



# Rolf Blijleven

IT-diensten voor  
**musea**

**M** 06 12 872 892

**E** [rolf@rolfblijleven.nl](mailto:rolf@rolfblijleven.nl)

# Thesaurus Schonen

Workshop

voor

Erfgoedinzicht & partners

22 november 2016

# Programma

10:30 – 11:00 Ontvangst

11:00 ~ 13:00 Theoriegedeelte

- Thesaurustheorie
- Adlib & Thesauri

13:00 - 13:45 lunch

13:45 ~ 15:30 praktijkgedeelte  
behandeling specifieke vragen

15:30 ~ 16:00 nabespreken

Vragen staat altijd vrij,  
maar...

# Wat zijn validatiebestanden?

database *collect*

(Museumcatalogus)

objectnaam: schilderij

objectnaam\_linkref: 11223

objectcategorie: schilderijen

objectcategorie\_linkref: 2842

vervaardiger: Rijn, Rembrandt van

vervaardiger\_linkref: 0003

database *thesau* (Thesaurus)

recordnummer: 11223

term: schilderij

soort term: objectnaam

recordnummer: 2842

term: schilderijen

soort term: objectcategorie

database *people*

(Personen en Instellingen)

recordnummer: 0003

naam: Rijn, Rembrandt van

soort naam: vervaardiger

Onder water kijken: Ctrl-r

# Welke velden zijn “gevalideerd”?

- Alle *gevalideerde* velden zijn ook *gekoppelde*, d.w.z. *opgehaalde* velden  
‘Opgehaald’ uit een andere database:
  - vb: objectnaam van thesau in collect
  - vb: vervaardiger van people in collect
  - vb: auteur van people in document
- Kenmerk van gevalideerde velden in thesau en people:
  - a. én gekoppeld, dus onderstreept in Adlib
  - b. én hebben meestal een domein..
    - soort term (objectnaam, objectcategorie, trefwoord, materiaalsoort...)
    - soort naam (persoon, vervaardiger, verwervingsbron, auteur...)
    - een term kan meerdere domeinen hebben
  - c. ...meestal is niet altijd
    - In Adlib bibliotheek zijn domeinen variabel. Je kunt ze invullen, maar dat is niet verplicht.
    - Domein in thesau is geen verplicht veld. Kies je dan die term in een catalogusrecord voor een veld waarvoor wel een domein is gedefinieerd, dan wordt dat domein voor die term alsnog ingevuld.

# Welk validatiebestand, welk domein?

1. wijzig record (F11)

2. rechtsklik in het veld -> Eigenschappen -> tab Linkinformatie

The screenshot shows a software interface with a menu bar (Start, Wijzigen, Beeld) and a toolbar. Below the toolbar are several tabs: Gebruik | Objectgeschiedenis | Eigendom, Standplaats | Toekomstige verplaatsingen, Standplaatshistorie, Identificatie, Vervaardiging | Datering, Fysieke kenmerken, Iconografie, Opschriften | Merken, and Associat. The main area displays a record for '04418 tijdschrift' with the title '73 amateur radio today 1997 1998'. The 'Vervaardiging' section shows 'Vervaardiger: Green, dhr. Wayne'. A 'Veldeigenschappen' dialog box is open, showing the 'Linkinformatie' tab. The 'Algeme' section of the dialog box contains the following information:

Veld op schema	Data dictionary	Linkinformatie
<b>Algeme</b>		
Strikte validatie		Ja
Gekoppelde dataset		\\AdlibUG\Adlib Applications\data+people
Gekoppelde domeinnaam		MAKER
Link indextag		BA
Linkreferentie tag		l1
Reverse tag		
Forceren toegestaan		Ja
Meerdere occurrences		Nee



# Welke velden zijn gevalideerd? (2)

1. Adlib Designer
2. Werkmap insteller 
3. Application browse 
4. + voor mapje data
5. + voor collect of document
6. Klik Fields
7. Klik Linked databas (totdat thesau bovenaan staat)

Zo zie je *alle* gevalideerde velden, maar zoals ze 'onder water' heten, in de data dictionary.

Boven water, in Adlib, heten ze anders.

Een overzichtsl lijst maken is heel veel werk. Niet doen. Leer Designer.



Name	Tag	Linked database	Domain	Group
gebruik	US	thesau	USAGE	
afmeting.eenheid	EE	thesau	UNIT	Dimension
techniek	TK	thesau	TECHN	Technique
stratigrafie.grootheid	UN	thesau	STRATI_UNIT	Stratigraphy
stratigrafie.soort	UT	thesau	STRATI_TYPE	Stratigraphy
inscriptie.schrift	IS	thesau	SCRIPT	Inscription
school_stijl	SC	thesau	SCHOOL	
vervaardiger.rol	FU	thesau	ROLE	Production
inscriptie.vervaardiger.rol	IB	thesau	ROLE	Inscription
vondst.naam.rol	NM	thesau	ROLE	Collector
rechten.tvpe	Rt	thesau	RIGHTS	Rights

# thesau-velden *use, used for*

Use - Used for (gebruik - gebruikt voor)

- jargon:
  - voorkeursterm of preferred term PT
  - niet-voorkeursterm of non-preferred term NPT
- **Used for** ingevuld: de term is een **PT**
- **Use** ingevuld: de term is een **NPT**
- binnen een record maar één van beide ingevuld
- zo'n record heeft altijd een gespiegeld record (zie Recordnummer onderin)

Thesaurusterm	
Beheergegevens	
Term	
Term	<u>weefgetouw</u> PT
Soort term	objectnaam
Term code	
Status	niet gedefinieerd
Relaties	
Use	
Used for	<u>weefmachine</u> NPT

Recordnummer 1 van 1      Recordnummer 1565

Thesaurusterm	
Beheergegevens	
Term	
Term	<u>weefmachine</u> NPT
Soort term	objectnaam
Term code	
Status	niet gedefinieerd
Relaties	
Use	<u>weefgetouw</u> PT
Used for	

Recordnummer 1 van 1      Recordnummer 518

# thesauvelden *broader*, *narrower*, *equivalent* en *related* term

Thesaurusterm	
Beheergegevens	
Term	stalenweefmachine
Soort term	objectnaam
Term code	
Status	niet gedefinieerd
Relaties	
Use	
Used for	
Broader Term	<a href="#">weefgetouw</a> <a href="#">weefmachine</a>
Narrower Term	<a href="#">Jacquard-stalenweefmachne</a>
Related Term	<a href="#">Jacquard-weefmachine</a>
Equivalent Term	<a href="#">testweefmachine</a>

*broader*  
*narrower*  
*equivalent*  
*t*


spelen een rol  
bij zoeken in  
catalogi én de  
thesaurus

*Related* speelt alleen een rol bij  
zoeken in de thesaurus zelf



# Waarom validatiebestanden

- 1 x correct invoeren, daarna overal gebruiken.
  - scheelt invoerwerk
  - zeer efficiënt met geheugen
- Gebruik voorkeurstermen en niet-voorkeurstermen
  - Bij twijfel: kies een niet-voorkeursterm, die wordt dan automatisch vervangen door de voorkeursterm.
- Verbeter de vindbaarheid met
  - equivalente termen
  - bredere en nauwere termen
  - semantische factoren  
wie gebruikt dit?



Dus om te schonen  
hoef ik alleen maar  
het thesau-record  
aan te passen?

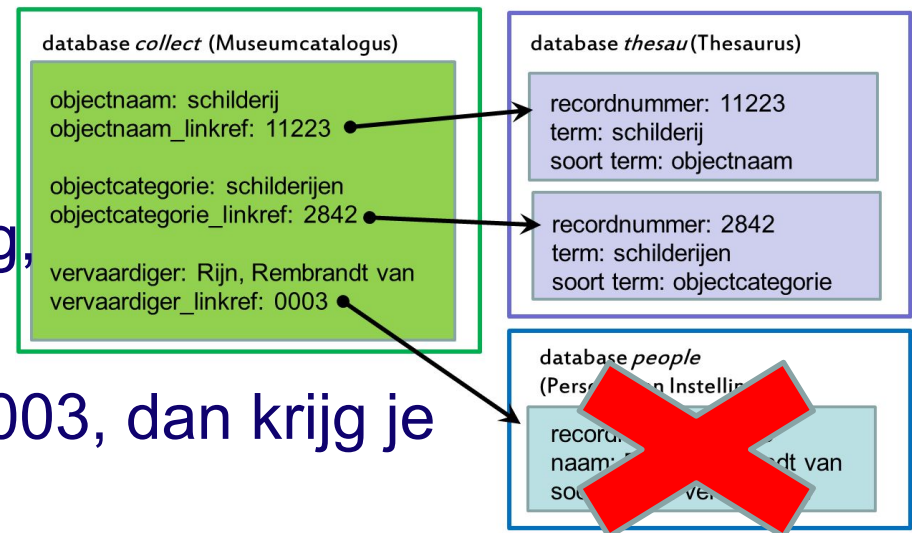
*Nou, dat ligt eraan*

# Verwijder niet zomaar een thesau-record

In “standaard” CBF Adlib t/m applicatie versie 3.5 is er alleen een heen-verwijzing, geen terugverwijzing.

Verwijder je people-record 0003, dan krijg je

- geen waarschuwing  
“record is in gebruik”
- wel een dode link  
“datacorruptie”



# Feedback- of terugkoppel-links

Applicatieversie vanaf 4.2 (eigenlijk 3.6) heeft *feedback-links*.

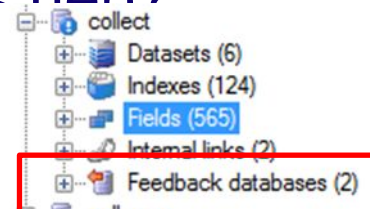
Er is dan een heenverwijzing én een terugverwijzing.

- Een thesaurusrecord oude stijl weet niet in welke records hij voorkomt. Een thesaurusrecord nieuwe stijl weet dat wel.
- niet elke Adlib heeft dit, vanwege 'performance'  
record opslaan vergt meer administratie, dus ook tijd.
- Tip! Applicatie-update met overstap naar SQL

Verwijder je een thesau-record dat in gebruik is, dan krijg je een waarschuwing. Lees die!

Hoe weet ik of ik feedback-links heb?

- techneuten kijken met Designer





# ..Lees de handleiding..

Adlib gebruikersgids voor versie 7.3

Paragraaf 12.4.2

Hoofdstuk 13



# Plannen van aanpak

## Stap 1: vind de fout

- als project: werk vanuit de thesaurus, selecteer op domein en dan op alfabet
- ad-hoc: je ziet een fout, die herstel je (maar dan wel grondig)

## Stap 2: Maak gebruik van PT en NPT

- Wordt een populaire term NPT?
  - Kan lang duren
  - werk verder (aan iets anders) in een ander Adlibje
- Completeer het validatie-records met BT's, NT's, ET's, RT's, scope notes, etc.
- Termstatus verouderd, verworpen -> niet meer te zien.

# Kijk eerst even rustig rond..

- Uw wijzigingen kunnen drastischer zijn dan u denkt.
- Vindt records met [veld]=[deze term]
- gebruik voorkeurs/equivalente relaties – of juist niet!
- Gebruik knop *Records met trefwoord* bij nieuwe of additionele invoer

**Recordopties**

Forceer record

Creëer record

Records met trefwoord

Technieken	Onderdeel	Materialen	Afmetingen
Techniek	Techniek	Materiaal	Afmeting
zinkets		papier	hoogte
			breedte

**Geavanceerd zoeken**

Velden	Vergelijking	Booleaanse vergelijking
adres	is	And
beheerder	groter	Or
<b>broader_term</b>	kleiner	And not
coörd	groter of gelijk	When
<b>equivalent_term</b>	kleiner of gelijk	When not
externe_code	bevat	
invoer	generiek	
<b>narrower_term</b>	nauwer	
<b>priréf</b>	gerelateerd	
<b>related term</b>		

Toon ook niet-geïndexeerde velden

Gebruik voorkeurs- / equivalente relaties

Zoekzin

# Zoeken en vervangen

Gebruik 'delen van woorden' voor een *vaste* tekenreeks *korter* dan de hele veldinhoud

Vervang [iets] door [iets anders]

o.a. om gewijzigde padnamen aan te passen

Yabba [dibi] doe ⇒ Yabba [dabba] doe

Vervang [iets] door [niets]

Yabba [dibi] dabba doe ⇒ Yabba dabba doe

Vervang \* om elke willekeurige volledige veldinhoud compleet te vervangen door iets anders

Niet mogelijk: iets nieuws voor of achter een willekeurige veldinhoud plakken

\* ⇒ \*[afstoten] werkt niet

Vervang "" om een leeg veld in een serie records te vullen

Versie 7: occurrence toevoegen

Voorbeeld: alle Verwervingsbronnen in één klap tot Persoon bombarderen

# Onderhoud zal nodig blijven

Benut de mogelijkheden:

Gebruik

Gebruikt voor

Equivalente term

termstatus

verouderd of verworpen: term is niet meer zichtbaar als keuze..

groep: term heeft nut in thesaurus, niet in record

“zelfreinigend”: kies bij invoeren een niet-voorkeursterm, dan neemt Adlib automatisch de juiste voorkeursterm

# Gereedschappen voor iedereen

- een groot beeldscherm
- een instance van Adlib open op de thesau, nog een instance van Adlib open op de catalogus, en evt Window aanpassen.
- de zoekassistent, sorteren, AND, OR en NOT, hiërarchie-browser
- de selectietaal, bijvoorbeeld alle thesaurustermen zonder domein:  
te = \* and do = ""
- externe bronnen!
  - [AAT](#), online-catalogi van andere instituten, woordenboeken, wikipedia, Google...
- de [handleiding](#)



# Gereedschappen voor gevorderden en/of applicatiebeheerders

## Exporteren

- Word-sjabloon -> Excel
- Exporteren vanuit Adlib
  - Comma Separated (CSV) -> Excel
  - XML
  - Adlib Tagged -> editor
- Exporteren vanuit Adlib Designer

## een editor

- bijvoorbeeld Notepad++

## Externe thesauri koppelen

## Bulk update

## Open Refine

## Erfgoedstats

(wat [handige links en boeken](#))



# Praktijkvoorbeeld

# projectmatige aanpak

1. Wat wil ik eigenlijk?
  - a. Welke velden en wat voor termenlijsten?
  - b. Welke afspraken voor consistentie?
  - c. Worden er externe bronnen (als AAT) gebruikt?
2. Inventarisatie
  - a. Maak een plan en een overzicht, zorg dat de afspraken duidelijk zijn
  - b. Welke termen heb je al, hoe erg is de vervuiling & ga je aanvullen?
3. Opschonen
  - a. Hou de voortgang bij
  - b. Eerst voorzichtig: werken de afspraken?
  - c. Verwijder of verwijst 'foute' termen, hiërarchie in orde maken: broader/narrower
4. Verrijking
  - a. Cultuurlink – thesauri koppelen
  - b. Erfgeo - geografie

# Velden

## Kenmerken

- Objectnaam
- Materiaal, Techniek
- Vervaardiger (personen), Plaats

## Onderwerp

- Wie (personen), Iconclass, Gebeurtenis, Trefwoord, Plaats, Datum

## Beheer

- Standplaats, Toestand


## Verwerving

- Verwervingsmethode, Van wie verworven (personen)

Microsoft Excel ribbon: BESTAND, START, INVOEGEN, PAGINA-INDELING, FORMULES, GEGEVENS, CONTROLEREN, BEELD, ONTWIKKELAARS, ACROBAT. Includes font settings (Calibri, 11), text alignment, and other formatting options.

Formulas bar: C8, WORDT

A	B	C	E	F	G	H	I	J	K
1	geel	naar fysieke kenmerken, globalere term gebruiken							
2		blijft techniek							
3		wordt materiaal							
4	rood	gaat eruit, willen we niet meer gebruiken							
5	cursief	hoort bij een andere deelcollectie							
6		is lastig! Verder uitzoeken.							
7		helemaal afgerond, term verwijderd							
8	<b>TERM</b>	<b>WORDT</b>	<b>Wat moet e</b>	<b>Aantal re</b>	<b>aantal re</b>	<b>Verand</b>	<b>Nog</b>	<b>In andere collecties</b>	<b>aantal</b>
54	TK fotomechanisch procedé								
55	TK frotteertechniek								
56	TK geel krijt	krijt		34	33			TK = *krijt* bevat 734 objecten	
57	TK gegriffeld	doorgriffelen		1		doorgr	0		
58	TK gehooagd met wit	krijt		379	352			Bibliotheek; China (RPK); silhouetportretten Barbiers (RPK); LJ Veen (RPK)	27
59	TK gehooagd met wit krijt	krijt		89	87				
60	TK gehooagd met witte dekverf	dekverf		101	101				
61	TK gekleurd lithografisch krijt	lithografisch krijt		11					
62	TK gekleurd met sjablonen			40	4			sierpapier = documenten en objecten (RPK);	36
63	TK gekleurd vetkrijt	vetkrijt		2		vetkrijt	0		
64	TK gekleurd waskrijt	waskrijt		1		"gekleu	0		
65	TK gekwadreed								
66	TK gele dekverf	dekverf		3			0		
67	TK geolied zwart krijt	krijt		3			0		
68	TK geoliede houtskool	houtskool		17	17		0		
69	TK gewassen			22	22		0		
70	TK gewassen inkt	inkt		10	10		0		
71	TK giclée	gicléedruk		1		giclée i	0		
72	TK gom								
73	TK gouache			542	0			gouache (waterverf) = accessoire; zuid-Azie (RPK); meubelen; foto's; papier (GS); kostuum en modeprenten; varia 18e eeuw	16
74	TK gouache au pochoir								
75	TK goud op snee	op snee		32	32			0 klopt dit, gezien instructie "kleur op snee"?	
76	TK goudopdruk								
77	TK grafiet			483	483			0 *grafiet* 9	
78	TK graveren (drukprocedé)								
79	TK graveren (inkerven)								
80	TK griffelsporen	doorgriffelen		6	0		0		
81	TK grijs gewassen	penseel in grijs		61	61		0		
82	TK grijs krijt	krijt		14	14		0		
83	TK grijs dekverf	dekverf					0		



*En nu aan de  
slag!*