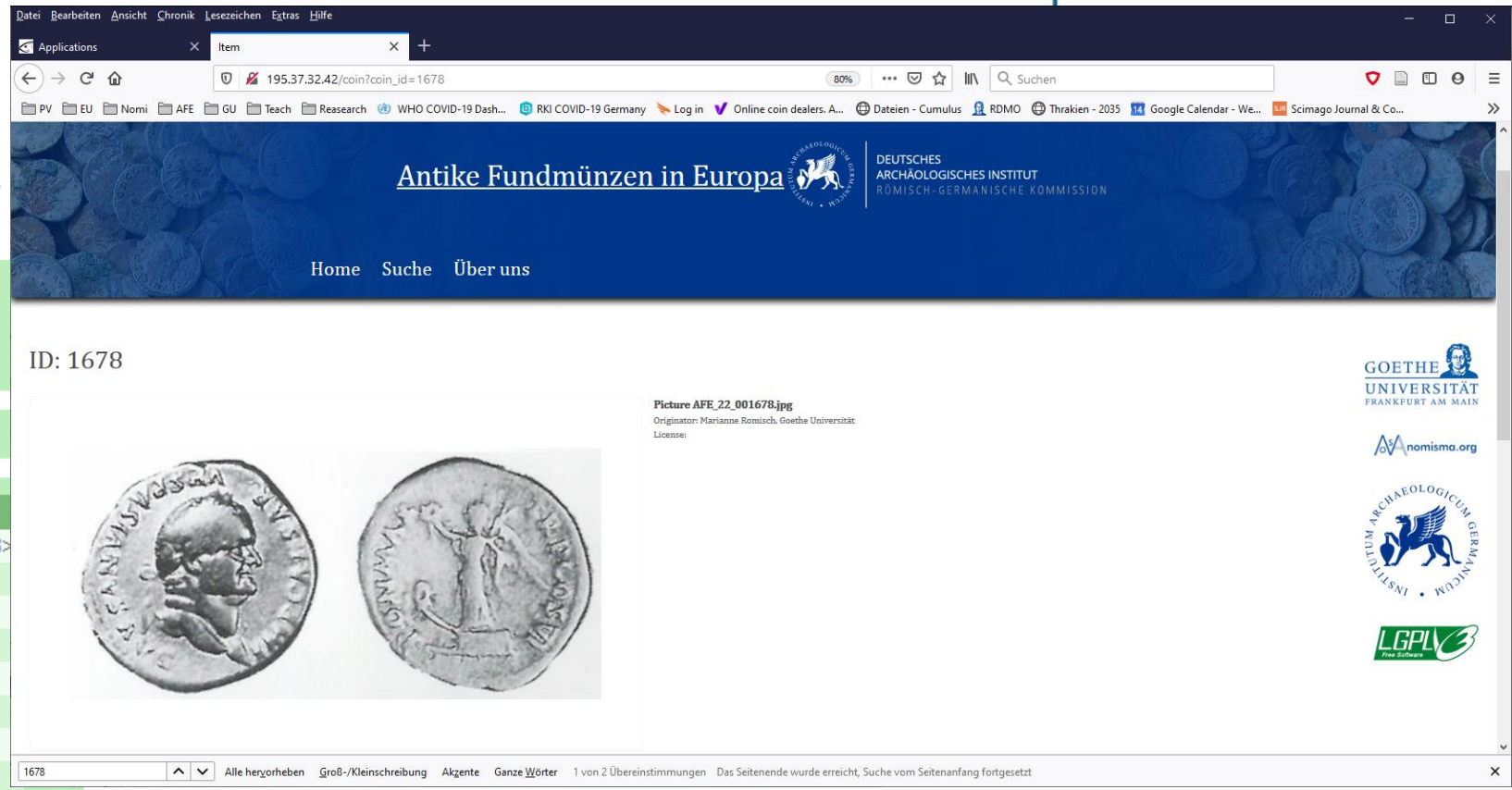
Property	Value
skos:exactMatch	<http://afe.dainst.org:8080/AFE_RGK/show_coin?coin_id=1678>
nmo:hasDenomination	nm:denarius
nmo:hasEndDate	75 (xsd:string)
nmo:hasFindspot	db:AFE_place_131
nmo:hasIssuer	nm:vespasian
nmo:hasMaterial	nm:ar
nmo:hasMint	nm:rome
nmo:hasObjectType	nm:coin
nmo:hasStartDate	75 (xsd:string)
nmo:hasTypeSeriesItem	<http://numismatics.org/ocre/id/ric.2_1(2).ves.777>
rdf:type	nmo:NumismaticObject

The server is configured to display only a limited number of values (limit per property bridge: 100).

Metadata

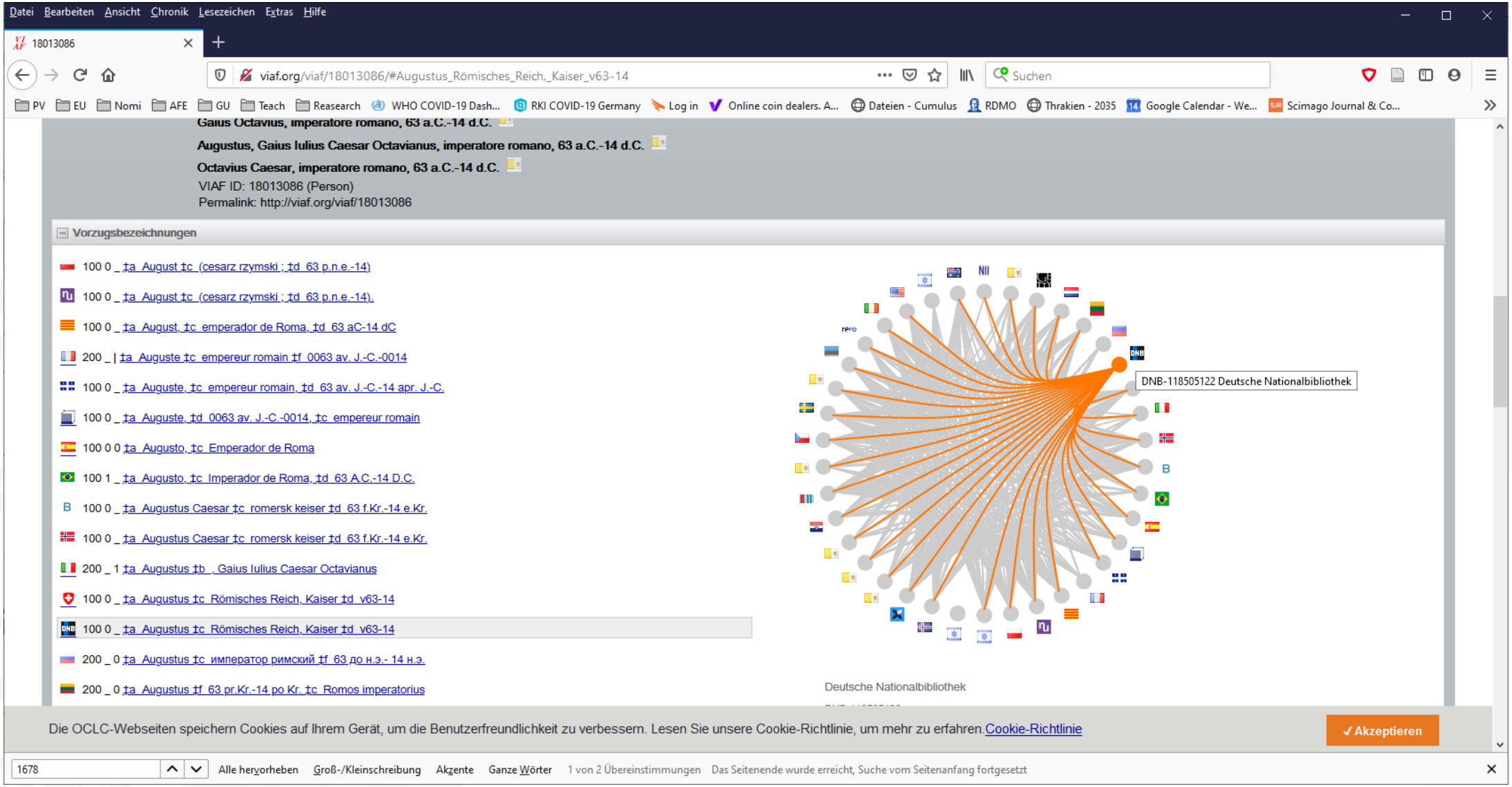
<http://afe.dainst.org:8080/d2rq/data/AFE_coin_1678>

dc:date	2015-03-26T15:39:47.027Z
prv:containedBy	<http://afe.dainst.org:8080/d2rq/dataset>
void:inDataset	<http://afe.dainst.org:8080/d2rq/dataset>
rdf:type	prv:DataItem
rdf:type	foaf:Document



Maschinen verstehbare Schnittstelle

Probleme …


Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe
18013086
viaf.org/viaf/18013086/#Augustus_Römisches_Reich,_Kaiser_v63-14
Suchen
PV EU Nomi AFE GU Teach Reasearch WHO COVID-19 Dash... RKI COVID-19 Germany Log in Online coin dealers. A... Dateien - Cumulus RDMO Thrakien - 2035 Google Calendar - We... Scimago Journal & Co...
Gaius Octavius, imperatore romano, 63 a.C.-14 d.C.
Augustus, Gaius Iulius Caesar Octavianus, imperatore romano, 63 a.C.-14 d.C.
Octavius Caesar, imperatore romano, 63 a.C.-14 d.C.
VIAF ID: 18013086 (Person)
Permalink: http://viaf.org/viaf/18013086
Vorzugsbezeichnungen
100 0 _ ‡a August ‡c (cesarz rzymski ; ‡d 63 p.n.e.-14)
100 0 _ ‡a August ‡c (cesarz rzymski ; ‡d 63 p.n.e.-14).
100 0 _ ‡a August, ‡c emperador de Roma, ‡d 63 aC-14 dC
200 _ | ‡a Auguste ‡c empereur romain ‡f 0063 av. J.-C.-0014
100 0 _ ‡a Auguste, ‡c empereur romain, ‡d 63 av. J.-C.-14 apr. J.-C.
100 0 _ ‡a Auguste, ‡d 0063 av. J.-C.-0014, ‡c empereur romain
100 0 0 ‡a Augusto, ‡c Emperador de Roma
100 1 _ ‡a Augusto, ‡c Imperador de Roma, ‡d 63 A.C.-14 D.C.
100 0 _ ‡a Augustus Caesar ‡c romersk keiser ‡d 63 f.Kr.-14 e.Kr.
100 0 _ ‡a Augustus Caesar ‡c romersk keiser ‡d 63 f.Kr.-14 e.Kr.
200 _ 1 ‡a Augustus ‡b , Gaius Iulius Caesar Octavianus
100 0 _ ‡a Augustus ‡c Römisches Reich, Kaiser ‡d v63-14
100 0 _ ‡a Augustus ‡c Römisches Reich, Kaiser ‡d v63-14
200 _ 0 ‡a Augustus ‡c император римский ‡f 63 до н.э.- 14 н.э.
200 _ 0 ‡a Augustus ‡f 63 pr.Kr.-14 po Kr. ‡c Romos imperatorius
DNB-118505122 Deutsche Nationalbibliothek
Deutsche Nationalbibliothek
Die OCLC-Webseiten speichern Cookies auf Ihrem Gerät, um die Benutzerfreundlichkeit zu verbessern. Lesen Sie unsere Cookie-Richtlinie, um mehr zu erfahren. Cookie-Richtlinie
✓ Akzeptieren
1678
Alle hervorheben Groß-/Kleinschreibung Akzente Ganze Wörter 1 von 2 Übereinstimmungen Das Seitenende wurde erreicht, Suche vom Seitenanfang fortgesetzt

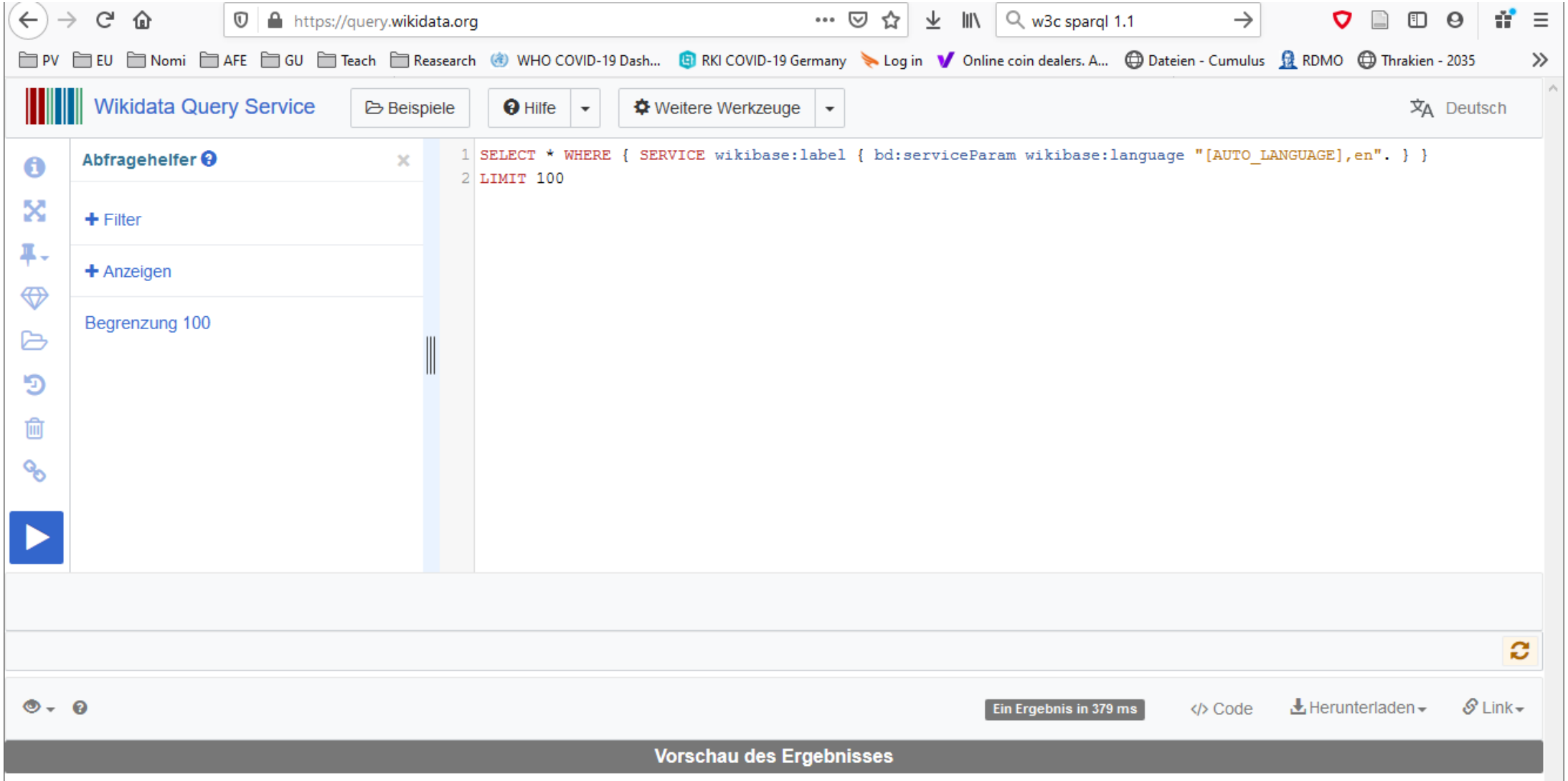
… nichts ist sicherer als Veränderung!

Amtlicher Gemeindeschlüssel – dient zum Identifizieren von Bundesländern, Kreisen, Gemeinden
(https://de.wikipedia.org/wiki/Amtlicher_Gemeindeschlüssel)

Gebietsreform: z.B. https://de.wikipedia.org/wiki/Kreisgebietsreform_Mecklenburg-Vorpommern_2011

Auswirkungen und Umgang?

Wikidata Anfrage SPARQL



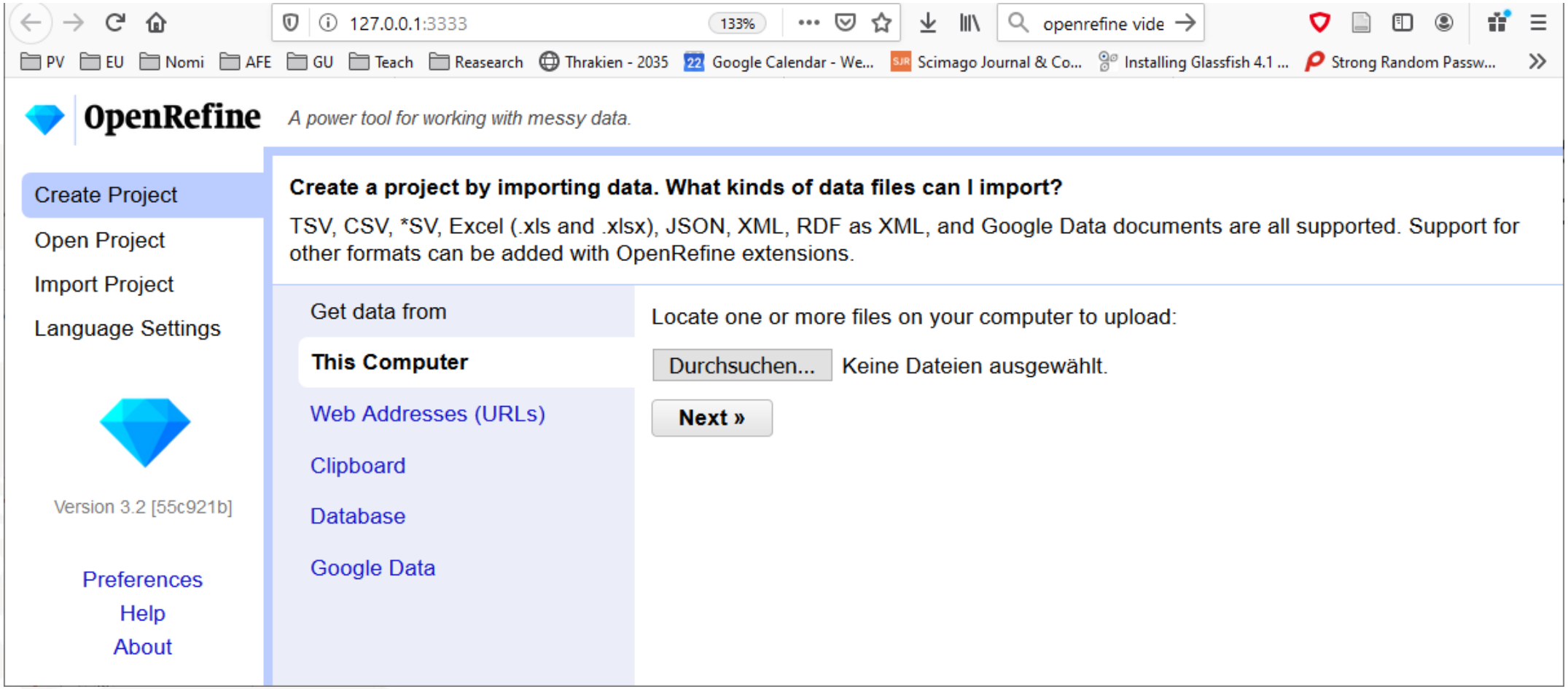
Openrefine …

1. Basics
   Home, Download, Documentation and Training: http://openrefine.org/

2. Reconciliation
   A Reconciliation API returns a ranked list of potential entities matching input labels

http://127.0.0.1:3333/

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.



Reconciliation API

Reconciliation API helps to find matches between your data and existing data

Once the matches are found, you can enrich your data with extra information

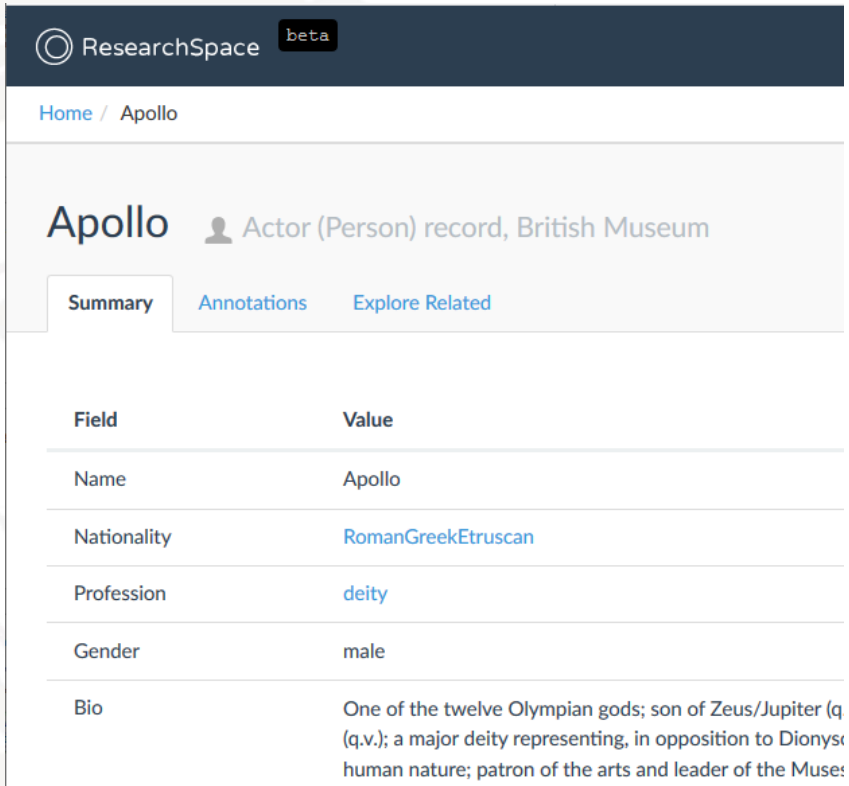
Existing systems supporting the Reconcilation API:

- Wikidata (www.wikidata.org)
- Getty (https://www.getty.edu/research/tools/vocabularies/obtain/openrefine.html)
  https://www.getty.edu/research/tools/vocabularies/obtain/getty_vocabularies_openrefine_tutorial.pdf
  The J. Paul Getty Trust is a cultural and philanthropic institution dedicated to the presentation, conservation, and interpretation of the world's artistic legacy.

Example: Deities from British Museum

Apollo(BM) – ???(Wikidata)

https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/56988



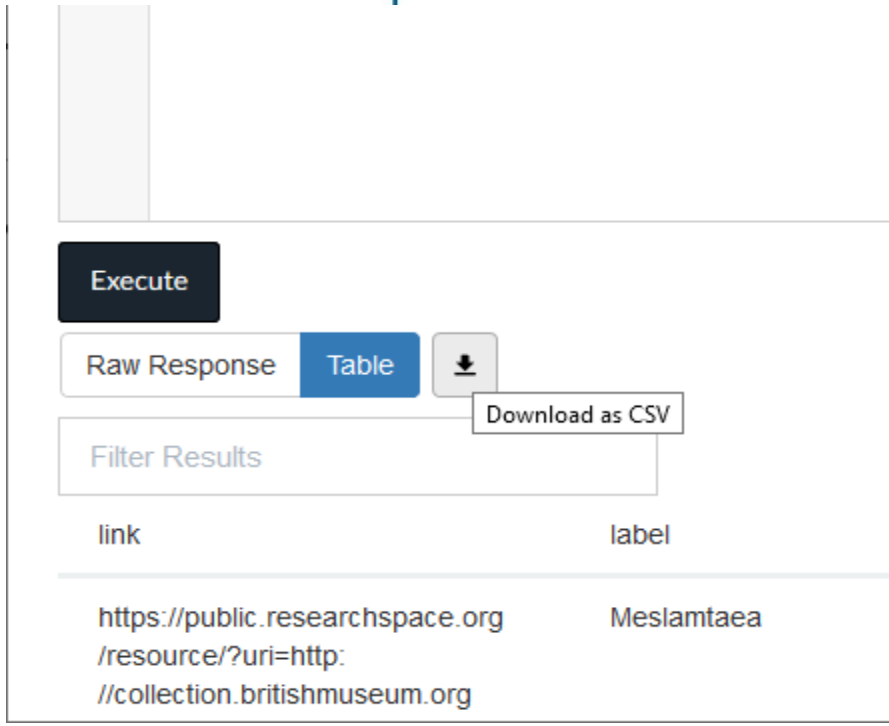
Example: Get all Deities from British Museum

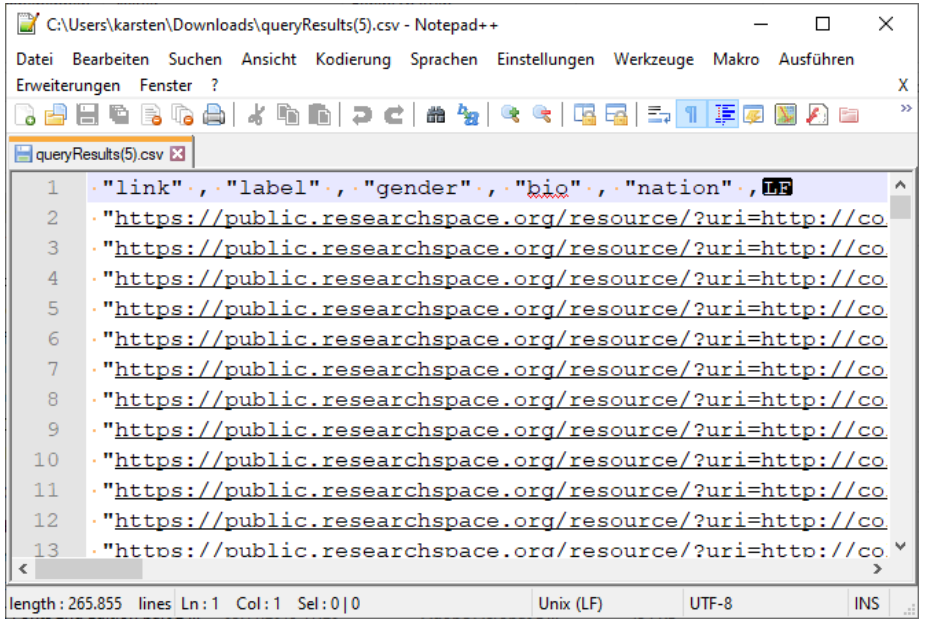
SPARQL endpoint: https://public.researchspace.org/sparql

```
SELECT ?link ?label ?gender ?bio ?nation
WHERE {
  ?l <http://www.cidoc-crm.org/cidoc-crm/P3_has_note> ?bio.
  ?l <http://www.researchspace.org/ontology/displayLabel> ?label .
  ?l <http://www.researchspace.org/ontology/PX_profession>
<http://collection.britishmuseum.org/id/thesauri/profession/deity>.
  ?l <http://www.researchspace.org/ontology/PX_nationality> ?n.
  ?l <http://www.researchspace.org/ontology/PX_gender> ?g.
  BIND(CONCAT("https://public.researchspace.org/resource/?uri=", STR(?l)) AS ?link).
  BIND(STRAFTER(STR(?g), "gender/") AS ?gender) .
  FILTER((LANG(?label)) = "en") .
  BIND(STRAFTER(STR(?n), "nationality/") AS ?nation) .
}
```

Hands on …

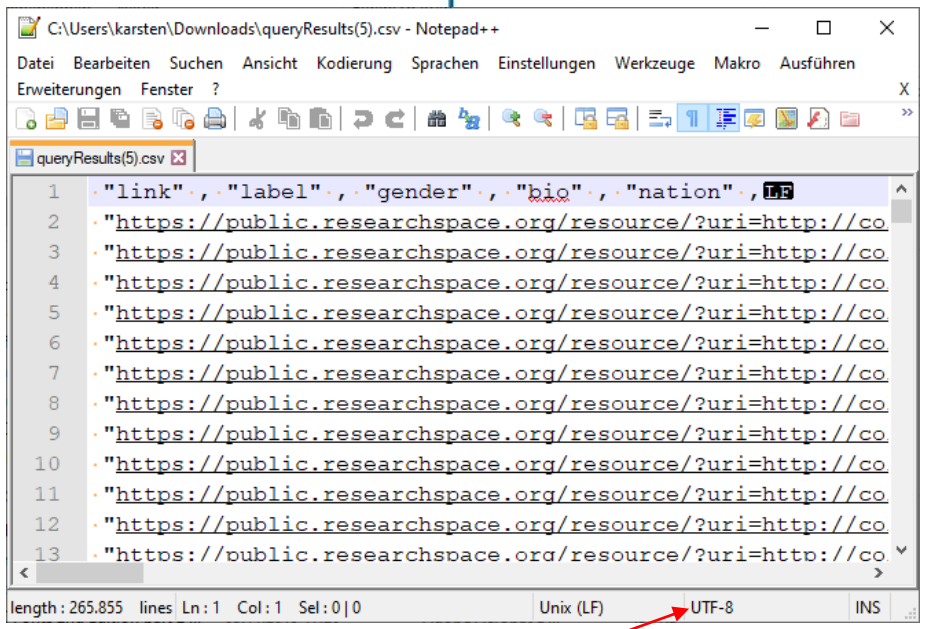
1. Execute the statement …
2. You might need to click on „Table“ in order to download the result as CSV.
3. Save CSV on your disc:

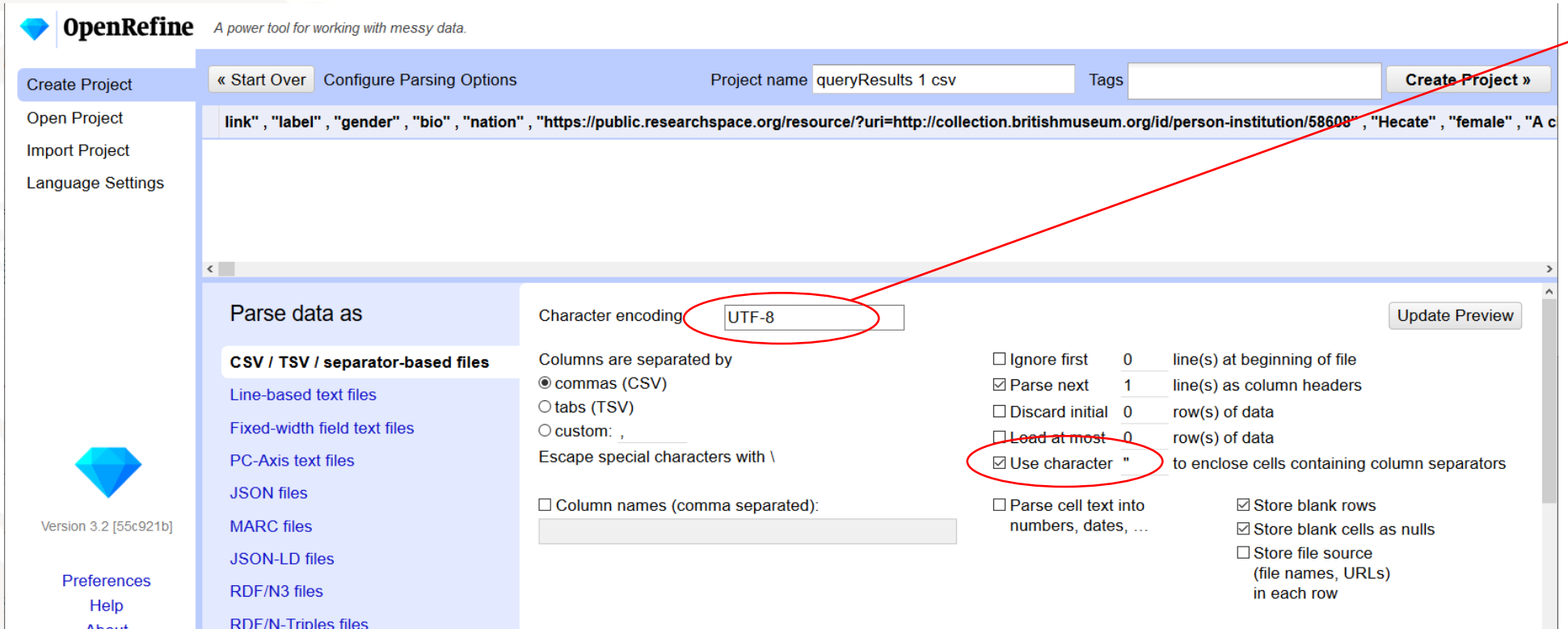




Problems … use Notepad++ (https://notepad-plus-plus.org/downloads/)

1. Since each cell is enclosed by "
   "link" , "label" , "gender" ,…
2. Make sure the encoding fits!

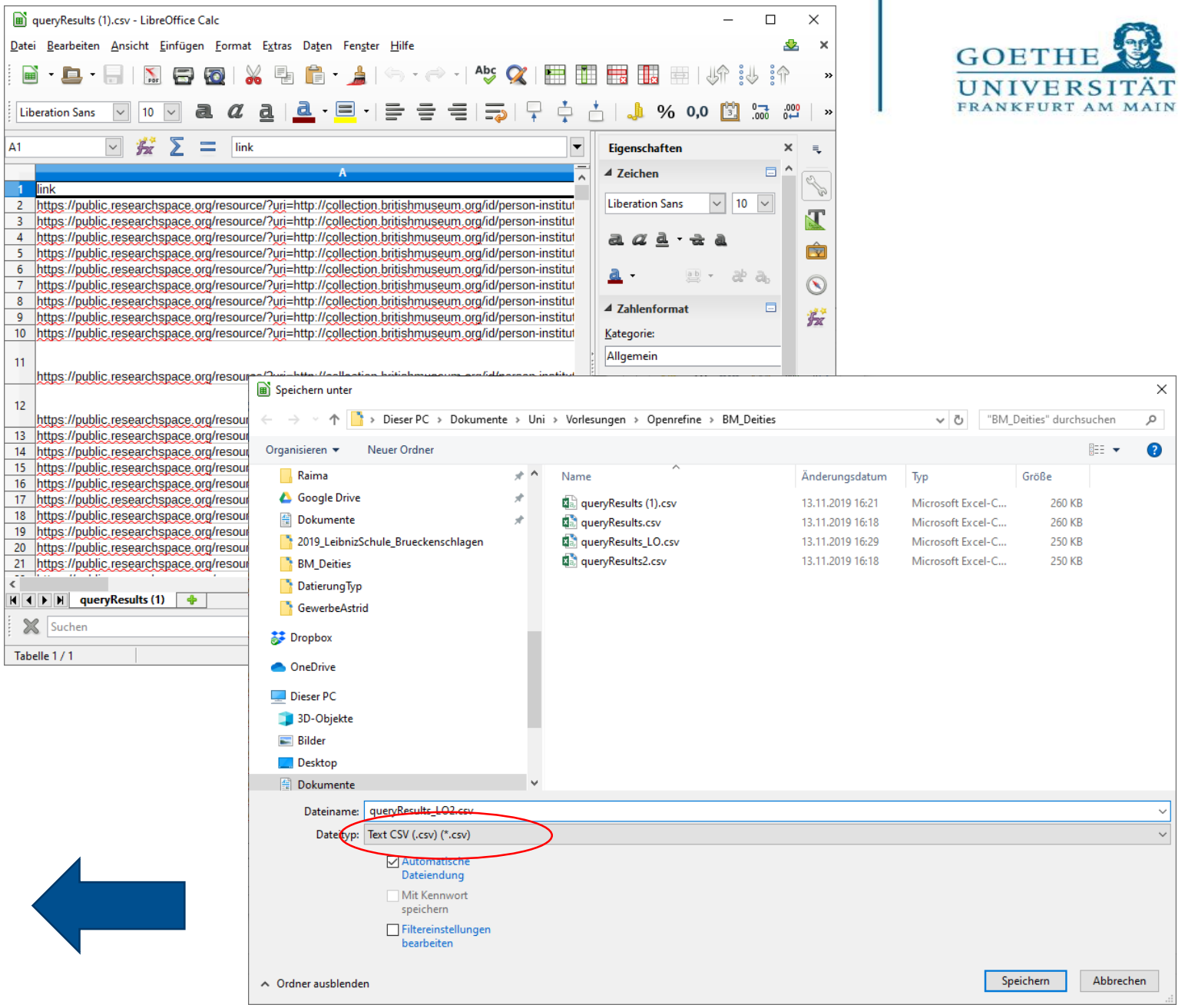


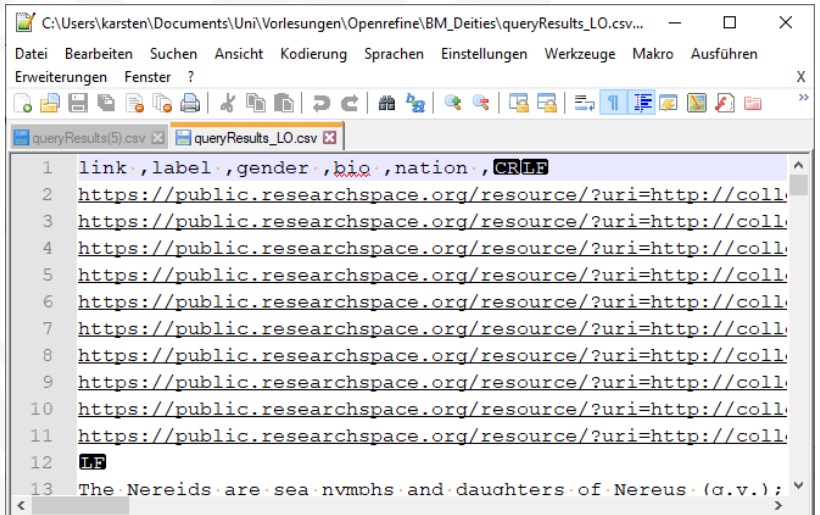


Solution …

Use **LibreOffice Calc** to
open CSV file
(https://www.libreoffice.org/download/download/)

Save it as Text CSV





Create project

Select some values

via text facet ...

OpenRefine queryResults_LO csv Permalink

Facet / Filter | Undo / Redo 0 / 0

Refresh | Reset All | Remove All

label change invert reset

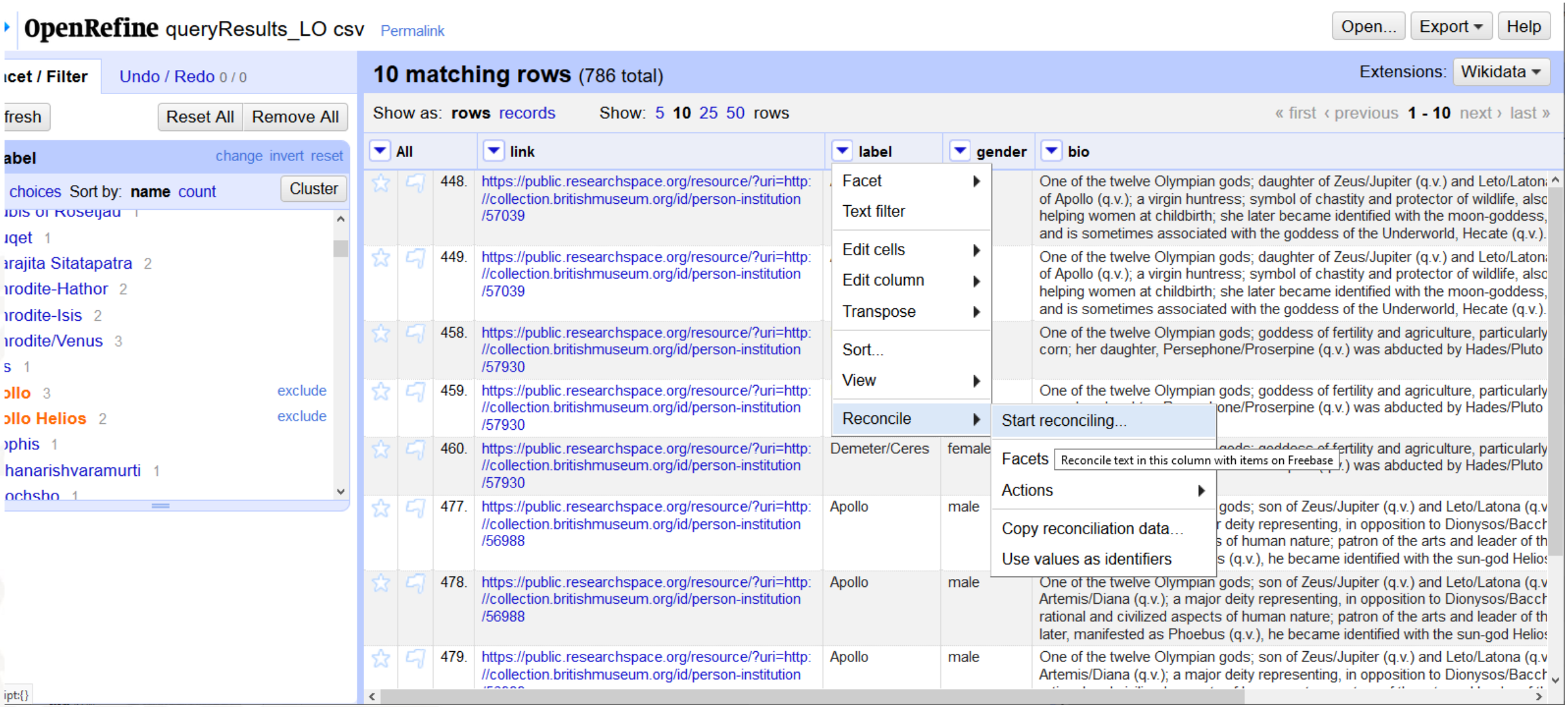
621 choices Sort by: **name** count | Cluster

Anubis of Roseljau 1
Anuqet 1
Aparajita Sitatapatra 2
Aphrodite-Hathor 2
Aphrodite-Isis 2
Aphrodite/Venus 3
Apis 1
Apollo 3 exclude
Apollo Helios 2 exclude
Apophis 1
Ardhanarishvaramurti 1
Ardochsho 1

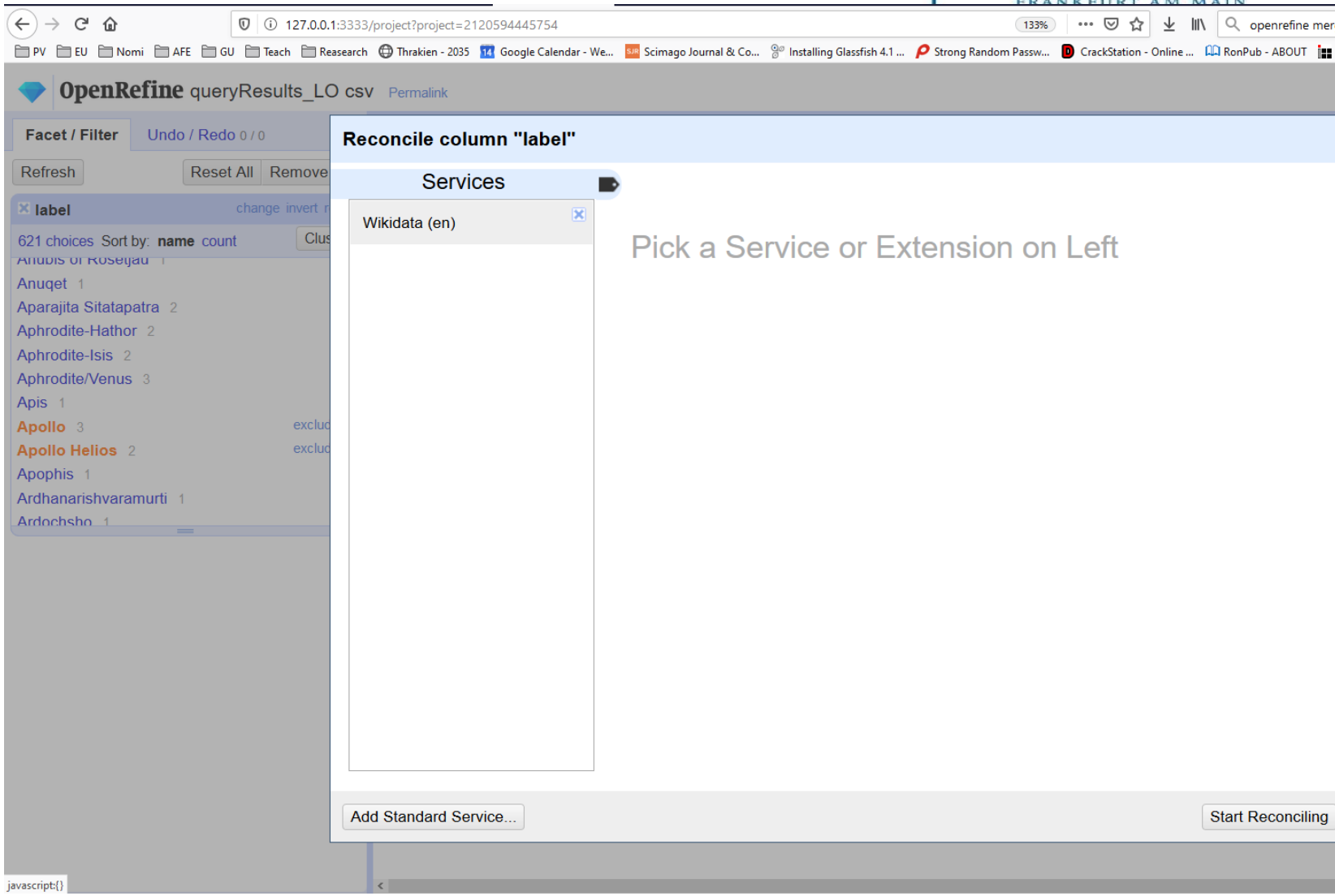
10 matching rows (786 total)

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

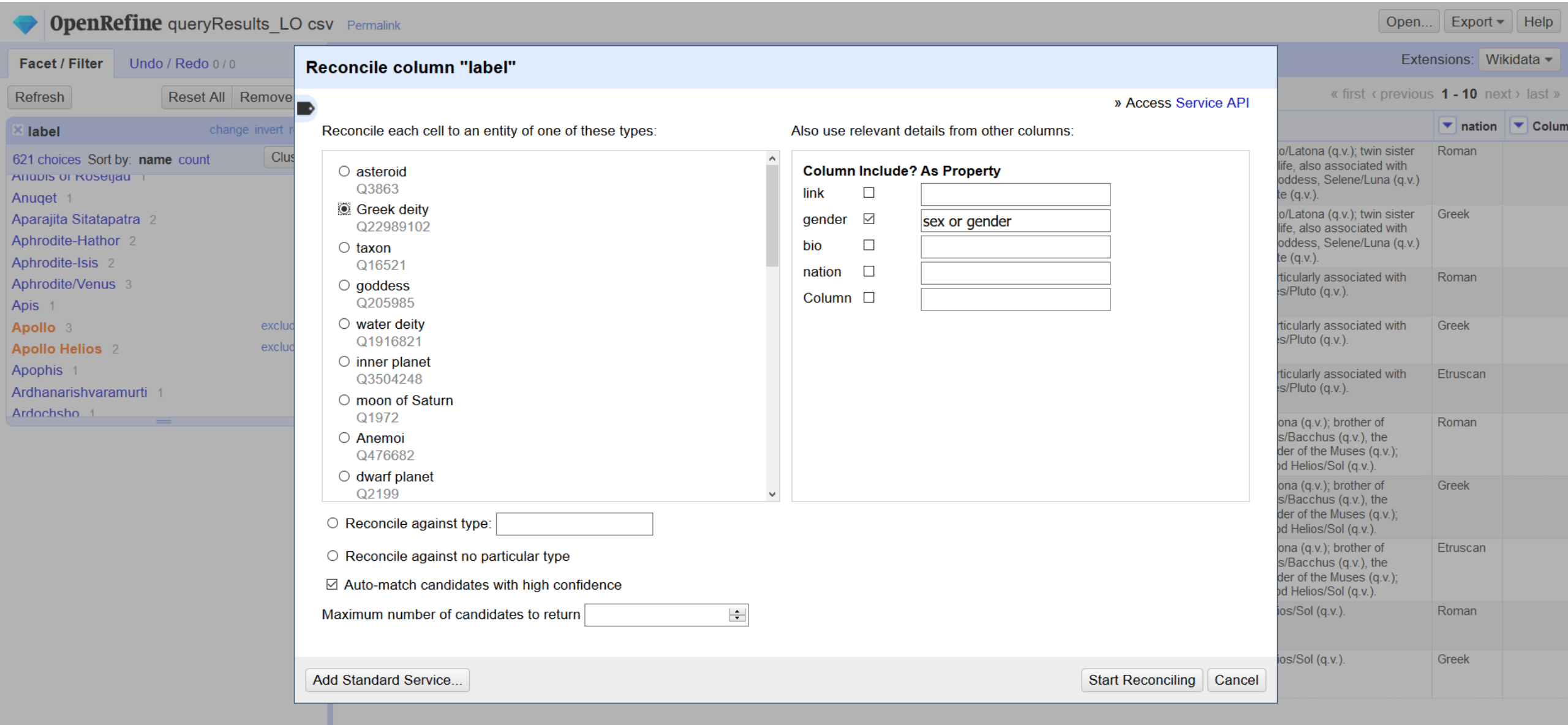
All			link	label	gender	bio
		448.	https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/57039	Artemis/Diana	female	One of the tw of Apollo (q.v.) helping wome and is someti
		449.	https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/57039	Artemis/Diana	female	One of the tw of Apollo (q.v.) helping wome and is someti
		458.	https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/57930	Demeter/Ceres	female	One of the tw corn; her daug
		459.	https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/57930	Demeter/Ceres	female	One of the tw corn; her daug
		460.	https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/57930	Demeter/Ceres	female	One of the tw corn; her daug
		477.	https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/56988	Apollo	male	One of the tw Artemis/Diana rational and c later, manifes
		478.	https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/56988	Apollo	male	One of the tw Artemis/Diana rational and c later, manifes
		479.	https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/56988	Apollo	male	One of the tw Artemis/Diana rational and c later, manifes
		491.	https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/56991	Apollo Helios	male	Syncretic deit
		492.	https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/56991	Apollo Helios	male	Syncretic deit


OpenRefine queryResults_LO csv
Permalink
Open...
Export
Help
cet / Filter
Undo / Redo 0 / 0
fresh
Reset All
Remove All
abel
change invert reset
choices Sort by: name count
Cluster
arajita Sitatapatra 2
rodite-Hathor 2
rodite-Isis 2
rodite/Venus 3
s 1
ollo 3
exclude
ollo Helios 2
exclude
phis 1
hanarishvaramurti 1
10 matching rows (786 total)
Extensions: Wikidata
Show as: rows records
Show: 5 10 25 50 rows
« first ‹ previous 1 - 10 next › last »
All
link
label
gender
bio
448. https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/57039
449. https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/57039
458. https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/57930
459. https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/57930
460. https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/57930
Demeter/Ceres
female
477. https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/56988
Apollo
male
478. https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/56988
Apollo
male
479. https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution
Apollo
male
Facet
Text filter
Edit cells
Edit column
Transpose
Sort...
View
Reconcile
Start reconciling...
Facets
Reconcile text in this column with items on Freebase
Actions
Copy reconciliation data...
Use values as identifiers

Pick a Service
(only Wikidata predefined)

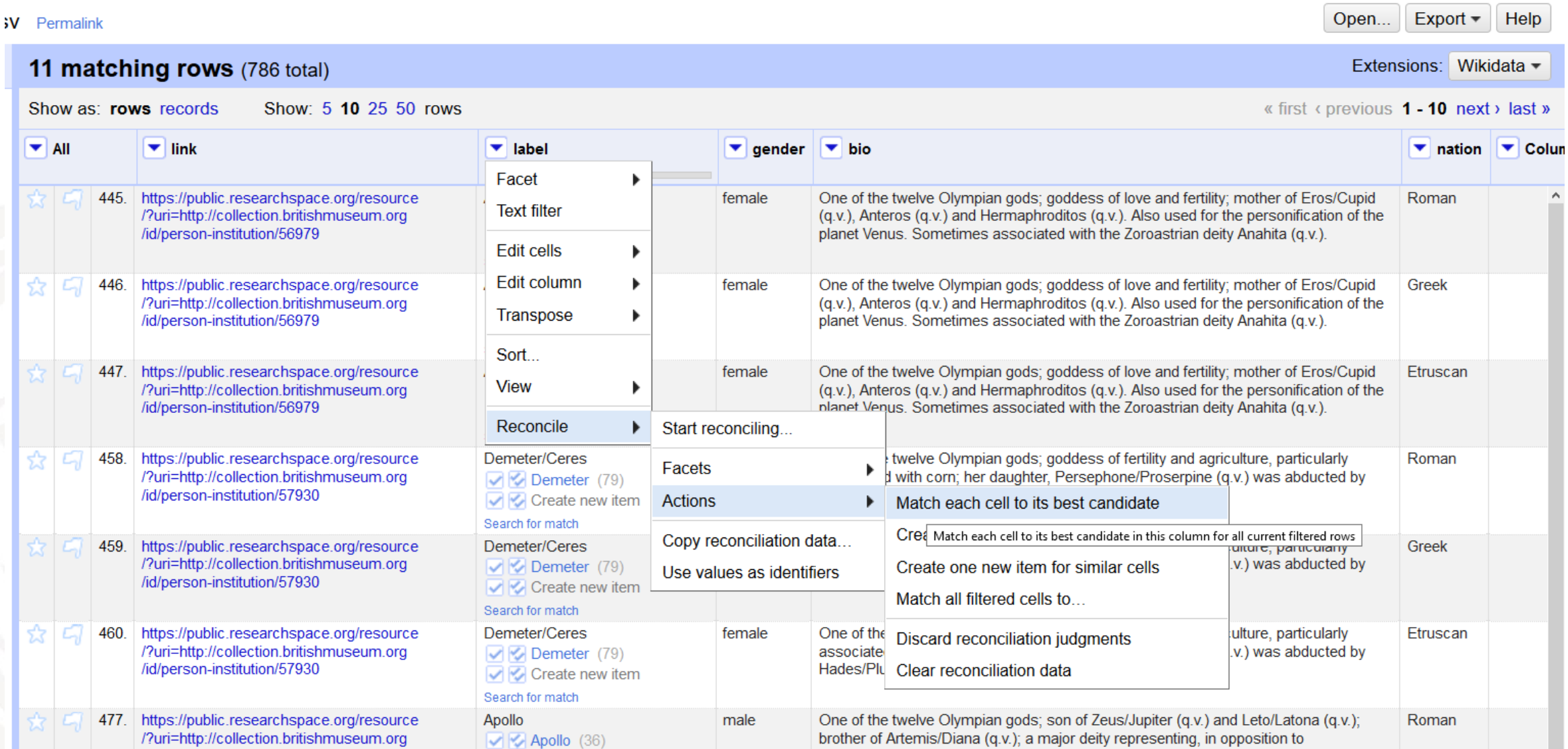


(left) select best proposed fitting classes (or enter it under „Reconcile against type“)
(right) add additional columns of „your“ data that are entered into the reconciliation process

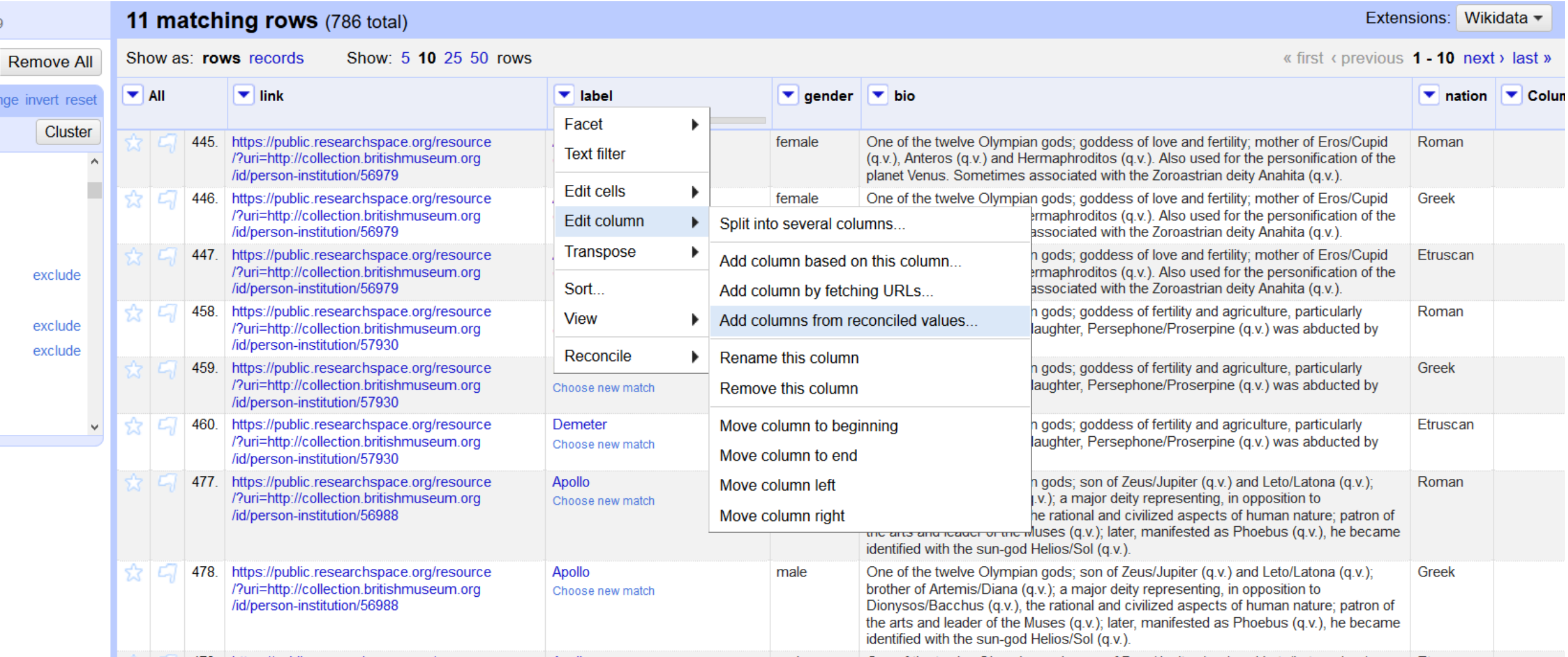


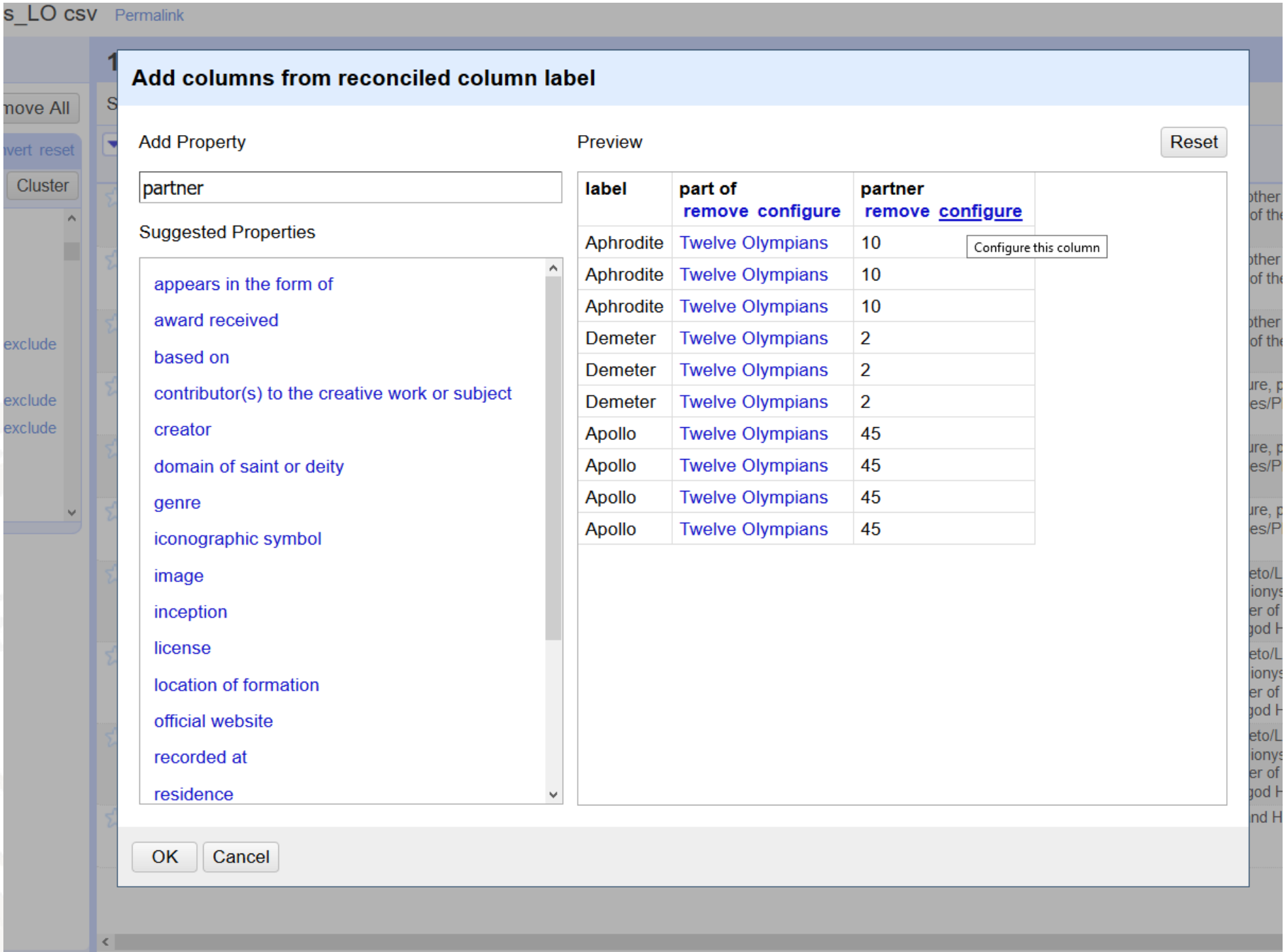
Match manually or filter and match to best candidate …

… also possible to clear reconciliation data and start from scratch

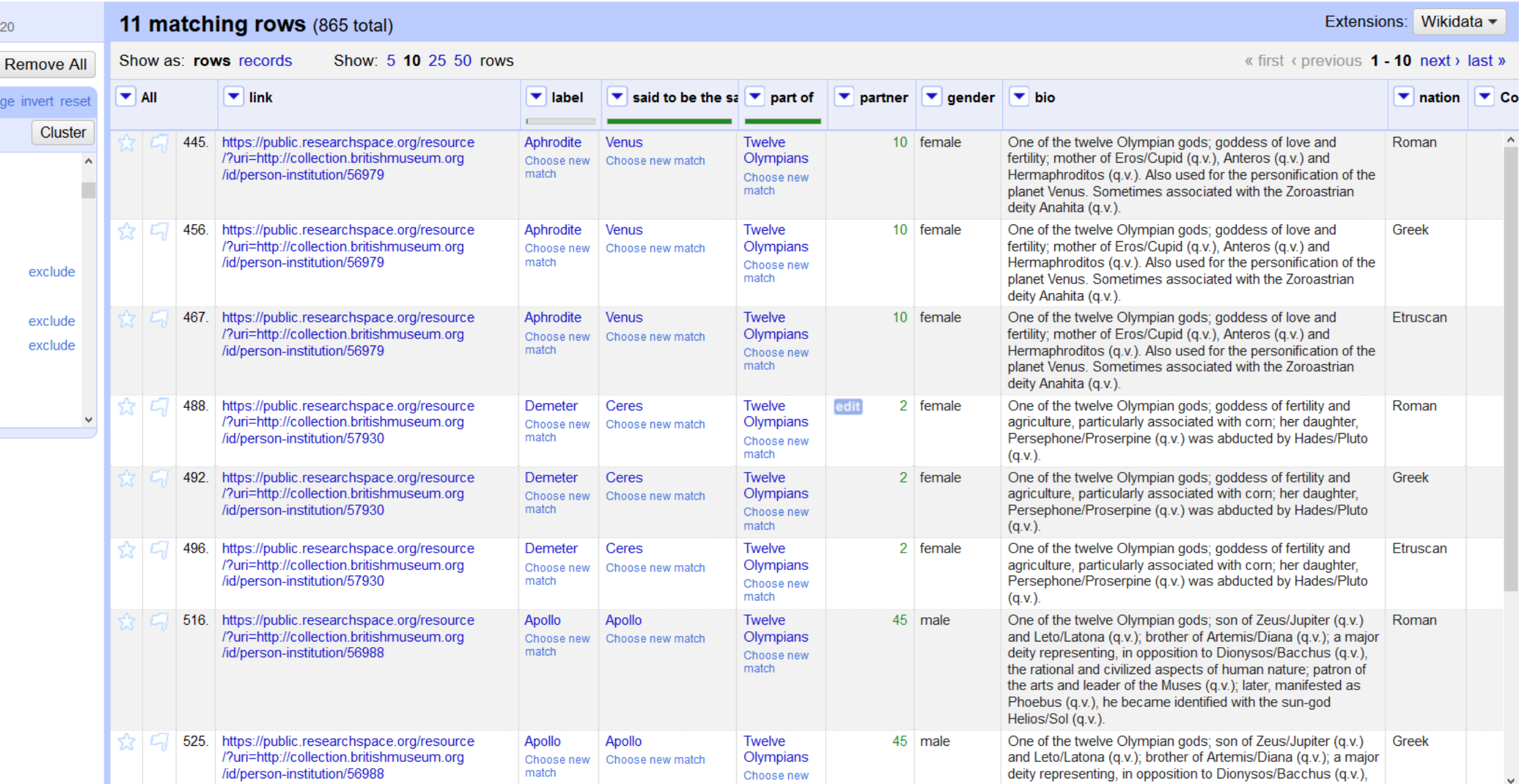


Enrich your data …




s_LO csv
Permalink
move All
vert reset
Cluster
exclude
exclude
exclude
Add columns from reconciled column label
Add Property
partner
Suggested Properties
appears in the form of
award received
based on
contributor(s) to the creative work or subject
creator
domain of saint or deity
genre
iconographic symbol
image
inception
license
location of formation
official website
recorded at
residence
Preview
Reset
label
part of
remove configure
partner
remove configure
Configure this column
Aphrodite Twelve Olympians 10
Aphrodite Twelve Olympians 10
Aphrodite Twelve Olympians 10
Demeter Twelve Olympians 2
Demeter Twelve Olympians 2
Demeter Twelve Olympians 2
Apollo Twelve Olympians 45
Apollo Twelve Olympians 45
Apollo Twelve Olympians 45
Apollo Twelve Olympians 45
OK
Cancel

… your enriched data …



20

Remove All

ge invert reset

Cluster

exclude

exclude

exclude

11 matching rows (865 total)

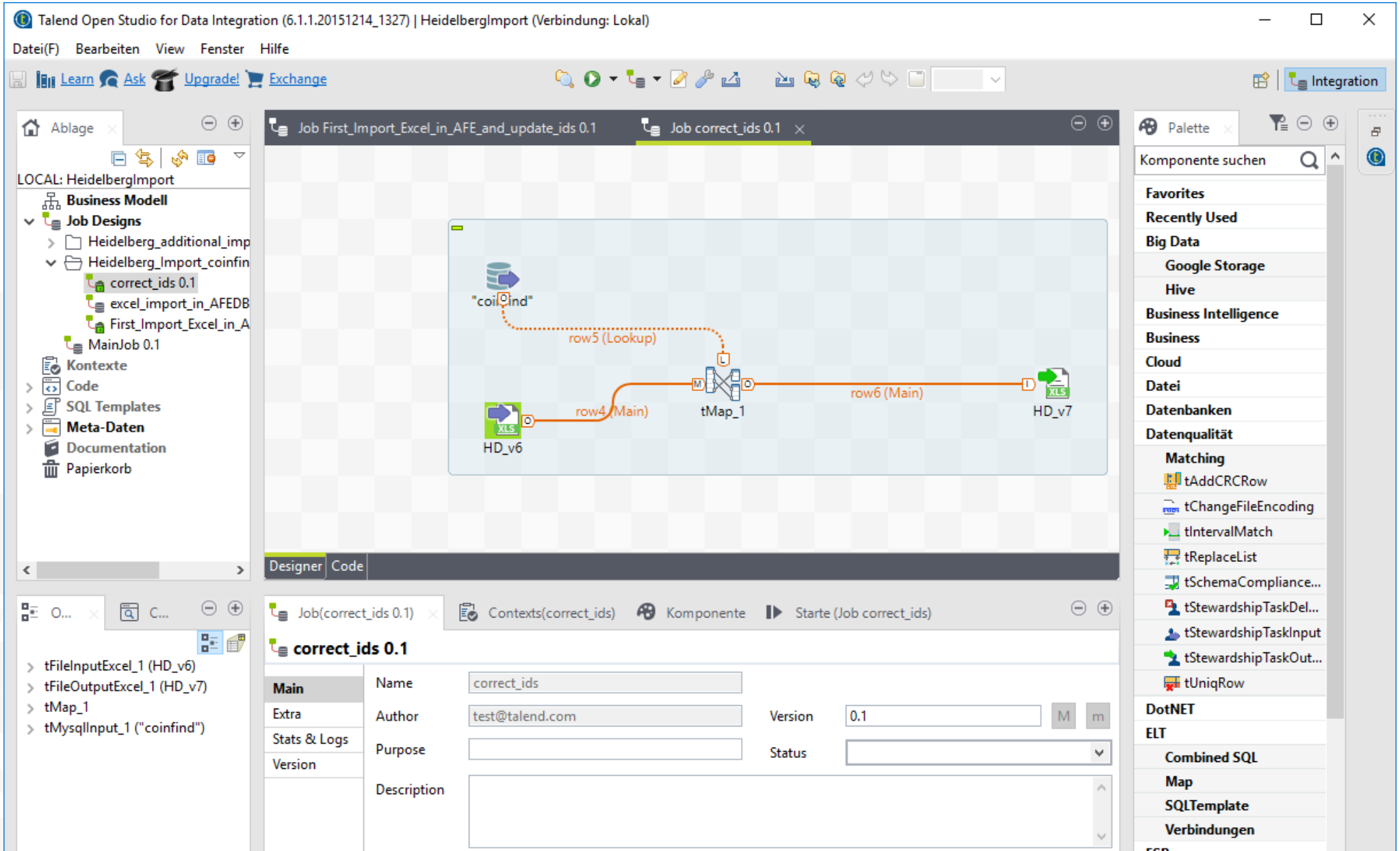
Extensions: Wikidata

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

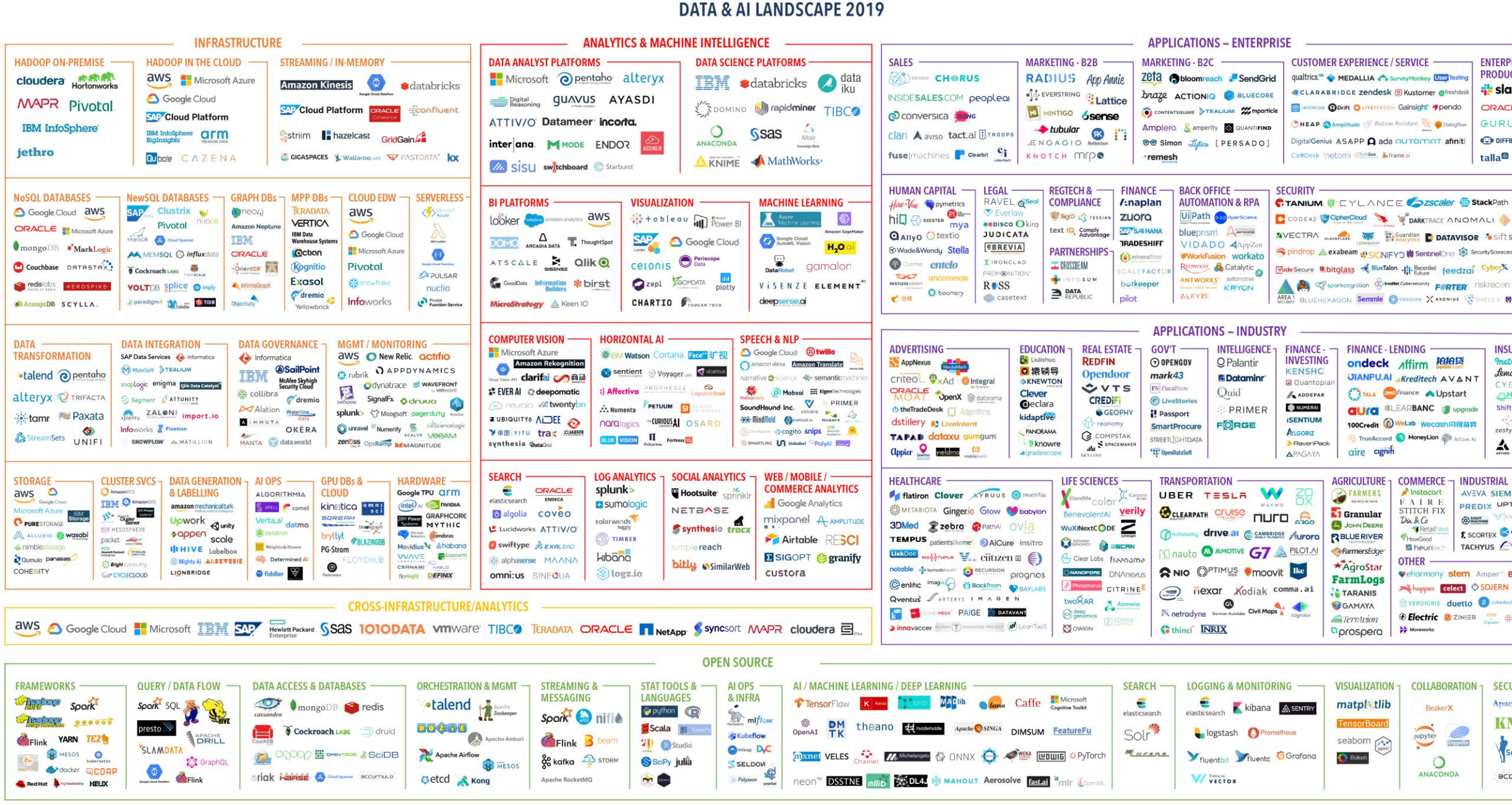
« first ‹ previous **1 - 10** next › last »

All		link	label	said to be the sa	part of	partner	gender	bio	nation	Co
445.	https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/56979	Aphrodite Choose new match	Venus Choose new match	Twelve Olympians Choose new match	10	female	One of the twelve Olympian gods; goddess of love and fertility; mother of Eros/Cupid (q.v.), Anteros (q.v.) and Hermaphroditos (q.v.). Also used for the personification of the planet Venus. Sometimes associated with the Zoroastrian deity Anahita (q.v.).	Roman	
456.	https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/56979	Aphrodite Choose new match	Venus Choose new match	Twelve Olympians Choose new match	10	female	One of the twelve Olympian gods; goddess of love and fertility; mother of Eros/Cupid (q.v.), Anteros (q.v.) and Hermaphroditos (q.v.). Also used for the personification of the planet Venus. Sometimes associated with the Zoroastrian deity Anahita (q.v.).	Greek	
467.	https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/56979	Aphrodite Choose new match	Venus Choose new match	Twelve Olympians Choose new match	10	female	One of the twelve Olympian gods; goddess of love and fertility; mother of Eros/Cupid (q.v.), Anteros (q.v.) and Hermaphroditos (q.v.). Also used for the personification of the planet Venus. Sometimes associated with the Zoroastrian deity Anahita (q.v.).	Etruscan	
488.	https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/57930	Demeter Choose new match	Ceres Choose new match	Twelve Olympians Choose new match	edit 2	female	One of the twelve Olympian gods; goddess of fertility and agriculture, particularly associated with corn; her daughter, Persephone/Proserpine (q.v.) was abducted by Hades/Pluto (q.v.).	Roman	
492.	https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/57930	Demeter Choose new match	Ceres Choose new match	Twelve Olympians Choose new match	2	female	One of the twelve Olympian gods; goddess of fertility and agriculture, particularly associated with corn; her daughter, Persephone/Proserpine (q.v.) was abducted by Hades/Pluto (q.v.).	Greek	
496.	https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/57930	Demeter Choose new match	Ceres Choose new match	Twelve Olympians Choose new match	2	female	One of the twelve Olympian gods; goddess of fertility and agriculture, particularly associated with corn; her daughter, Persephone/Proserpine (q.v.) was abducted by Hades/Pluto (q.v.).	Etruscan	
516.	https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/56988	Apollo Choose new match	Apollo Choose new match	Twelve Olympians Choose new match	45	male	One of the twelve Olympian gods; son of Zeus/Jupiter (q.v.) and Leto/Latona (q.v.); brother of Artemis/Diana (q.v.); a major deity representing, in opposition to Dionysos/Bacchus (q.v.), the rational and civilized aspects of human nature; patron of the arts and leader of the Muses (q.v.); later, manifested as Phoebus (q.v.), he became identified with the sun-god Helios/Sol (q.v.).	Roman	
525.	https://public.researchspace.org/resource/?uri=http://collection.britishmuseum.org/id/person-institution/56988	Apollo Choose new match	Apollo Choose new match	Twelve Olympians Choose new	45	male	One of the twelve Olympian gods; son of Zeus/Jupiter (q.v.) and Leto/Latona (q.v.); brother of Artemis/Diana (q.v.); a major deity representing, in opposition to Dionysos/Bacchus (q.v.),	Greek	

Talend Open Studio - https://de.talend.com/products/talend-open-studio/



Tool landscape - https://mattturck.com/data2019/



Tools

- Fuseki (RDF-Datenbank)
- D2R (Mapping zwischen relationaler DB und RDF)
- OpenRefine (Umgang mit verschiedenen Formaten, auch für große Dokumente, Reconciliation)

Weitere …

- Notepad ++ (generischer Texteditor)
- PilotEdit Lite (wenn die Dateien zu groß werden – free bis 10GB)
- Libre Office (gut zum Einlesen von CSV-Dateien)

52 52

Hausaufgabe (bis nächste Woche):

1. Team – Steckbrief bis diesen Freitag eintragen!

2. Lesen:
   https://de.wikipedia.org/wiki/Beurteilung_eines_bin%C3%A4ren_Klassifikators
   bzw. (möglichst UND)
   https://en.wikipedia.org/wiki/Evaluation_of_binary_classifiers
   – Sie sollten die folgenden Begriffe erklären können:
     – Recall
     – Precision
     – Sensitivity
     – Specificity
     – F-Maß/F-Score

53 53

Hausaufgabe (sobald es die Zeit erlaubt) ….

Ethical Implications of AI

http://www.bigdata.uni-frankfurt.de/data-challenge-ss-2020/

Prof. Gemma Roig http://www.cvai.cs.uni-frankfurt.de/

Fairness, Bias and Discrimination in AI (Prof. Gemma Roig)

[slides http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2020/01/Fairness_Bias_Discrimination_in_AI_GemmaRoig.pdf]
[video https://www.youtube.com/watch?v=LE38HWZz8NU]