Data opschonen in Adlib

Collegagroep Digitale collectieregistratie
25 juni 2020

./DATABLE

./DATABLE

DATABLE ondersteunt sinds 2010 erfgoedorganisaties
bij het beheren en optimaliseren van data en dataprocessen
door strategisch advies, onderzoek en dienstverlening.

digipolis

europeana fashion
INTERNATIONAL ASSOCIATION

MM MODEMUSEUM

RIJKS MUSEUM

DEPARTEMENT CULTUUR, JEUGD & MEDIA
Vlaanderen

KUNSTEN PUNT

vanabbemuseum

erf goed Nederland

Provincie Antwerpen

Universiteit Antwerpen

– Museumstichting –

museum Hof van Busleyden
Bourgondisch stadspaleis Mechelen

Het Nieuwe Instituut
architectuur design e-cultuur

Vlaams Architectuur instituut

ETWIE

GEMEENTE MUSEUM DEN HAAG

Ministerie van Onderwijs, Cultuur en Wetenschap

PACKED
Expertisecentrum Digitaal Erfgoed

RKD NEDERLANDS INSTITUUT VOOR KUNSTGESCHIEDENIS

Van Gogh Museum
Amsterdam

DIVA
ANTWERP HOME OF DIAMONDS

Stad KORTRIJK

FOMU foto museum

DEPARTEMENT
CULTUUR,
JEUGD & MEDIA

Handleiding data opschonen in Adlib

Stappenplan voor data-opschoning in Adlib software

./DATABLE

agenda

1. Datakwaliteit
2. Dataplanning
3. Gegevensvalidatie in Adlib
4. Analyse
5. Opschonen in Adlib
6. Reconciliëren
7. Automatiseren

1. DATAKWALITEIT

datakwaliteit: dimensies

1. Consistentie
2. Validiteit
3. Uniciteit
4. Volledigheid
5. Accuraatheid
6. Actualiteit

datakwaliteit: dimensies

1. Consistentie
2. Validiteit
3. Uniciteit
4. Volledigheid
5. Actualiteit
6. Accuraatheid

inconsistenties

grammaticale varianten	`bas-reliëf`**`s`**
synoniemen	`halfverheven` `beeldhouwwerk`
specifiek/generiek	`beeldhouwwerk`
hoofdletters	**`B`**`as-reliëf`

ongeldig

tikfouten	`bas-relif`
witruimte	`_bas-reliëf_`
leestekens	`bas-reliëf?`
spellingsvarianten	`bas reliëf`
meervoudige term	`reliëf en kader`
type error	`brons`

validiteit

- intern regelwerk en invoerlijsten
- Invulboek: projectcest.be/wiki/invulboek
- thesauri:
  - Art & Architecture Thesaurus
  - RKDartists, ULAN
  - Geonames

Welkom op de vernieuwde versie van Project CEST! Eventuele pro

Term objectnaam

Veld	**term objectnaam**
Definitie	Het veld 'term objectnaam' documenteert de meest specifieke benaming van het object.
Invulregels	Noteer, indien mogelijk, met één term om welk soort object het gaat. wereldbol affiche sieraad fiets wiel scherfgranaat hals Een objectnaam wordt meestal vanuit een functionele benadering aan het object toegekend. Zo is een "spaarpot" een objectnaam voor een spaarvarken en niet "varken", wat een Afgebeeld concept voor h Analoog hieraan is "beeld" een objectnaam voor een beeldje van een clown en niet "clown", aangezien het niet Aan een baal wol geef je "baal" als objectnaam en niet "wol", want dat is het materiaal van het object. Gebruik een term ontleend aan een publiek toegankelijke bron of standaard, zoals bv. AAT. Relaties tussen aanverwante termen kunnen in de thesaurus worden gelegd.

23:41 06/24/2020 24°C 100%
Not secure | getty.edu/vow/AATFullDisplay?find=lithogra*&logic=AND&note=&english=N&prev_page=1&subjectid=300041379

Click the icon to view the hierarchy.

Semantic View (JSON, RDF, N3/Turtle, N-Triples)

Representative Images: 1

ID: 300041379
Page Link: http://vocab.getty.edu/page/aat/300041379

Record Type: concept

lithographs (planographic prints, <prints by process: transfer method>, ... Visual and Verbal Communication (hierarchy name))

Note: Prints made using the process of lithography, which is a planographic printing process in which a design is deposited on the stone or plate with a greasy substance which will accept ink.

Terms:

lithographs (**preferred**,C,U,English-P,D,U,PN)
lithograph (C,U,English,AD,U,SN)
平版印刷版畫 (C,U,Chinese (traditional)-P,D,U,U)
píng bǎn yìn shuā bǎn huà (C,U,Chinese (transliterated Hanyu Pinyin)-P,UF,U,U)
ping ban yin shua ban hua (C,U,Chinese (transliterated Pinyin without tones)-P,UF,U,U)
p'ing pan yin shua pan hua (C,U,Chinese (transliterated Wade-Giles)-P,UF,U,U)
lithografieën (prenten) (C,U,Dutch-P,D,U,U)
lithografie (prent) (C,U,Dutch,AD,U,SN)
litho (C,U,Dutch,UF,U,SN)
steendruk (C,U,Dutch,UF,U,SN)
litho's (C,U,Dutch,UF,U,U)
steendrukken (C,U,Dutch,UF,U,U)

2. DATAPLANNING

planning

- Wat wil ik bereiken?
  - vindbaarheid
  - uitwisselbaarheid
  - data geschikt voor publicatie
  - georeferencing
- Hoeveel tijd heb ik?
- Wat zijn prioriteiten?

dataplan

- Welke gegevens?
  - objectnaam
  - locaties
- Welke acties?
  - opschonen
  - normaliseren volgens standaard
  - vervolledigen
- Termijn en prioriteiten?

3. GEGEVENSVALIDATIE IN ADLIB

Adlib Bibliotheek+Museum 4.3

Bestand | Start | Wijzigen | Beeld | Publiceren

Terug | Herstart | Eerste | Vorig | Volgend | Laatste | Wijzigen | Nieuw

Terug | Navigatie

Identificatie | Vervaardiging | Datering | Fysieke kenm

01015 beeldhouwwerk

Sint-Michiel verslaat d

Vervaardig

Rec

Datering

Record 20 van 204 | Recordnummer 550007559

Zoomscherm

Zoomscherm Thesaurus

Term

Term	beeldhouwwerk
Soort term	objectcategorie
	objectnaam
	onderwerp
Termcode	
Status	descriptor

Zoomscherm

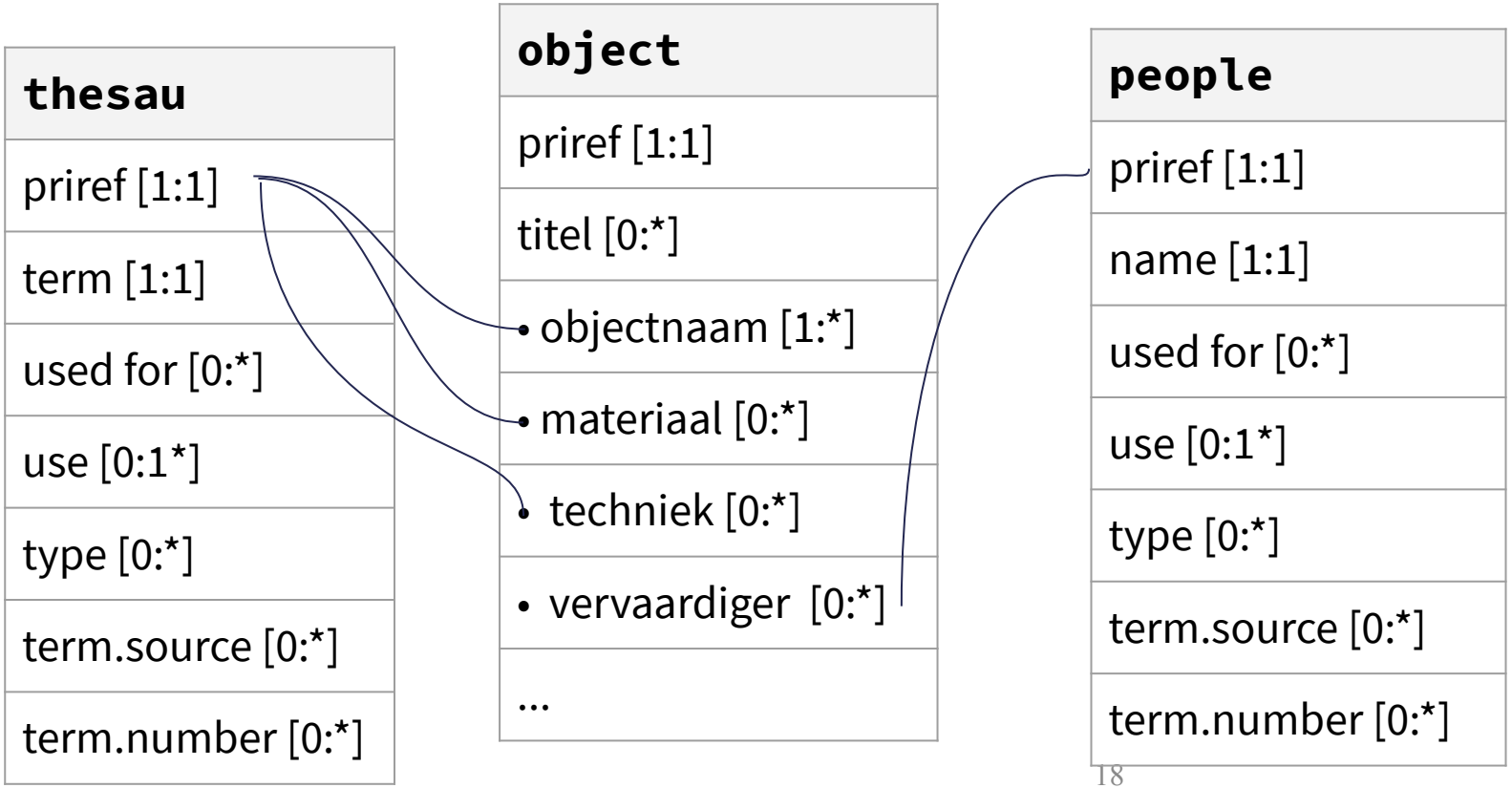
Zoomscherm Persoon/Instelling

Naam

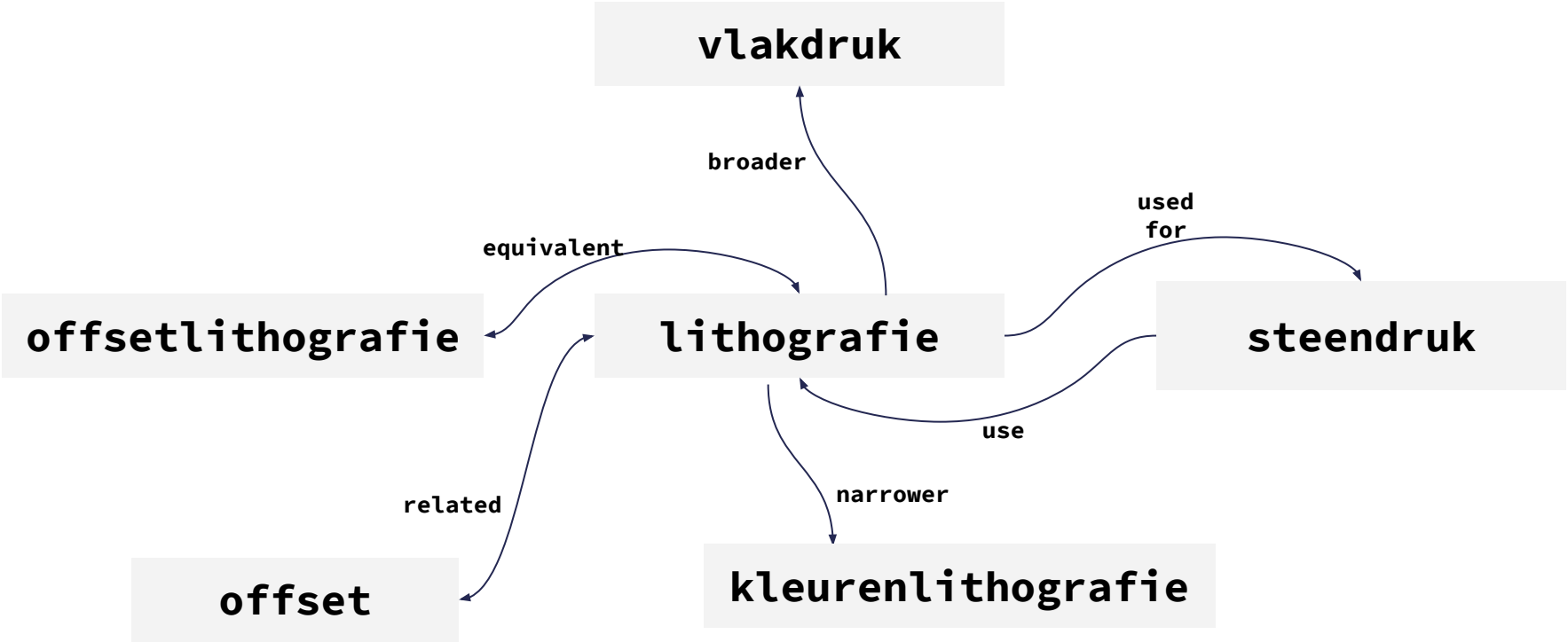
Naam	Bate, Judouis		
Soort naam	vervaardiger		
Referentienummer		Instellingscode	
Status	descriptor		
Opmerking naam			

Relaties met andere namen

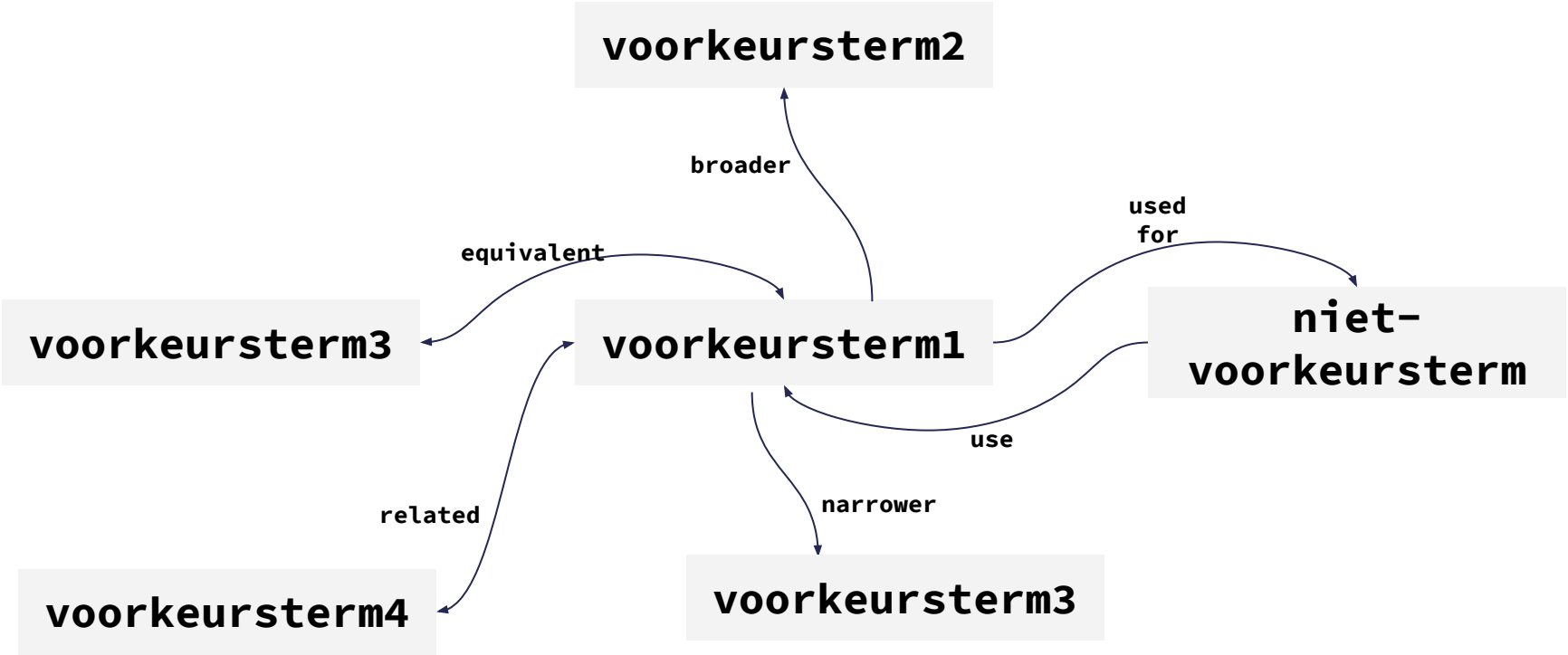
relaties in Adlib



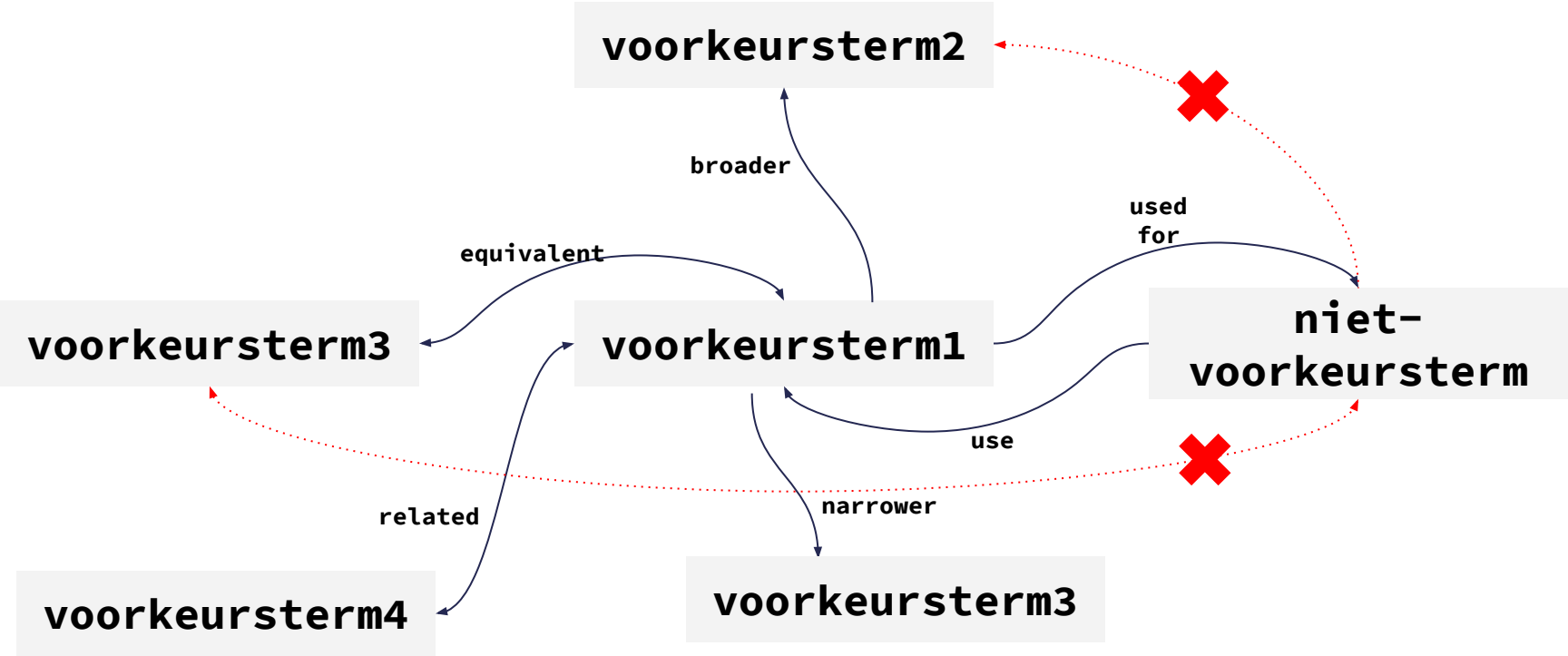
thesaurus: relaties



thesaurus: relaties

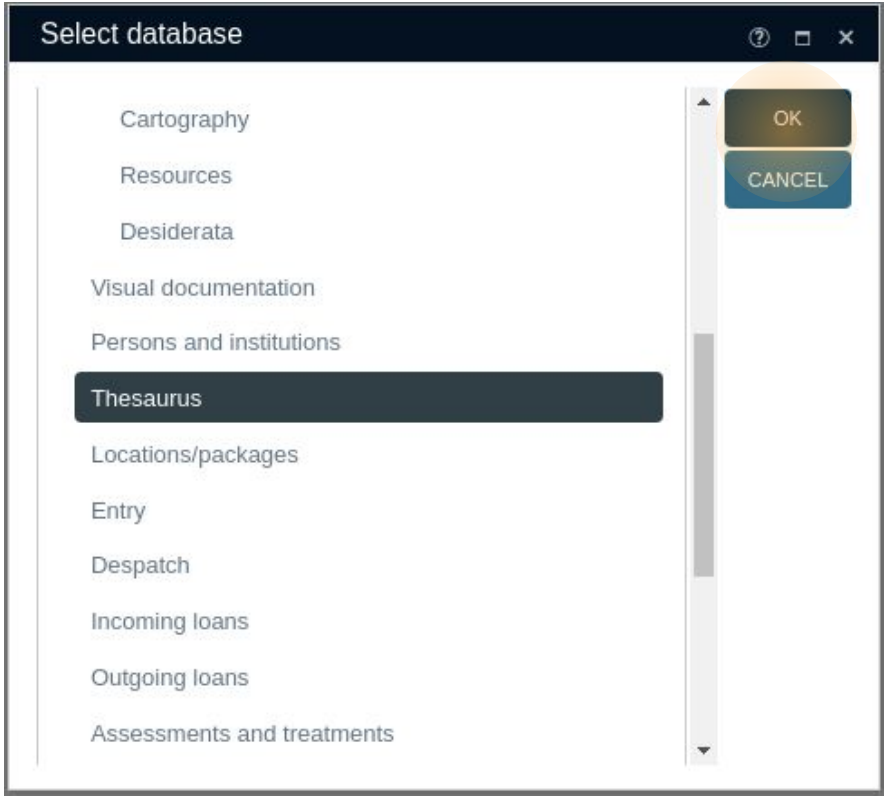


thesaurus: relaties



4. ANALYSE

export via Axiell Collections



13:12 06/12/2020 23°C 100%

cjm-web.adlibhosting.com/Collections/

axiell

Thesaurus

New

Search

Language switch

Help

Account

Search Items Features

Record details

Action Clipboard Occurrence

Features

term	term type	term status
België		
Izegem		
Brugge		
Sint-Eloois-Winkel		
Nederland		
Hasselt		
Damme		
S-schoenen		
schoenmaat		
breedte bal		
hakhoogte		
cm		
fabrikant		
reliëfstempel		
leer		
Pump		
Dames		
Handwerk - handgenaaid		
licht beschadigd		
zeer goed		
Schenking		

3888 ITEMS

Search Thesaurus

STANDARD ADVANCED SAVED SEARCHES

Term	equals	
Term type	equals	
Status	equals	
Input date	equals	
Input name	equals	
Input from	equals	
Edit date	equals	
Edit name	equals	
Edit from	equals	
Record number	equals	

FIND

CANCEL

CLEAR

SETTINGS

Fields

And

Or

Functions

Expand

Num...

For a period of time

Licensed by: - Copyright © 2014-2020 Axiell Group. All rights reserved. - Version: 1.8.1.848 (5/13/2020 11:38:38 AM)

export (2).csv

Show all

13:00 06/12/2020 23°C 100%

cjm-web.adlibhosting.com/Collections/

axiell

- New
- Search
- Language switch
- Help
- Account

Thesaurus

Search | Items | Features

term	term.type	term.status
België	country	candidate
Izegem	place	candidate
Brugge	place	candidate
Sint-Eloois-Winkel	place	candidate
Nederland	country	candidate
Hasselt	place	candidate
Damme	place	candidate
S-schoenen	collection	candidate
schoenmaat	dimension	candidate
breedte bal	dimension	candidate
hakhoogte	dimension	candidate
cm	unit	candidate
fabrikant	role	candidate
reliëfstempel	inscription t…	candidate
leer	material	candidate
Pump	object name	candidate
Dames	object categ…	candidate
Handwerk - handgenaaid	technique	candidate
licht beschadigd	condition	candidate
zeer goed	condition	candidate
Schenking	acquisition …	candidate
buitenzool	inscription p…	candidate
nieuw	condition	candidate

3888 ITEMS

Record details

Action | Clipboard | Occurrence | Features

Thesaurus term

…

Term: België
Term type: country
geographical keyword
place
Term code:
Status: candidate

Relati…

Use:
Used for:
Broader Term:
Narrower Term:
Related Term:
Equivalent Term:

Source and definition

Source: Num…
Scope Note:
Content date:
History note:
Notes:

Text for display in Library OPAC

For a period of time

Period start:

Licensed by: . - Copyright © 2014-2020 Axiell Group. All rights reserved. - Version: 1.8.1.848 (5/13/2020 11:38:38 AM)

13:01 06/12/2020 23°C 100%

cjm-web.adlibhosting.com/Collections/

axiell

- New
- Search
- Language switch
- Help
- Account

Thesaurus

Search | Items | Features

term	term.type	term.status
België	country	candidate
Izegem	place	candidate
Brugge	place	candidate
Sint-Eloois-Winkel		
Nederland		
Hasselt		
Damme		
S-schoenen		
schoenmaat		
breedte bal		
hakhoogte		
cm		
fabrikant		
reliëfstempel		
leer		
Pump		
Dames	object categ…	candidate
Handwerk - handgenaaid	technique	candidate
licht beschadigd	condition	candidate
zeer goed	condition	candidate
Schenking	acquisition …	candidate
buitenzool	inscription p…	candidate
nieuw	condition	candidate

3888 ITEMS

Export

Format: Csv

Columns: Start typing to filter

address
address.place
address.postal_code
broader_term
contact.name
content.date
edit.date
edit.name
edit.notes
edit.source
edit.time
equivalent_term
external_code
external_code.date
external_code.institution
external_code.type
field_coll_place.notes
grid.X_reference
grid.Y_reference

<< REMOVE
ADD >>
MOVE UP
MOVE DOWN

Start typing to filter

term
term.type
term.status
used_for
use
source
priref
term.number

OK
CANCEL
SETTINGS

- All
- Marked
- Current

Record details

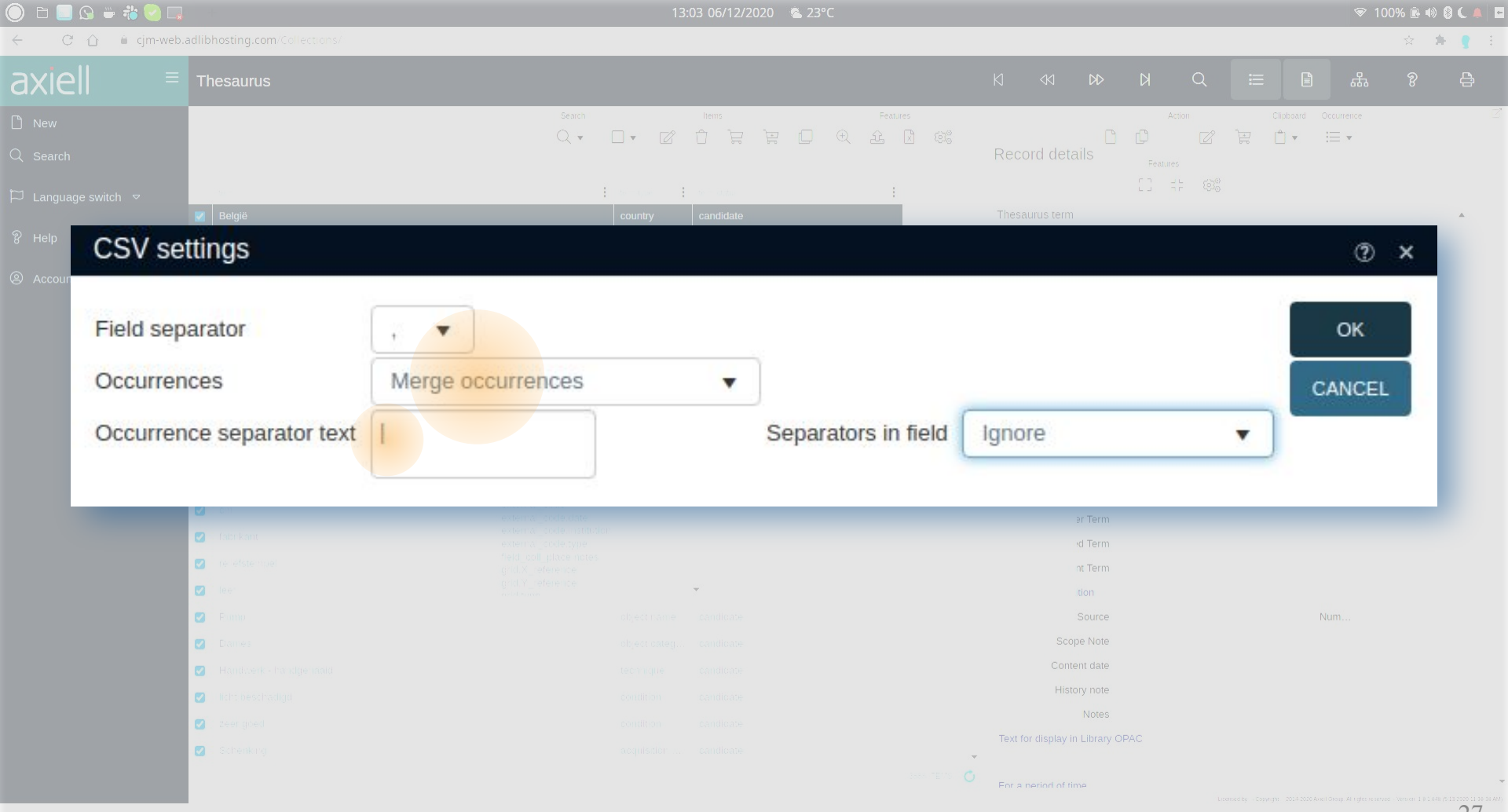
Action | Clipboard | Occurrence | Features

Thesaurus term

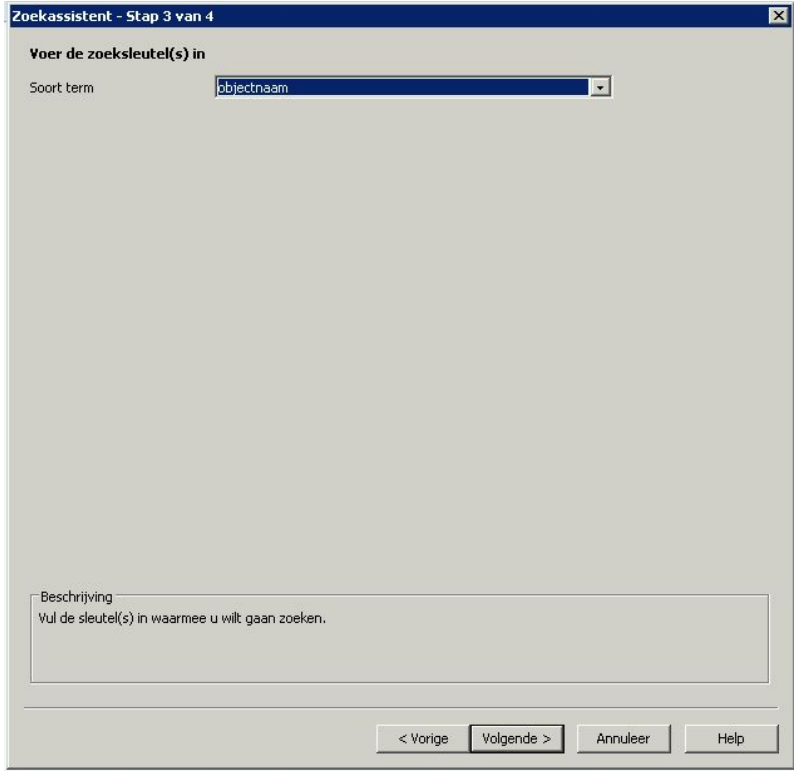
…

Term: België
…rm type: country
geographical keyword
place
…m code
Status: candidate
Use
…sed for
…er Term
…er Term
…d Term
…nt Term
…tion
Source
Num…
Scope Note
Content date
History note
Notes
Text for display in Library OPAC
For a period of time
Period start

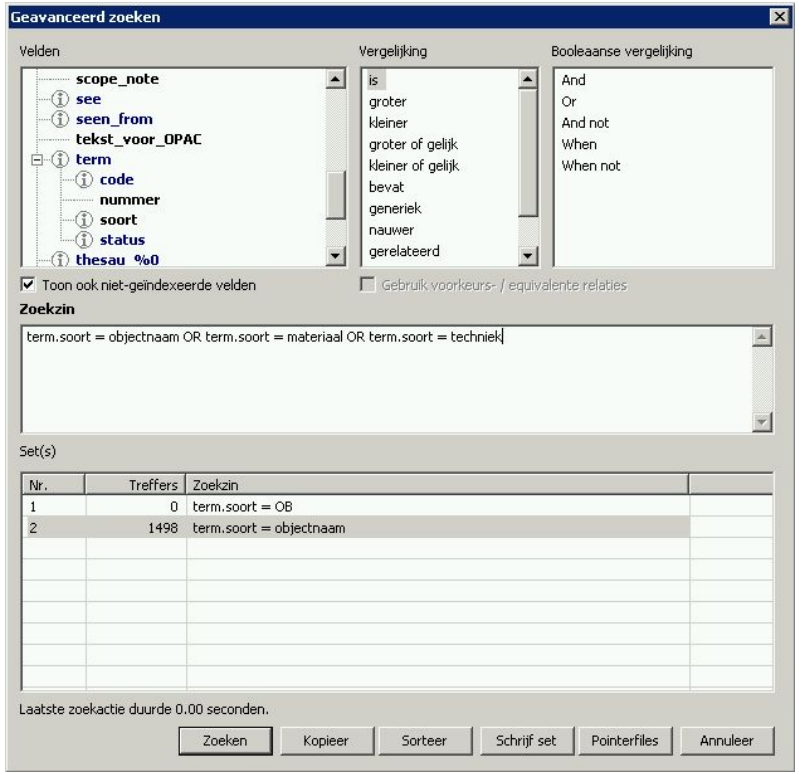
Licensed by: - Copyright © 2014-2020 Axiell Group. All rights reserved. - Version: 1.8.1.848 (5/13/2020 11:38:38 AM)


13:03 06/12/2020
23°C
100%
cjm-web.adlibhosting.com/Collections/
axiell
Thesaurus
New
Search
Language switch
Help
Record details
Thesaurus term
CSV settings
Field separator
,
Occurrences
Merge occurrences
Occurrence separator text
|
Separators in field
Ignore
OK
CANCEL
Source
Scope Note
Content date
History note
Notes
Text for display in Library OPAC
For a period of time
export (2).csv
Show all

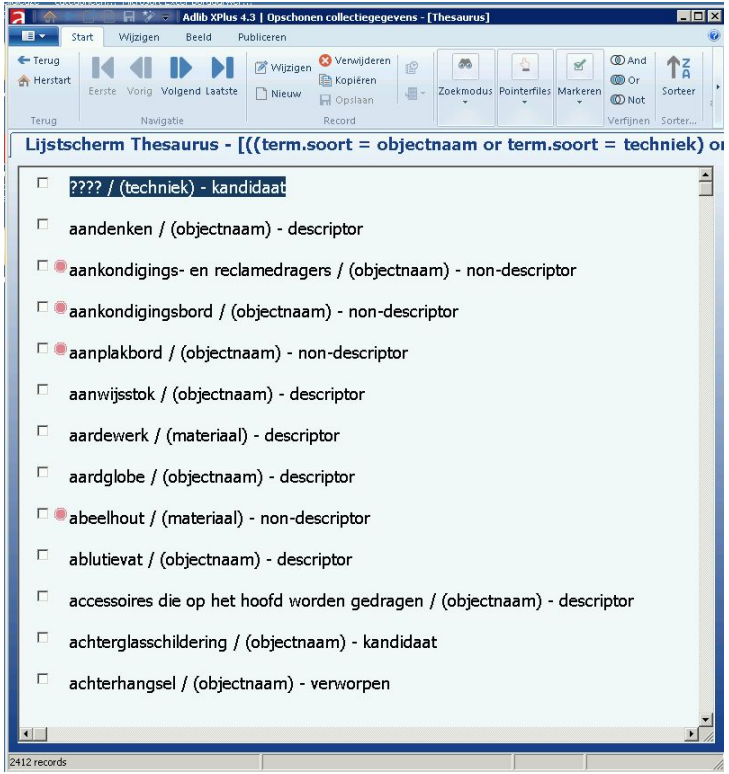
export via Adlib for Windows

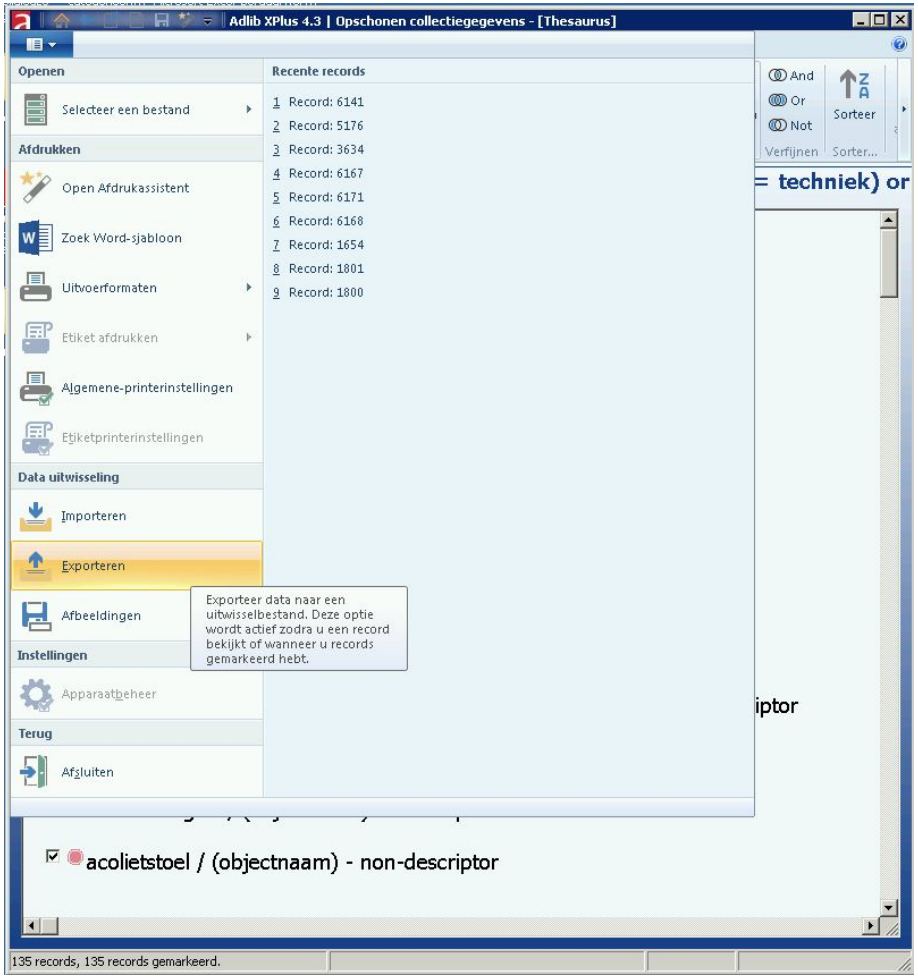


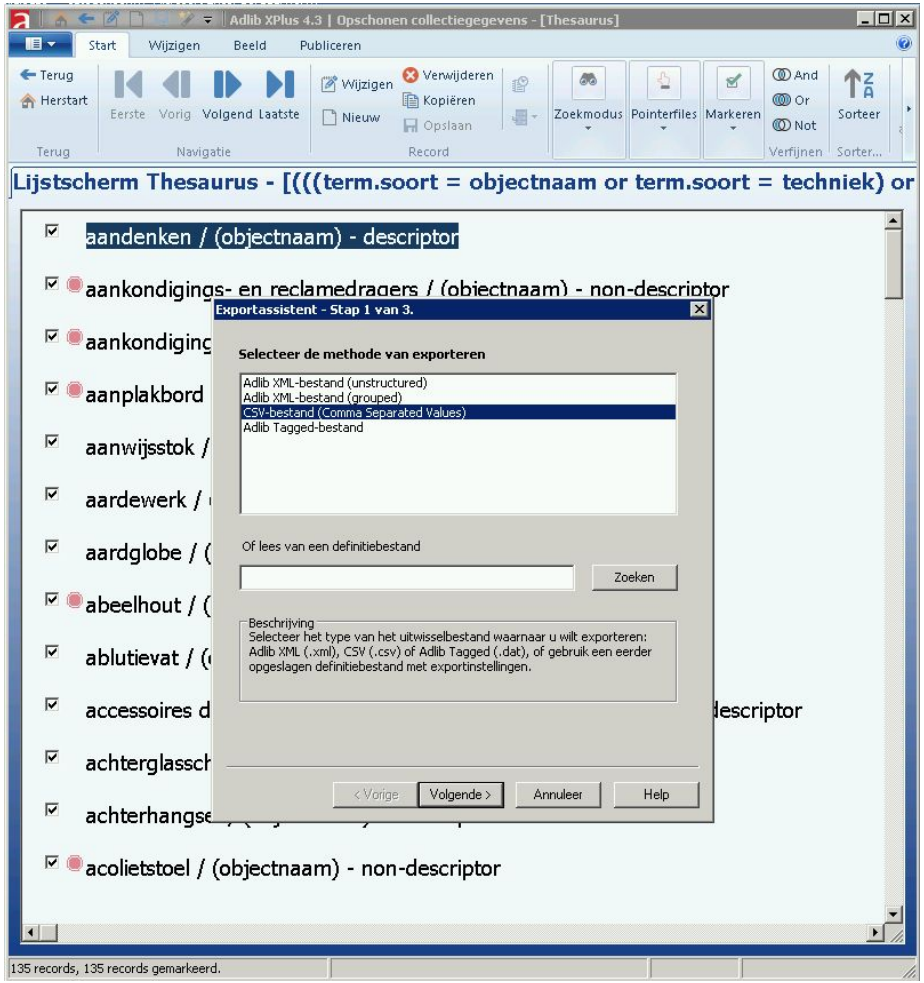
Adlib export



Adlib thesaurus




Adlib XPlus 4.3 | Opschonen collectiegegevens - [Thesaurus]
Openen
Selecteer een bestand
Afdrukken
Open Afdrukassistent
Zoek Word-sjabloon
Uitvoerformaten
Etiket afdrukken
Algemene-printerinstellingen
Etiketprinterinstellingen
Data uitwisseling
Importeren
Exporteren
Afbeeldingen
Instellingen
Apparaatbeheer
Terug
Afsluiten
Recente records
1 Record: 6141
2 Record: 5176
3 Record: 3634
4 Record: 6167
5 Record: 6171
6 Record: 6168
7 Record: 1654
8 Record: 1801
9 Record: 1800
Exporteer data naar een uitwisselbestand. Deze optie wordt actief zodra u een record bekijkt of wanneer u records gemarkeerd hebt.
And
Or
Not
Verfijnen
Sorteer
Sorter...
= techniek) or
iptor
acolietstoel / (objectnaam) - non-descriptor
135 records, 135 records gemarkeerd.


Adlib XPlus 4.3 | Opschonen collectiegegevens - [Thesaurus]
Start
Wijzigen
Beeld
Publiceren
Terug
Herstart
Eerste
Vorig
Volgend
Laatste
Wijzigen
Nieuw
Verwijderen
Kopiëren
Opslaan
Zoekmodus
Pointerfiles
Markeren
And
Or
Not
Sorteer
Terug
Navigatie
Record
Verfijnen
Sorter...
Lijstscherm Thesaurus - [(((term.soort = objectnaam or term.soort = techniek) or
aandenken / (objectnaam) - descriptor
aankondigings- en reclamedragers / (objectnaam) - non-descriptor
aankondiging
aanplakbord
aanwijsstok /
aardewerk /
aardglobe / (
abeelhout / (
ablutievat / (
accessoires d
achterglassch
achterhangse
acolietstoel / (objectnaam) - non-descriptor
Exportassistent - Stap 1 van 3.
Selecteer de methode van exporteren
Adlib XML-bestand (unstructured)
Adlib XML-bestand (grouped)
CSV-bestand (Comma Separated Values)
Adlib Tagged-bestand
Of lees van een definitiebestand
Zoeken
Beschrijving
Selecteer het type van het uitwisselbestand waarnaar u wilt exporteren: Adlib XML (.xml), CSV (.csv) of Adlib Tagged (.dat), of gebruik een eerder opgeslagen definitiebestand met exportinstellingen.
< Vorige
Volgende >
Annuleer
Help
135 records, 135 records gemarkeerd.

Excel

Henk Vanstappen

File | Home | Insert | Page Layout | Formulas | Data | Review | View | Help | Tell me what you want to do

Cut
Copy
Paste
Format Painter
Clipboard
Font
Wrap Text
AutoSum
Fill
Clear
Sort & Filter
Find & Select
Editing

Text Import Wizard - Step 1 of 3

The Text Wizard has determined that your data is Delimited.

If this is correct, choose Next, or choose the data type that best describes your data.

Original data type

Choose the file type that best describes your data:

- (●) Delimited - Characters such as commas or tabs separate each field.
- ( ) Fixed width - Fields are aligned in columns with spaces between each field.

Start import at row: 1

File origin: 65001 : Unicode (UTF-8)

[ ] My data has headers.

Preview of file \\tsclient\2020_webinar\data\PROD\export_bulskampveld_20200212.csv.

```
1 term;term.type;term.status;used_for;use;source;priref;term.number
2 België;COUNTRY|GEOKEYW|PLACE;3;;;;160000001;
3 Izegem;PLACE|GEOKEYW;3;;;;160000002;
4 Brugge;PLACE|GEOKEYW;3;;;;160000003;
5 Sint-Eloois-Winkel;PLACE|GEOKEYW;3;;;;160000004;
```

Cancel | < Back | Next > | Finish

Ready

Excel

Henk Vanstappen

File | Home | Insert | Page Layout | Formulas | Data | Review | View | Help | Tell me what you want to do

Cut
Copy
Paste
Format Painter
Clipboard
Font
Wrap Text
AutoSum
Fill
Clear
Sort & Filter
Find & Select
Editing

Text Import Wizard - Step 2 of 3

This screen lets you set the delimiters your data contains. You can see how your text is affected in the preview below.

Delimiters

- [ ] Tab
- [x] Semicolon
- [ ] Comma
- [ ] Space
- [ ] Other:

- [ ] Treat consecutive delimiters as one

Text qualifier: "

Data preview

term	term.type	term.status	used_for	use	source	p
België	COUNTRY|GEOKEYW|PLACE	3				1
Izegem	PLACE|GEOKEYW	3				1
Brugge	PLACE|GEOKEYW	3				1
Sint-Eloois-Winkel	PLACE|GEOKEYW	3				1

Cancel | < Back | Next > | Finish

Ready

export_bulskampveld_20200212.csv - Excel

Henk Vanstappen

File Home Insert Page Layout Formulas Data Review View Help Tell me what you want to do Share

Paste Cut Copy Format Painter Clipboard | Calibri 11 Font | Wrap Text Merge & Center Alignment | General Number | Conditional Formatting Format as Table | Normal Bad Good Neutral Calculation Check Cell Styles | Insert Delete Format Cells | AutoSum Fill Clear Sort & Filter Find & Select Editing

F1 source

	A	B	C	D	E
1	term	term.type	term.statu	used_for	use
2	België	COUNTRY|GEOKEYW|PLACE	3		
3	Izegem	PLACE|GEOKEYW	3		
4	Brugge	PLACE|GEOKEYW	3		
5	Sint-Eloois-Winkel	PLACE|GEOKEYW	3		
6	Nederland	COUNTRY|GEOKEYW|PLACE	3		
7	Hasselt	PLACE|GEOKEYW	3		
8	Damme	PLACE	3		
9	S-schoenen	COLL	3		
10	schoenmaat	DIM	3		
11	breedte bal	DIM	3		
12	hakhoogte	DIM	3		
13	cm	UNIT	3		
14	fabrikant	ROLE	3		
15	reliëfstempel	INSCTYPE	3		
16	leer	MATER|OBJCAT	3		
17	Pump	OBJECT|NAMTYPE	3		
18	Dames	OBJCAT	3		
19	Handwerk - handgenaaid	TECHN	3		
20	licht beschadigd	COND	3		
21	zeer goed	COND	3		
22	Schenking	ACQMETH	3		
23	buitenzool	POSITION	3		
24	nieuw	COND	3		
25	buitenzijde hiellap	POSITION	3		
26	insigne	INSCTYPE	3		
27	stof en leder	MATER	3		
28	Heren	OBJCAT|OBJECT	3		
29	Leuven	GEOKEYW|PLACE	3		
30	Aankoop	ACQMETH	3		
31	voering en buitenhak	POSITION	3		
32	kalfsleder	MATER	3		
33	Schachtschoen	OBJECT|NAMTYPE|OBJCAT	3		

export_bulskampveld_20200212

Ready Average: 161934522 Count: 4008 Sum: 6,38022E+11 100%

20190618.xlsx [Shared] - Microsoft Excel

File Home Insert Page Layout Formulas Data Review View

Normal Page Layout Page Break Preview Custom Views Full Screen Workbook Views | Ruler Formula Bar Gridlines Headings Show | Zoom 100% Zoom to Selection Zoom | New Window Arrange All Freeze Panes Split Hide Unhide View Side by Side Synchronous Scrolling Reset Window Position Window | Save Workspace Switch Windows | Macros Macros

G86

	A	B	C	D	E	F	G	H	I
1	priref	term	term.type	term.status	use	used_for	USE	Correctie	Opmerking
86	710000183	kruisbeeld	objectnaam	descriptor		Kruisbeeld met 7 koor	kandelaars		
87	710000185	offerblok	objectnaam	kandidaat		offerblok-reliekhouder			
88	710000186	kruis (object)	objectnaam	verworpen	kruis (beeldend werk)				
89	710000188	rozenkrans	objectnaam	non-descriptor	bidsnoer				
90	710000189	sokkel	objectnaam	descriptor		sokkel met wandkandelaar			
91	710000191	armstoel	objectnaam	descriptor		armstoelen			
92	710000193	kerkstoel	objectnaam	non-descriptor	stoel				
93	710000196	halfedelsteen	materiaal	descriptor		siersteen			
94	710000197	kroon (hoofdtooi)	objectnaam	descriptor		Kroon voor beeld			
95	710000198	koorlezenaar	objectnaam	non-descriptor	koorlessenaar				
96	710000199	harmonium	objectnaam	descriptor		harmoniums			
97	710000201	orgelkast	objectnaam	non-descriptor	orgel				
98	710000202	doophek	objectnaam	descriptor		doophekken			
99	710000203	doopvont	objectnaam	descriptor		doopbekken			
100	710000204	reliekkruis	objectnaam	non-descriptor	reliekhouder				
101	710000205	reliekmonstrans	objectnaam	descriptor					
102	710000206	vaas	objectnaam	descriptor		siervaas			
103	710000208	wijdingskruis	objectnaam	descriptor		wijdingskruisen			
104	710000209	flambeeuw	objectnaam	non-descriptor	flambeeuw (liturgische objecten)				
105	710000210	misdienaarskruk	objectnaam	descriptor		acolietstoel			
106	710000212	altaarbel	objectnaam	descriptor		altaarbellen			

20190618

Ready 150%

5. OPSCHONEN

Opschonen

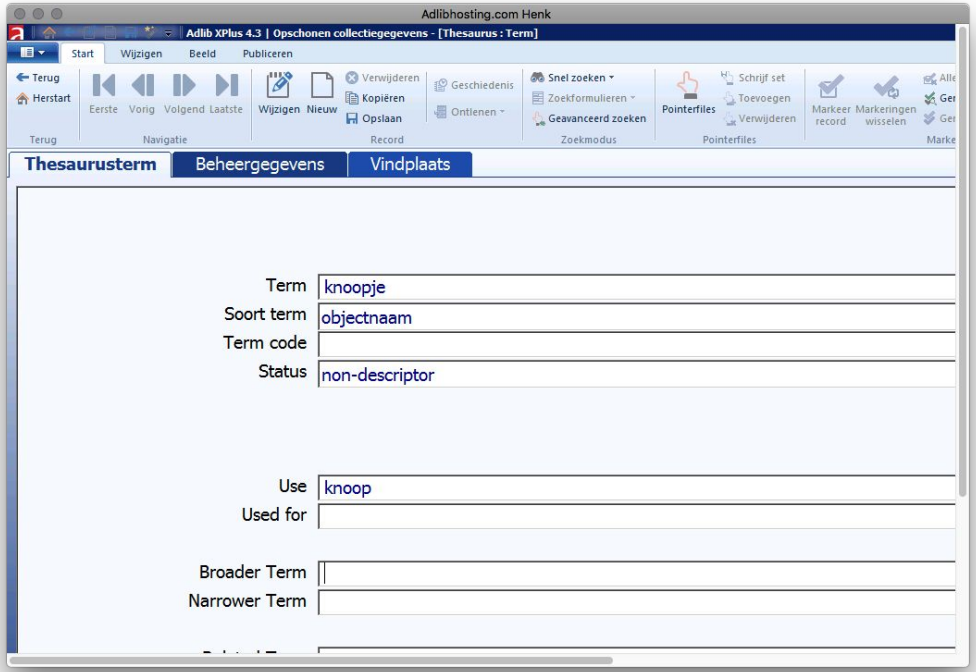
via objectrecords

+ objectniveau
+ context
+ zoek/vervang
+ termen splitsen

via thesaurus

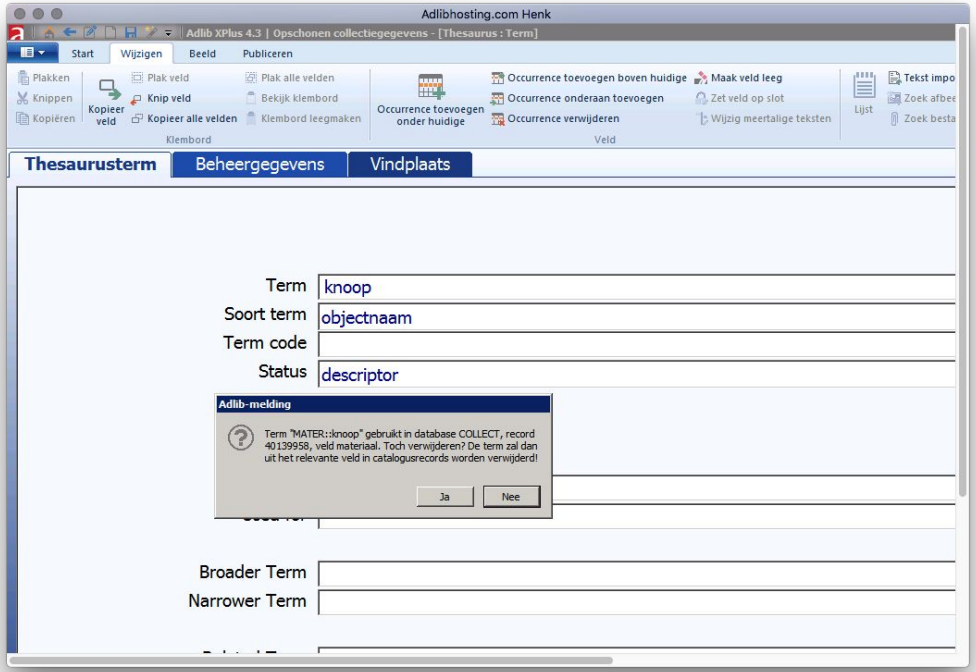
+ snel
+ herbruikbaar

Term wordt niet-voorkeursterm



niet-voorkeursterm wordt automatisch aangepast

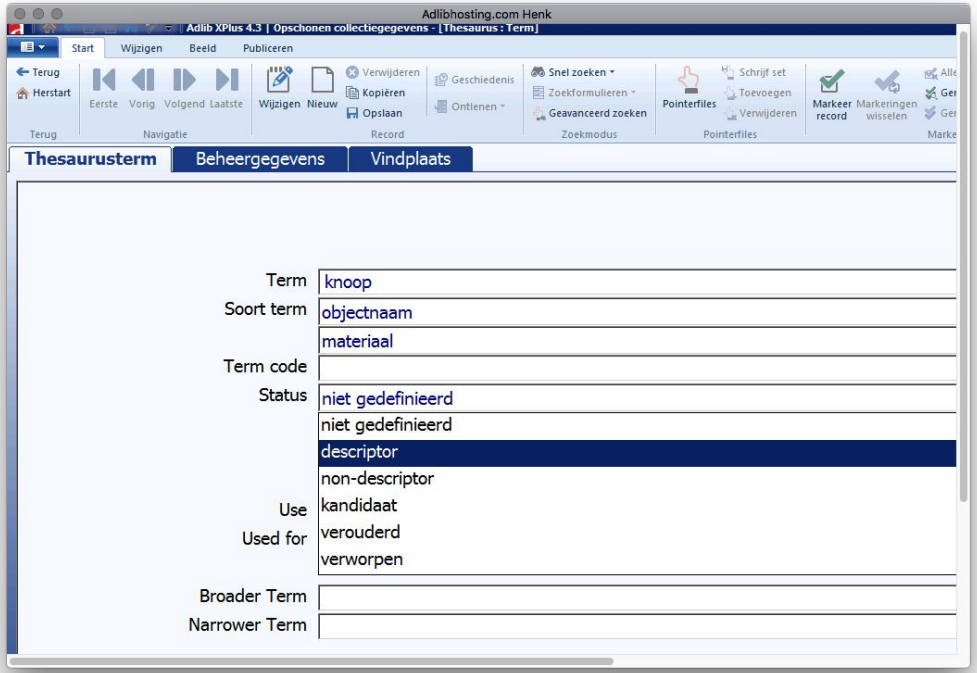
Meerdere termsoorten



Term aanpassen geldt voor alle velden.

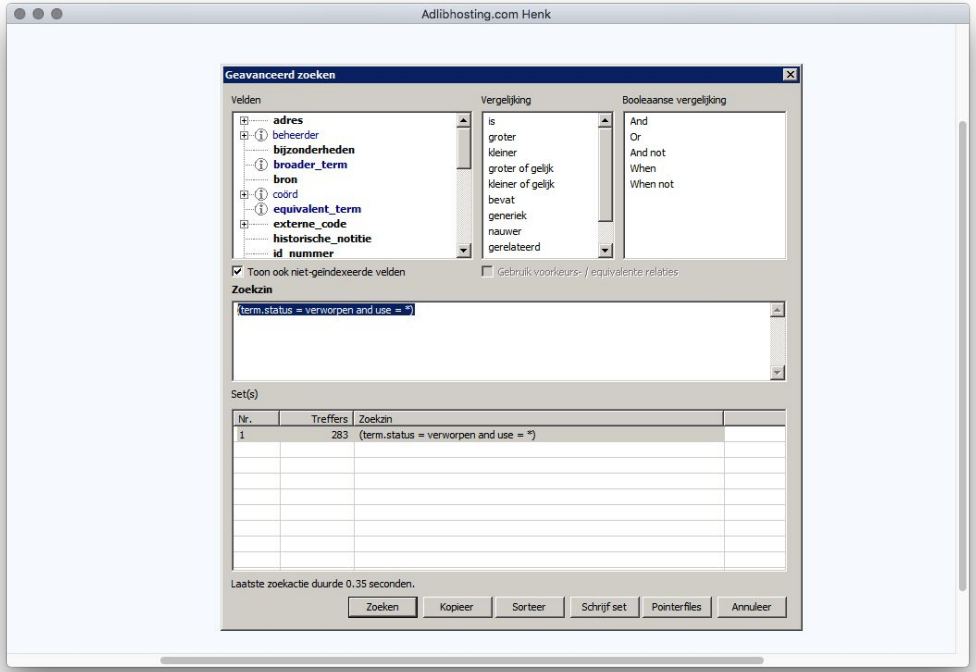
Pas objectrecords aan vooraleer soort term te verwijderen

Voorkeursterm status aanpassen



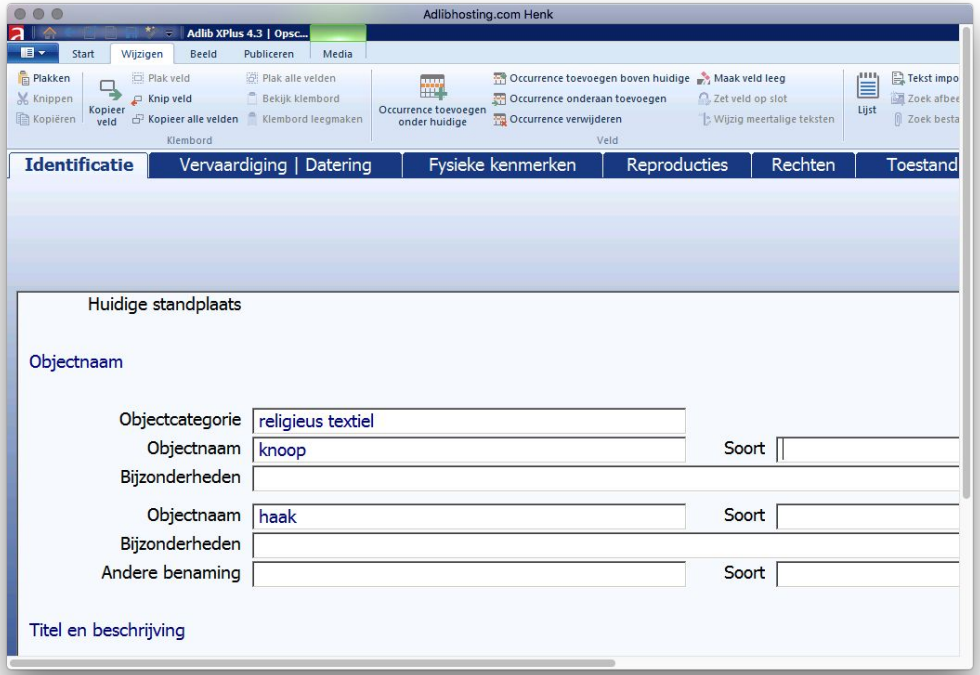
Status > descriptor

Term verwijderen



Selecteer status ‘verworpen’ (Optioneel)

Term ontdubbelen



Termen ontdubbelen kan enkel in objectrecords.

6. RECONCILIËREN

23:41 06/24/2020 24°C 100%
Not secure | getty.edu/vow/AATFullDisplay?find=lithogra*&logic=AND&note=&english=N&prev_page=1&subjectid=300041379

New Search Previous Page Help

Click the icon to view the hierarchy.

Semantic View (JSON, RDF, N3/Turtle, N-Triples)

Representative Images: 1

ID: 300041379

Record Type: concept

Page Link: http://vocab.getty.edu/page/aat/300041379

lithographs (planographic prints, <prints by process: transfer method>, ... Visual and Verbal Communication (hierarchy name))

Note: Prints made using the process of lithography, which is a planographic printing process in which a design is deposited on the stone or plate with a greasy substance which will accept ink.

Terms:

lithographs (**preferred**,C,U,English-P,D,U,PN)
lithograph (C,U,English,AD,U,SN)
平版印刷版畫 (C,U,Chinese (traditional)-P,D,U,U)
píng bǎn yìn shuā bǎn huà (C,U,Chinese (transliterated Hanyu Pinyin)-P,UF,U,U)
ping ban yin shua ban hua (C,U,Chinese (transliterated Pinyin without tones)-P,UF,U,U)
p'ing pan yin shua pan hua (C,U,Chinese (transliterated Wade-Giles)-P,UF,U,U)
lithografieën (prenten) (C,U,Dutch-P,D,U,U)
lithografie (prent) (C,U,Dutch,AD,U,SN)
litho (C,U,Dutch,UF,U,SN)
steendruk (C,U,Dutch,UF,U,SN)
litho's (C,U,Dutch,UF,U,U)
steendrukken (C,U,Dutch,UF,U,U)

waarom reconciliëren?

1. Unieke identificatie met concept
   - ondubbelzinnig
   - universeel
2. Verrijking
   - scope notes
   - meertalig
   - …

Art & Architecture Thesaurus

Not Secure | www.getty.edu/research/tools/vocabularies/...

The Getty: Explore

Getty360 | Calendar | Blog | Connect with Us | Shop | Support Us

The Getty Research Institute

Exhibitions & Events | Special Collections | Library | Search Tools & Databases | Scholars & Projects | Publication

Search Tools & Databases

- Library Catalog
- Getty Research Portal
- Collection Inventories & Finding Aids
- Photo Archive
- Research Guides & Bibliographies
- Digital Collections
- Article & Research Databases
- Collecting & Provenance Research
- BHA & RILA
- Getty Vocabularies
  - **Art & Architecture Thesaurus (AAT) ®**
    - About AAT
    - AAT: Frequently Asked Questions
    - How to use AAT
  - Cultural Objects Name Authority (CONA) ®
  - Getty Thesaurus of

Art & Architecture Thesaurus® Online

Search the AAT

Help

Find Term or ID: knoop

Search

AND OR

Clear

Note:

Pop-up Search

Browse the AAT hierarchies

Copyright information

Search Tips

For the Find Term or Note field, you may use **AND** and **OR** (all in upper case) [e.g., 1) windsor chairs, 2) chairs OR rockers, 3) chairs OR rockers OR armchairs, 4) bow-back AND windsor, 5) windsor AND (rockers OR chairs), 6) (windsor OR boston) AND (rockers OR chairs)]. Wildcard is the asterisk (*); right truncation only. To find an exact match rather than a key word in the Find Term field, use quotes [e.g., "chairs"]. If you wish to search the term and note together, click on the buttons for AND or OR.

Adlib XPlus 4.3 | Opschonen collectiegegevens - [Thesaurus : Term]

Start | Wijzigen | Beeld | Publiceren

Terug | Herstart | Eerste | Vorig | Volgend | Laatste | Wijzigen | Nieuw | Verwijderen | Kopiëren | Opslaan | Zoekmodus | Pointerfiles | Schrijf set | Toevoegen | Verwijderen | Markeren | And | Or | Not | Sorteer | Etiket afdrukken

Thesaurusterm | Beheergegevens | Vindplaats

Term	knoop
Soort term	objectnaam
	materiaal
Term code	
Status	descriptor
Use	
Used for	knoopje
Broader Term	
Narrower Term	
Related Term	
Equivalent Term	
Bron	http://vocab.getty.edu/aat
Nummer	300239261

Record 1 van 1 | Recordnummer 3634

export_20200212.csv - Excel

File | Home | Insert | Page Layout | Formulas | Data | Review | View | Help | Tell me what you want to do

Get Data | From Text/CSV | From Web | From Table/ Range | Recent Sources | Existing Connections | Refresh All | Queries & Connections | Properties | Edit Links | Sort | Filter | Clear | Reapply | Advanced | Text to Columns | Flash Fill | Remove Duplicates | Data Validation | Consolidate | Relationships | Manage Data Model | What-If Analysis | Forecast Sheet | Group | Ungroup

J1466

=HYPERLINK(CONCAT("http://www.getty.edu/vow/AATServlet?english=N&find=";A2;"&logic=AND&page=1&note="))

	A	B	J
1	term	term.type	
1465	Haar/vezelbewer	OBJCAT	
1466	haarborstel	OBJECT|NAMTYPE	http://www.getty.edu/vow/AATServlet?english=N&find=haarborstel&log
1467	haarborstel en ka	NAMTYPE	
1468	haar		
1469	haar		
1470	haar		
1471	haar		
1472	haar		
1473	Haar		
1474	haar		
1475	hage		
1476	hage		
1477	hage		

http://www.getty.edu/vow/AATServlet?english=N&find=haarborstel&logic=AND&page=1&note=

Zoeken...

Art & Architecture Thesaur...

Find Name: **haarborstel**

Logic:

Note:

1 result

View Selected Records | Select All Records | Clear All

First | Previous | Next | Last

Page: 1

Click the icon to view the hierarchy.
Check boxes to view multiple records at once.

1. **hairbrushes (grooming tools)**
(<brushes for personal use>, <equipment for grooming, hygiene and health care>, ... Furnishings and Equipment (hierarchy name)) [300236713]
haarborstel
haarborstels

VERBIND JE TERMEN

Link je Adlib-thesaurus in Excel aan de AAT

& Introductie CultuurLINK, Importeren met Adlib Designer, Eenvoudig opschonen

7. AUTOMATISEREN

Open Refine

OpenRefine 20190618 csv Permalink

Facet / Filter | Undo / Redo 0 / 0

Refresh | Reset All | Remove All

term change

2402 choices Sort by: name count Cluster

???? 1
's Hertogenbos 1
aandenken 1
aankondigings- en reclamedragers 1
aankondigingsbord 1
aanplakbord 1
aanwijsstok 1
aardewerk 1
aardglobe 1
abc 1
abeelhout 1

2412 rows

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

All		priref	term	term.type	term.status
1.	1235		????	techniek	kandidaat
2.	1487		aandenken		
3.	1488		aankondigings- en reclam		
4.	1489		aankondigingsbord		
5.	1491		aanplakbord		
6.	1492		aanwijsstok		
7.	1494		aardewerk		
8.	1496		aardglobe		
9.	1500		abeelhout		
10.	1501		ablutievat		
11.	1504		accessoires die op het hoo		
12.	1523		achterglasschildering		
13.	1526		achterhangsel		
14.	1563		acolietstoel		
15.	1564		acrylverf		
16.	1567		adventskandelaar		
17.	1568		aër		
18.	1570		afdruk		
19.	1571		affiche		
20.	1572		aflaatbrief		
21.	1573		afsluiting		
22.	1574		afsluitingsmiddelen		
23.	1575		afwerkings- en herafwerki		
24.	1576		afzelia		
25.	1577		agaatcamee		
26.	1578		agnus dei		

Cluster & Edit column "term"

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 | **20** clusters found

2	2	• geweven en (relief)borduurwerk (1 rows) • geweven en reliëfborduurwerk (1 rows)	geweven en (relief)borduurwerk
2	2	• Gevelsteen (1 rows) • gevelsteen (1 rows)	Gevelsteen
2	2	• Domper (1 rows) • domper (1 rows)	Domper
2	2	• glas en koper (1 rows) • koper en glas (1 rows)	glas en koper
2	2	• cache-pot (1 rows) • cachepot (1 rows)	cache-pot
2	2	• gemarmerd hout (1 rows) • hout, gemarmerd (1 rows)	gemarmerd hout
2	2	• katoen en polyester (1 rows)	katoen en polyester

 # Choices in Cluster

2 — 3

 # Rows in Cluster

2 — 3

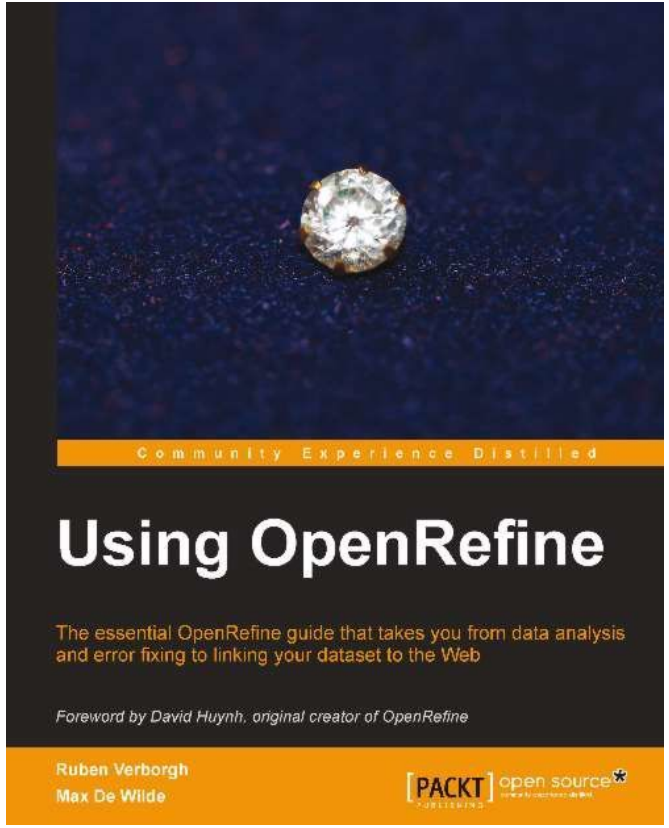
Average Length of Choices

6 — 44

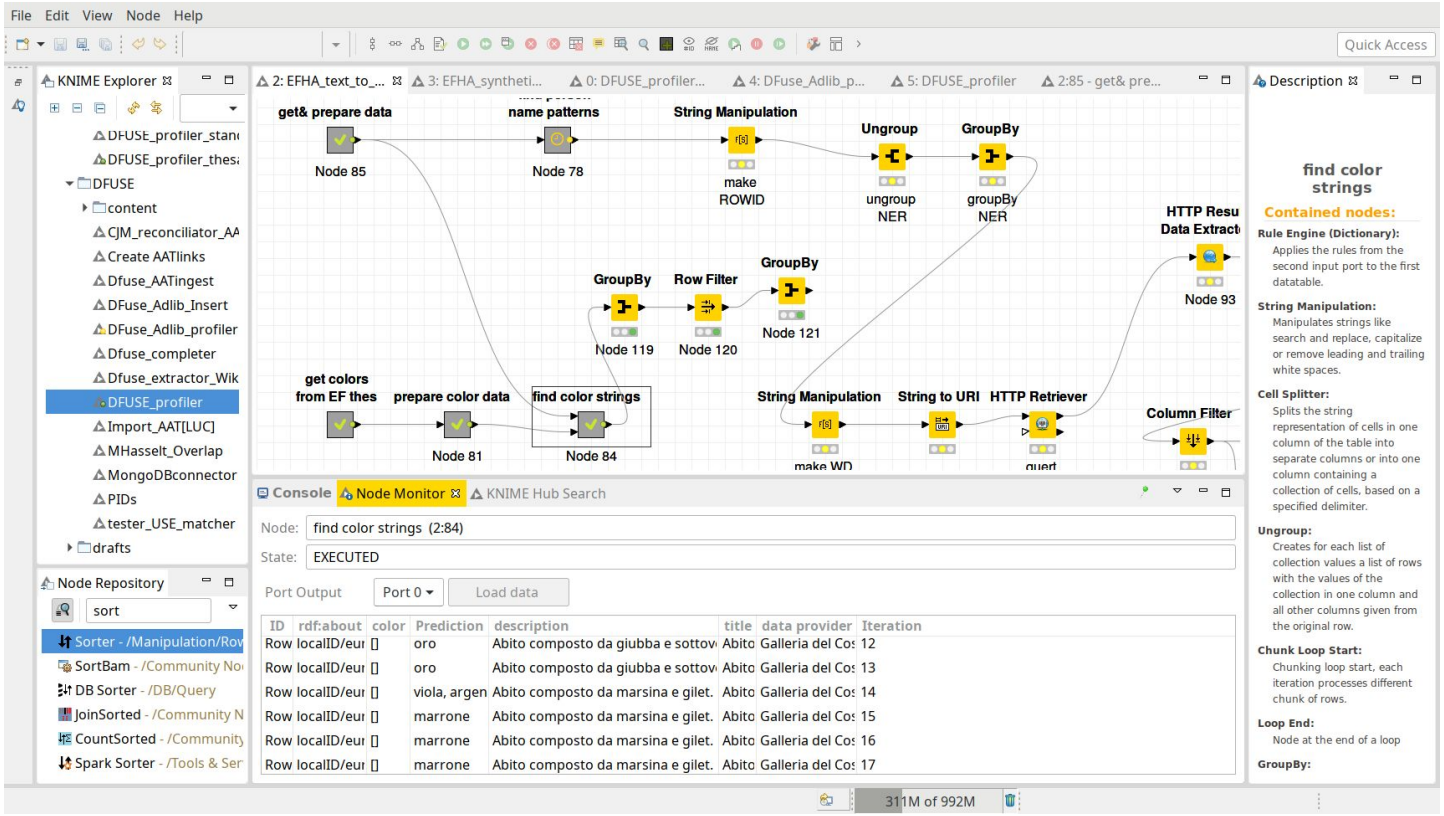
Length Variance of Choices

0 — 1

Open Refine


Community Experience Distilled
Using OpenRefine
The essential OpenRefine guide that takes you from data analysis and error fixing to linking your dataset to the Web
Foreword by David Huynh, original creator of OpenRefine
Ruben Verborgh
Max De Wilde
PACKT
open source

Knime


File Edit View Node Help
Quick Access
KNIME Explorer
DFUSE_profiler_stan
DFUSE_profiler_thes
DFUSE
content
CJM_reconciliator_AA
Create AATlinks
Dfuse_AATingest
DFuse_Adlib_Insert
DFuse_Adlib_profiler
Dfuse_completer
Dfuse_extractor_Wik
DFUSE_profiler
Import_AAT[LUC]
MHasselt_Overlap
MongoDBconnector
PIDs
tester_USE_matcher
drafts
Node Repository
sort
Sorter - /Manipulation/Row
SortBam - /Community No
DB Sorter - /DB/Query
JoinSorted - /Community N
CountSorted - /Community
Spark Sorter - /Tools & Ser
2: EFHA_text_to_...
3: EFHA_syntheti...
0: DFUSE_profiler...
4: DFuse_Adlib_p...
5: DFUSE_profiler
2:85 - get& pre...
get& prepare data
Node 85
name patterns
Node 78
String Manipulation
make ROWID
Ungroup
ungroup NER
GroupBy
groupBy NER
HTTP Resu
Data Extract
Node 93
GroupBy
Node 119
Row Filter
Node 120
GroupBy
Node 121
get colors from EF thes
Node 81
prepare color data
find color strings
Node 84
String Manipulation
make WD
String to URI
HTTP Retriever
quert
Column Filter
Console
Node Monitor
KNIME Hub Search
Node: find color strings (2:84)
State: EXECUTED
Port Output
Port 0
Load data
ID rdf:about color Prediction description title data provider Iteration
Row localID/eur [] oro Abito composto da giubba e sottov Abito Galleria del Cos 12
Row localID/eur [] oro Abito composto da giubba e sottov Abito Galleria del Cos 13
Row localID/eur [] viola, argen Abito composto da marsina e gilet. Abito Galleria del Cos 14
Row localID/eur [] marrone Abito composto da marsina e gilet. Abito Galleria del Cos 15
Row localID/eur [] marrone Abito composto da marsina e gilet. Abito Galleria del Cos 16
Row localID/eur [] marrone Abito composto da marsina e gilet. Abito Galleria del Cos 17
Description
find color strings
Contained nodes:
Rule Engine (Dictionary):
Applies the rules from the second input port to the first datatable.
String Manipulation:
Manipulates strings like search and replace, capitalize or remove leading and trailing white spaces.
Cell Splitter:
Splits the string representation of cells in one column of the table into separate columns or into one column containing a collection of cells, based on a specified delimiter.
Ungroup:
Creates for each list of collection values a list of rows with the values of the collection in one column and all other columns given from the original row.
Chunk Loop Start:
Chunking loop start, each iteration processes different chunk of rows.
Loop End:
Node at the end of a loop
GroupBy:
311M of 992M

DFUSE clean & reconcile

dFuse

web-dot-dfuse-250508.appspot.com/terminologyCleaner

DFUSE

- Home
- Terminology cleaner
- Agent cleaner
- Profiel
- Uitloggen

DFUSE Terminology cleaner and reconciliator

TERM	FREQ	TYPE	USE	USED FOR	AUTHORITY	REJECT	REMARK	STATUS
Carraramarmer	2	materiaal		1 Carrara marmer	Carrara marmer			
Carrara marmer	9	materiaal	Carraramarmer		0 Candidates			
Carrara-marmer	2	materiaal	steendrukgcfghcg		0 Candidates			
witte marmer uit Carrara	1	materiaal	9 Candidates		0 Candidates			
witte Carrara	1	materiaal	5 Candidates		0 Candidates			
witte carraramarmer	2	materiaal	11 Candidates		0 Candidates			
wit Carrara-marmer	1	materiaal	9 Candidates		0 Candidates			

Q&A

henk@datable.be