

CATCH



CONTINUOUS
ACCESS
TO
CULTURAL
HERITAGE
PLUS

OpenSKOS / VocabulaireBank

Hennie Brugman

Meertens Instituut

Technisch coordinator CATCHPlus



Overzicht

1. *Inleiding*
2. *Ontstaansgeschiedenis*
3. *“Technische acronymen”*
4. *Functionele beschrijving en demo*
5. *Toepassingen in de praktijk*
6. *Content en beschikbaarheid*
7. *Stakeholders en toekomstige ontwikkelingen*
8. *Basis/onderdeel van een (inter)nationaal platform?*

CATCH



CONTINUOUS
ACCESS
TO
CULTURAL
HERITAGE
PLUS

Inleiding

CATCH



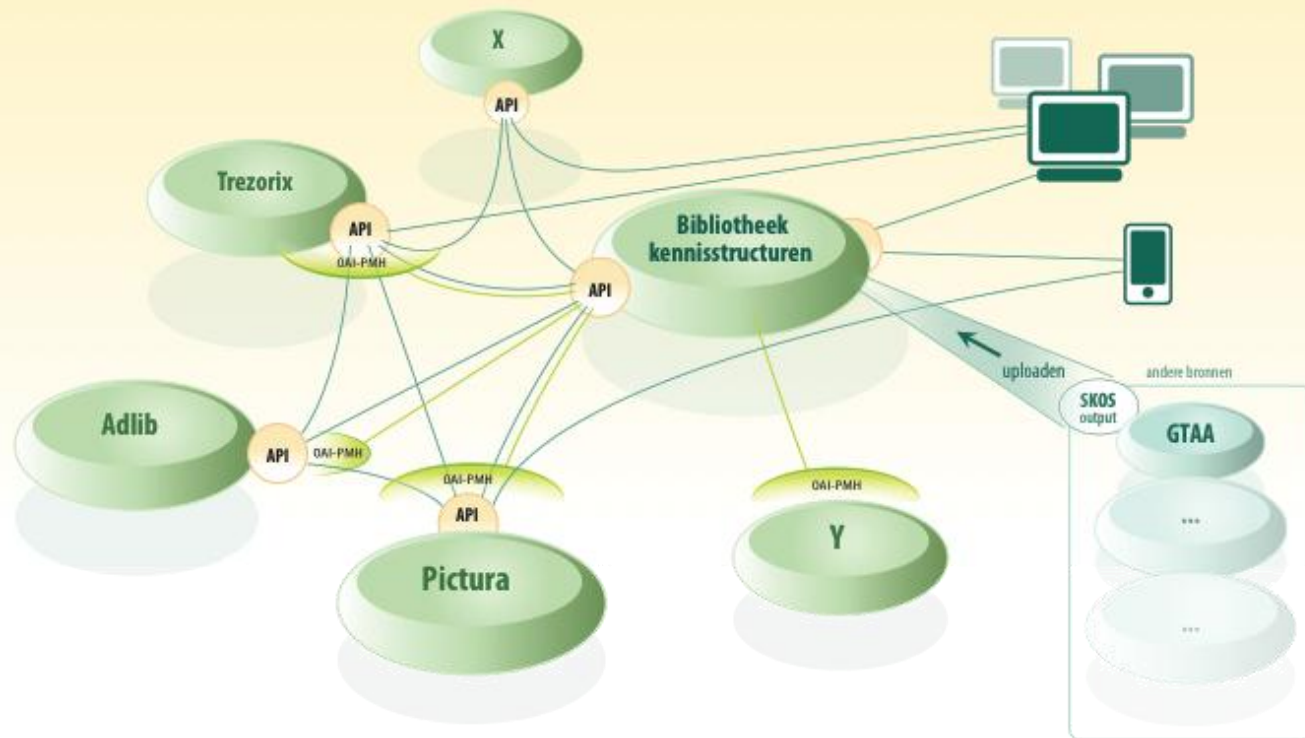
CONTINUOUS
ACCESS
TO
CULTURAL
HERITAGE
PLUS

Problemen mbt gebruik van vocabulaires

- Bouwen en beheren van thesauri is tijdrovend
 - Liever hergebruik
 - Hoe ontdek en evalueer je bestaande thesauri?
- Je software-tools moeten specifieke thesauri ondersteunen
- Meestal hooguit als browsable en doorzoekbare website gepubliceerd
 - Niet als data, niet in standaardvorm
- Toolbouwers moeten veel maatwerk verrichten

Visie

Op weg naar een wereldstandaard.
Voor uitwisseling van kennisstructuren.





OpenSKOS doelen

- Standaard *formaat* en *access* methoden
 - SKOS, SKOS gebaseerde REST API
- Web *publicatie* van vocabulaires
 - Als Linked Open Data
 - Als doorzoekbare en browse-bare dataset → REST API
 - Te gebruiken voor duurzame referenties naar concepten → persistente identifiers en/of stabiele URLs
 - Te harvesten door middel van OAI-PMH
- Bevorderen *semantische interoperabiliteit* door het ondersteunen van 'alignments'
- Stimuleren van Open *licenties*

CATCH



CONTINUOUS
ACCESS
TO
CULTURAL
HERITAGE
PLUS

Voordelen aanbieders

- Anderen je vocabulaires aanbieden kan met een simpele upload actie
 - Of via een eigen OpenSKOS instantie, of via OAI-PMH
- Het is mogelijk je eigen vocabulaire te gebruiken in andermans tools (al dan niet commercieel), als die tools OpenSKOS aanspreken
- Vocabulaires kunnen makkelijk en frequent door de bouwer/beheerder worden ge-update, zonder tussenkomst van derden
- Het is eenvoudig concepten uit je eigen vocabulaires te koppelen aan andere vocabulaires



Voordelen gebruikers

- Het is eenvoudig om vocabulaires te ontdekken, evalueren en te gebruiken
 - Minder noodzaak ze zelf te construeren
- Nieuwe mogelijkheden tot browsen en zoeken
 - In vocabulaires, in collectie-beschrijvingen
- Altijd up-to-date versies van vocabulaires beschikbaar



Voordelen toolbouwers

- Geen maatwerk per vocabulaire meer nodig, geen periodieke updates door te voeren
- Kunnen profiteren van werk van andere toolbouwers en vocabulaire-aanbieders
- Kunnen OpenSKOS gebruiken voor een reeks van nieuwe use cases



Ontstaansgeschiedenis



VAS, versie 1

- In eigen beheer gebouwd
 - Lourens van der Meij, VU
- Tekortkomingen
 - Ontbrekende functionaliteit
 - Twijfel aan schaalbaarheid
 - API definitie kon netter
- Desondanks succesvol
 - Jarenlang stabiel gedraaid en gebruikt
 - Toegepast tbv gebruik Beeld en Geluid thesaurus door Nationaal Archief (via Pictura's Memorix tool)



VAS, versie 2: OpenSKOS

- Gedurende half jaar intensief overlegd door CATCHPlus, RCE, Adlib, Pictura, Trezorix
 - Op initiatief van en gehost door RCE
 - Mede naar aanleiding van VAS, versie 1
- Doel: elkaars tools en omgevingen op elkaar aansluiten mbt vocabulaires/kennisstructuren
- Uitkomst: architectuur-concept en REST API *specificatie*
- OpenSKOS *implementatie* is helemaal conform de uitkomsten van dit overleg



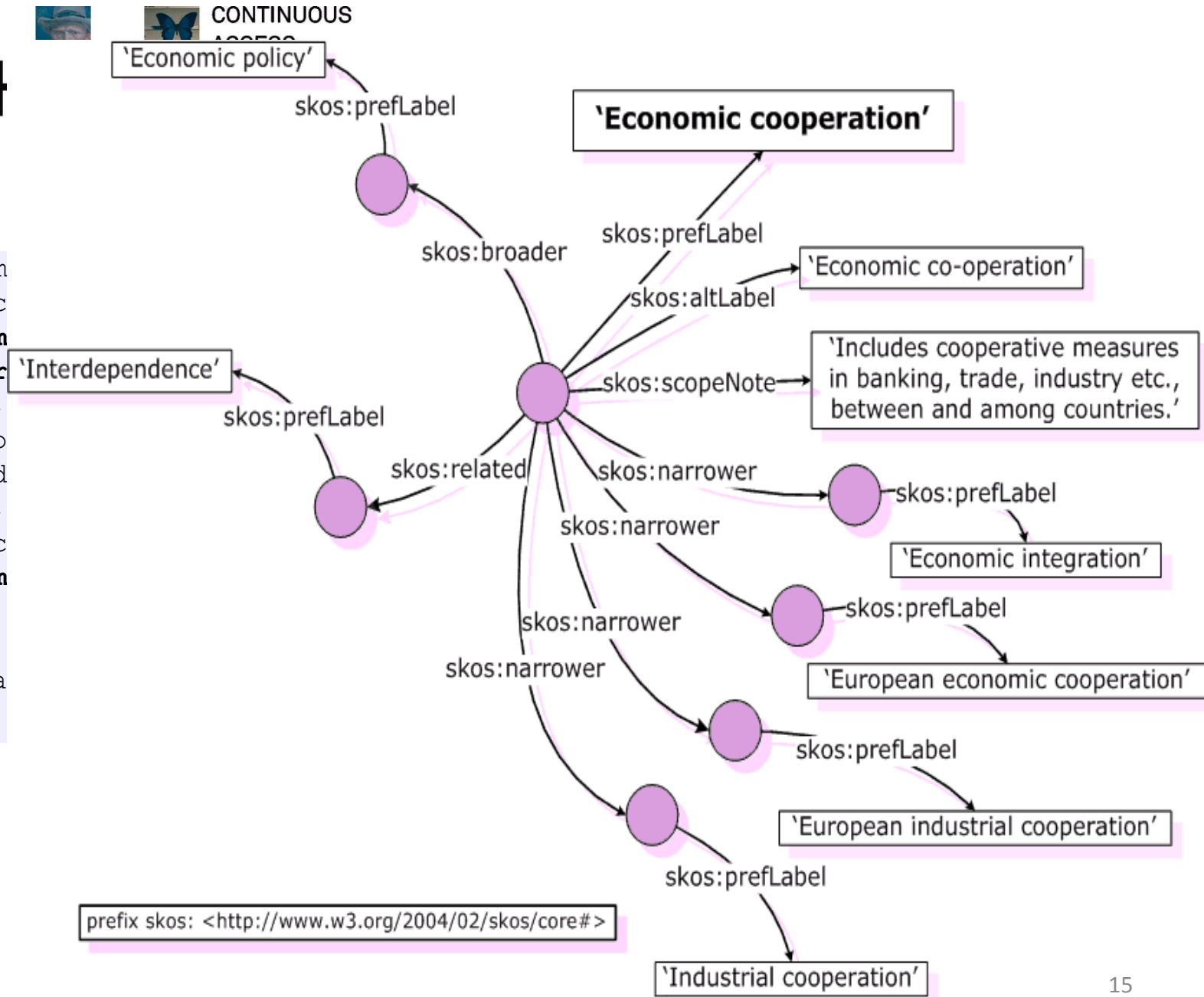
Technische acronymen

SKOS, een ultrakorte primer

- SKOS kennisstructuren bestaan uit Concepten, onderverdeeld in ConceptSchemes
- Concepten hebben een URL ter identificatie
- Concepten hebben labels in 1 of meer talen
- Concepten kunnen worden gedocumenteerd met 'notes'
- Concepten hebben onderling semantische relaties
 - broader, narrower, related

CATCH

Term: Econom
Used For: Ec
Broader term
Narrower term
 integration,
 European eco
 European ind
 cooperation,
 Industrial c
Related term
Scope Note:
 cooperative
 banking, tra
 between and



SKOS en het semantisch web

- Concepten kunnen over grenzen van ConceptSchemes worden gelinkt
 - exactMatch, closeMatch, broadMatch, narrowMatch, relatedMatch
- Mits een Concept een stabiele, resolvable URL heeft kan er *naar* gelinkt worden
- Concepten kunnen zelf linken naar (Linked Open Data op) het web



Nut/voordelen

- Standardisatie
- Sluit aan bij web
- Concept- versus term-gebaseerd

CATCH

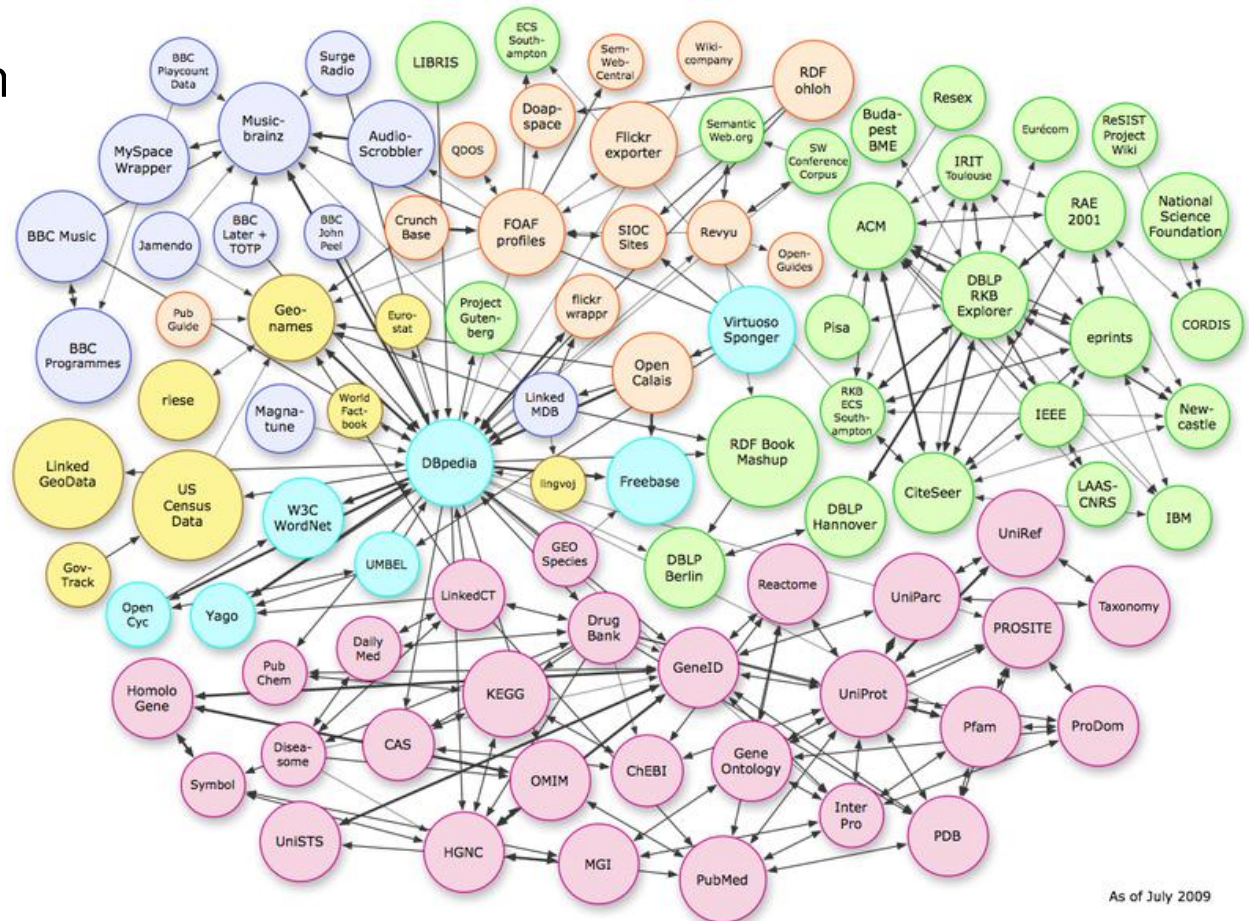


CONTINUOUS
ACCESS
TO
CULTURAL
HERITAGE
PLUS



Linked Open Data

- Een methode om data te tonen, delen en verbinden via 'dereferencable URIs' op het Web.





Nut/voordelen

- Eenvoudige, gestandaardiseerde manier van webtoegang tot data
- 'web van data'

RESTful web services

- Web service:
 - Software die interactie tussen machines over het web mogelijk maakt (“afspraken over vragen en antwoorden”)
- RESTful:
 - Gebruikt standaard protocollen van world wide web (http)
 - Beschrijft resultaten zoveel mogelijk als (online) ‘resources’
 - Aan te roepen via gewone URLs
 - In principe eenvoudig in gebruik en schaalbaar
- Voor gebruik vanuit software
 - Dus primair voor gebruik door programmeurs
 - Middleware
- Linked Data is een specifieke instantie van een RESTful web service



Nut/voordelen

- Standaard vorm van communicatie tussen machines
- Modulair, gelaagd
- Beter over verschillende machines en instellingen te verspreiden
- Allerlei verschillende toepassingen 'er bovenop' te bouwen

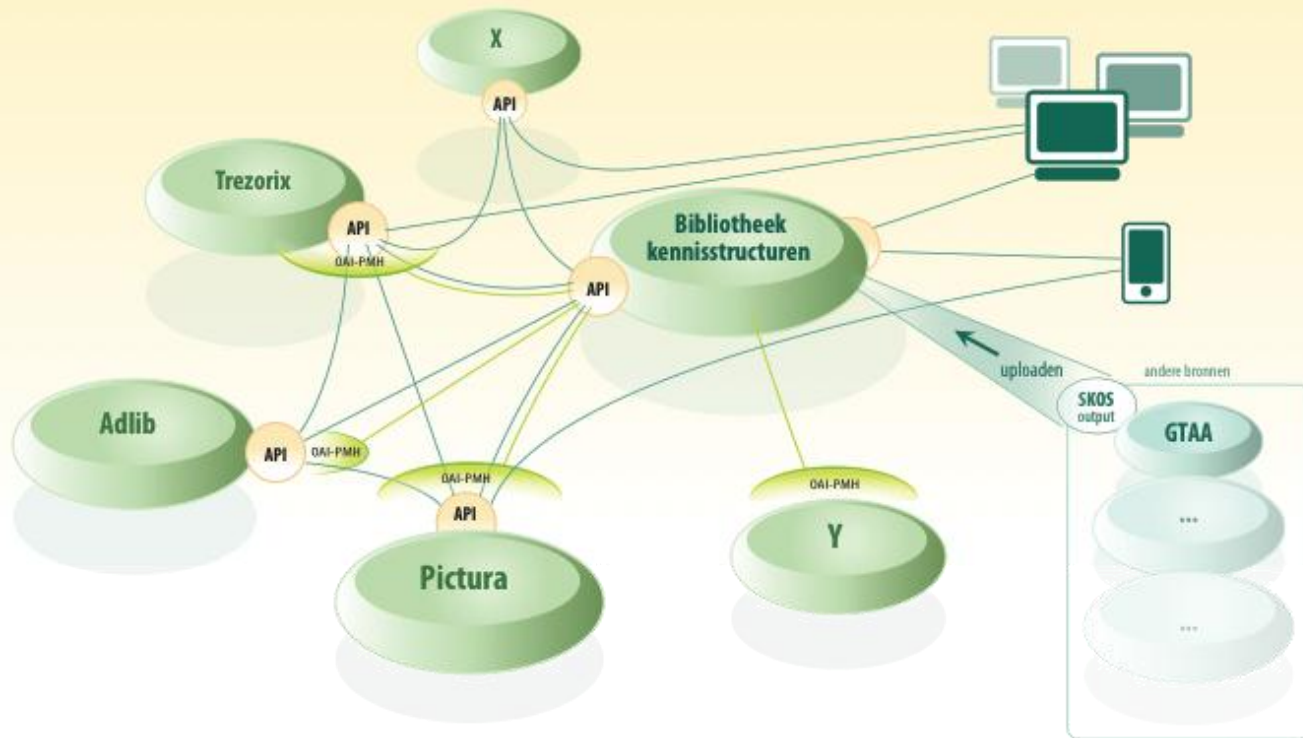


Functionele beschrijving en demo

Architectuur

Op weg naar een wereldstandaard.
Voor uitwisseling van kennisstructuren.

RESTful API
Dashboard voor interactieve toegang
OAI-PMH data provider en harvester
Upload module





RESTful API

- Concept resolve API
- Find API
- Autocomplete
- CRUD (create, retrieve, update, delete)
- Institutions
- Collections

CATCH



CONTINUOUS
ACCESS
TO
CULTURAL
HERITAGE
PLUS

Demo

- Concept resolve

- api/concept/a5994c9c-3368-9e52-494c-89b0fa8cd263.rdf
- api/concept?id=http://data.beeldengeluid.nl/gta/a/30135&format=html

CATCH



CONTINUOUS
ACCESS
TO
CULTURAL
HERITAGE
PLUS

Demo

- Find API
 - [api/find-concepts?q=crisis&format=rdf](http://api.find-concepts?q=crisis&format=rdf)
 - api/find-concepts?q=crisis+collection:4&format=json
 - api/find-concepts?q=crisis+tenant:rkd&format=json
 - api/find-concepts?q=prefLabelText:rutte+collection:4+inScheme:http*Persoonsnamen&format=json&fl=prefLabel,scopeNote



Demo

- Autocomplete API
 - <api/autocomplete/politi?returnLabel=prefLabel>
 - <api/autocomplete/speelf?returnLabel=prefLabel>

- Institutions, Collections API
 - in RDF: </api/institutions>
 - in HTML: </api/institutions?format=html>
 - </api/collections/beng:gtaa.html>



Demo

- Dashboard en upload module
 - <http://openskos.org/dashboard>

CATCH



CONTINUOUS
ACCESS
TO
CULTURAL
HERITAGE
PLUS

Demo

- OAI-PMH
 - [OAI voorbeelden](#)



Toepassingen in de praktijk



Client tools en diensten

- Bruikbaar voor CATCHPlus cases
- CLARIN(-NL) CLAVAS project
- Nationaal Archief gebruikt de API voor toegang tot de Beeld en Geluid GTAA thesaurus (via Memorix)
- Generieke browse- en zoek- web applicatie – Q42

CATCH Vocabulairebank Beta

Zoeken **Browse**

<ul style="list-style-type: none"> Gemeenschappelijke Thesaurus Audiovisuele Archieven Referentie Collectie Glas 	Classification	Zoeken op <input type="text" value="ku"/> Gezocht concepten <ul style="list-style-type: none"> kubisme kuikens kunst kunstacademies kunstdiefstal kunstenars kunstenars kunstenars kunstenarskolonies 	SKOS bredere <ul style="list-style-type: none"> 05E4 bedrijfseconomie en arbeid 12K0 kunst en cultuur algemeen beroepen SKOS nauwere <ul style="list-style-type: none"> architecten artiesten decorontwerpers diskjockeys fotografen grafisch kunstenaars 	SKOS bredere <ul style="list-style-type: none"> 12K1 beeldende kunst artistieke beroepen SKOS nauwere <ul style="list-style-type: none"> persfotografen SKOS gerelateerde <ul style="list-style-type: none"> foto's fotografie vrije beroepen
	Genre			
	GeografischeNamen			
	Maker			
	Namen			
	Onderwerpen			
	OnderwerpenBeeldGeluid			
Persoonsnamen				

Details

Concept: fotografen

Identifier
<http://data.beeldengeluid.nl/gtaa/28577>

Voorkeurstermen

Nederlands fotografen

SKOS bredere
[12K1 beeldende kunst](#)
[artistieke beroepen](#)

SKOS nauwere
[persfotografen](#)

Bewerk Foto's 902-8340

16 van 1157

Formulier **Geschiedenis** Auteursrechten

Serie_Collectie (webserie) Fotocollectie Anefo

Reportage/serie [onbekend]

Verbandtype

Beschrijving Kees de Bok in Artis. Publiek geeft hem voer

Annotatie

Begindatum (jjjj-mm-dd) 1948-07-08

Einddatum (jjjj-mm-dd) 0000-00-00

Datum aanduiding

Trefwoorden GTAA

Geogr. trefw. GTAA

Amsterdam Noord-Holland

Persoonsnamen GTAA

Namen GTAA **Artis**

Vervaardiger Schenkel, [...]

vervaardiger

Vervaardiger [onbekend]

Publicatie

Materiaalstype

Media Thesaurus

Memorix Vocabulairebank

Gemeenschappelijke Thesaurus Audiovisuele Archieven

Onderwerpen	Zoek op <input type="text"/>	Artis
GeografischeNamen	Artis	Voorkeursterm
Genre	Artishock	Artis
Namen	Artistieke staalkaart	Alternatieve term(en)
Persoonsnamen	Artists United Against Apartheid	Geen
Maker		ScopeNote
Classification		dierentuin Amsterdam
OnderwerpenBeeldGeluid		<input type="button" value="Koppel deze term"/>

Bewerk Foto's 902-8340

16 van 1157

Formulier **Geschiedenis** Auteursrechten

Serie_Collectie (webserie) Fotocollectie Anefo

Reportage/serie [onbekend]

Verbandtype

Beschrijving Kees de Bok in Artis. Publiek geeft hem voer

Annotatie

Begindatum (jjjj-mm-dd) 1948-07-08

Einddatum (jjjj-mm-dd) 0000-00-00

Datum aanduiding

Trefwoorden GTAA

Geogr. trefw. GTAA

Amsterdam Noord-Holland

Persoonsnamen GTAA

Namen GTAA

Vervaardiger

vervaardiger

Vervaardiger [onbekend]

Publicatie

Materiaalsoort

classificatie

Media Thesaurus

1 totaal 10 per pagina

pagina 1 van 1



1 totaal 10 per pagina

pagina 1 van 1

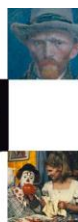


Content en beschikbaarheid

Content

- GTAA (via Beeld en Geluid)
 - 2 eigen OpenSKOS instanties
 - Open Database Licentie
- AAT-NED (via RKD)
 - Door RKD beschikbaar gesteld
 - Licentie?

CATCH



CONTINUOUS
ACCESS
TO
CULTURAL
HERITAGE
PLUS

Beschikbaarheid

- Broncode is onder Open source licentie en online beschikbaar (GitHub)
- We promoten Open Database licenties en Creative Commons
- We promoten vrij gebruik van services
 - M.u.v. schrijf-acties via API en Dashboard toegang
 - Iedereen kan eigen instantie van OpenSKOS draaien
- Pictura biedt 10 jaar hosting van OpenSKOS aan



Stakeholders en toekomstige ontwikkelingen



Stakeholders

- Adlib
- Beeld en Geluid
- Clarin/Clavas
- DEN
- KB
- Meertens
- Nationaal Archief
- Naturalis
- Pictura
- RCE
- RKD
- Trezorix
- Werkgroep erfgoedthesauri

Toekomst

- Linken met nieuwe “DEN inventarisatie van erfgoedthesauri”
- Uitbreiden van de inhoud
- Meerdere OpenSKOS instanties
- Beeld en Geluid: bouwen thesaurusbeheerapplicatie
- Pilot in overweging
 - RCE, Beeld en Geluid, Meertens
- Presentaties
 - KVAN dagen (juni 2012, met Pictura en BenG)
 - LREC conferentie Istanbul (mei 2012)



Basis/onderdeel van een (inter)nationaal platform?