

Continuous Access To Cultural Heritage

a computer science research programme



Véronique Malaisé

Luit Gazendam

Hennie Brugman

Guus Schreiber

Mettina Veenstra

Annemieke de Jong

Johan Oomen



**SKOS: a model for metadata
representation and interoperability –
Dutch Cultural Heritage Institution
thesaurus conversion use case**



Introduction

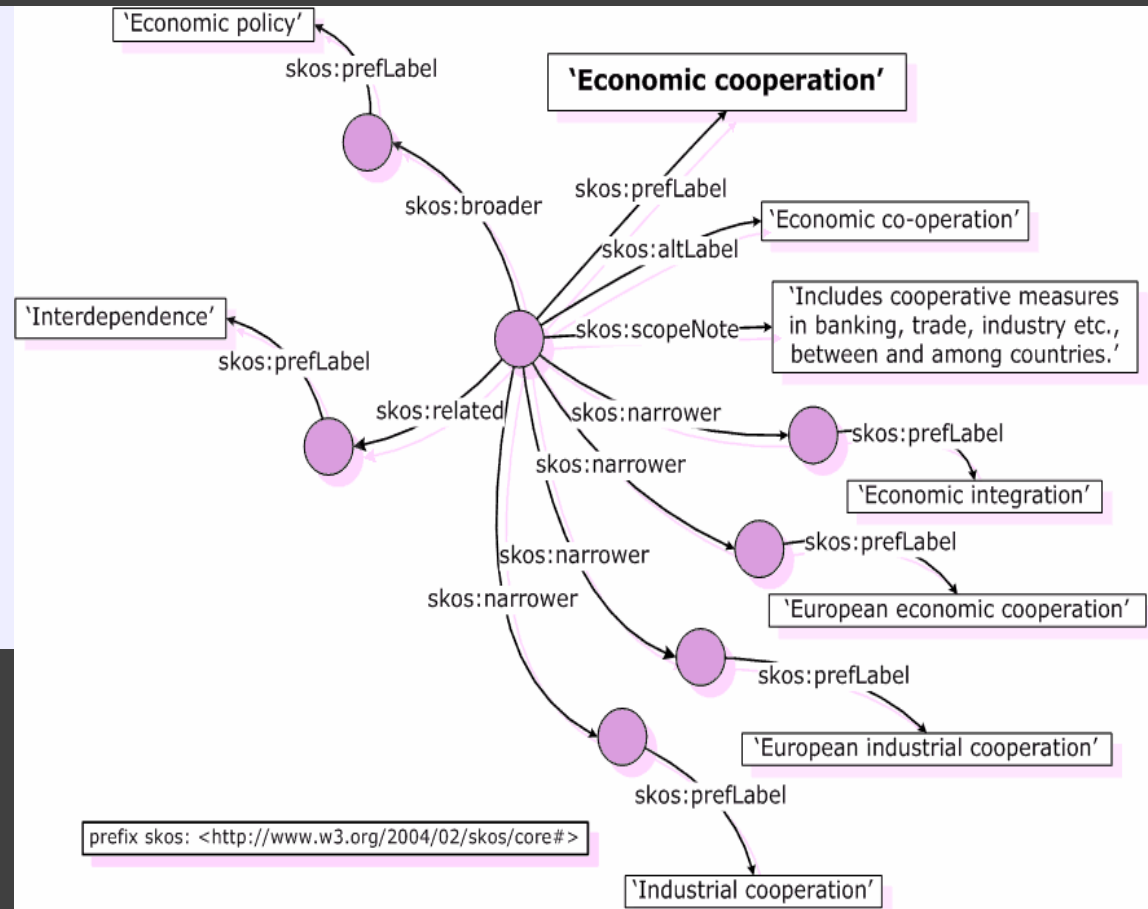
- Context: CATCH program, CHOICE project
- Collaboration between different research institutes and a Cultural Heritage Institution: the Netherlands Institute for Sound and Vision
- Focus: helping the cataloguers in their indexing process
- Central point of this activity: the GTAA thesaurus
 - How to represent it to enable its access within and outside of S&V?
 - How to represent it to enable reasoning and interoperability?





The SKOS model: Simple Knowledge Organization Systems

Term: Economic cooperation
Used For: Economic co-operation
Broader terms: Economic policy
Narrower terms: Economic integration, European economic cooperation, European industrial cooperation, Industrial cooperation
Related terms: Interdependence
Scope Note: Includes cooperative measures in banking, trade, industry etc., between and among countries.





The GTAA thesaurus

- Faceted thesaurus: Subject, Person, Name, Location, Genre, Maker
- Standard constructs: broader term/narrower term, related term, use/use for, scope note
- Standard construct not yet standardised in SKOS: qualifiers
- Non standard features: Categories grouping different Subjects and relationships between terms across facets



Conversion method: see [van Assem *et al.* ESWC2006]

- Interpret the intended meaning of the different constructs (some are not instantiated in the GTAA)
- Check their counterpart in SKOS
 - Broader term <-> skos:broader, narrower term <-> skos:narrower, related term <-> skos:related, scope note <-> skos:scopeNote
 - Facets <-> sub-concepts of skos:Concept, all terms are instances of these concepts
 - Categories <-> skos:Collection/skos:member
 - Qualifiers: open question!
- Transform the text file into RDF: accessing the different elements of plain text and transforming them into RDF-OWL structured data

Conversion example 1: Subject facet

- aalmoezeniers
 - US geestelijk verzorgers
- aalscholvers
 - 1D14.01
 - BT watervogels
- aandeelhouders
 - 1D05.00
 - 1D05.04
 - RT beleggingen
 - RT ondernemingen
 - RT waardepapieren
- aandelen
 - US waardepapieren
- aankomst en vertrek
 - 1D10.00
 - RT reizen
 - UF vertrek
 - SN NB: op de rode lijst

```
<Subject rdf:about="&gtaa;aalscholvers">
<skos:prefLabel>aalscholvers_</skos:prefLabel>
<skos:broader rdf:resource="#watervogels_"/>
</Subject>
```

```
<Subject rdf:about="&gtaa;aandeelhouders">
<skos:prefLabel>aandeelhouders_</skos:prefLabel
>
<skos:related rdf:resource="#beleggingen"/>
<skos:related rdf:resource="#ondernemingen"/>
<skos:related rdf:resource="#waardepapieren"/>
</Subject>
```

```
<Subject
  rdf:about="&gtaa;aankomst_en_vertrek">
<skos:prefLabel>aankomst_en_vertrek</skos:pref
Label>
<skos:related rdf:resource="#reizen"/>
<skos:altLabel>vertrek_</skos:altLabel>
<skos:scopeNote>NB_: op de rode lijst
  </skos:scopeNote>
</Subject>
```

Conversion example 2: Person facet


- A.F.Th. [zie] Heijden, A.F.Th. van der
- Aa, F. vd
- Aa, Frans vd {acteur}
- Aa, Herman vd {gitarist}
- Aa, Jaap vd {acteur}
- Abdullah (koning Jordanië) {zie ook}
 - Abdullah (kroonprins Jordanië)
- Abdullah (kroonprins Jordanië) {zie ook}
 - Abdullah (koning Jordanië)

```

<PersonName rdf:about="#PersonName_A.F.Th._">
<skos:prefLabel>A.F.Th._</skos:prefLabel>
<skos:related rdf:resource="#PersonName_Heijden,_A.F.Th._van_der_"/>
</PersonName>
<PersonName rdf:about="#PersonName_Aa,_F._vd_">
<skos:prefLabel>Aa,_F._vd_</skos:prefLabel>
</PersonName>
<PersonName rdf:about="#PersonName_Aa,_Frans_vd_">
<skos:prefLabel>Aa,_Frans_vd_</skos:prefLabel>
<skos:scopeNote>acteur_</skos:scopeNote>
<linkedTerm rdf:resource="#Subject_Acteurs"/>
</PersonName>
<PersonName rdf:about="#PersonName_Aa,_Herman_vd_">
<skos:prefLabel>Aa,_Herman_vd_</skos:prefLabel>
<skos:scopeNote>gitarist_</skos:scopeNote>
<linkedTerm rdf:resource="#Subject_Guitarists"/>
</PersonName>
<PersonName rdf:about="#PersonName_Aa,_Jaap_vd_">
<skos:prefLabel>Aa,_Jaap_vd_</skos:prefLabel>
<skos:scopeNote>acteur_</skos:scopeNote>
<linkedTerm rdf:resource="#Subject_Acteurs"/></PersonName>
<PersonName rdf:about="#PersonName_Abdullah__koning_Jordanië__">
<skos:prefLabel>Abdullah__koning_Jordanië__</skos:prefLabel>
<skos:altLabel>Abdullah__kroonprins_Jordanië_</skos:altLabel>
<linkedTerm
  rdf:resource="#PersonName_Abdullah__kroonprins_Jordanië__"/>
<skos:scopeNote>zie_ook_Abdullah__kroonprins_Jordanië__</skos:scopeNote>
</PersonName>
<PersonName
  rdf:about="#PersonName_Abdullah__kroonprins_Jordanië__">
<skos:prefLabel>Abdullah__kroonprins_Jordanië__</skos:prefLabel>
...

```

Interoperability possibility: the Thesaurus browser

BEELD EN GELUID  GTAA Web Browser (VERSION AUGUST 3 2006)

Onderwerpen | Genres | Personen | Namen | Makers | Geografische Namen

Zoeken

Case sensitive Filter Spelling
 In Selekteer Synoniemen

Rubrieken

[11 communicatie en media](#)

- ⊕ (2) [01 levensbeschouwing](#)
- ⊕ (6) [02 maatschappij](#)
- ⊕ (5) [04 ruimtelijke omgeving](#)
- ⊕ (34) [05 economie](#)
- ⊕ (6) [06 gezondheid](#)
- ⊕ (6) [07 opvoeding en onderwijs](#)
- ⊕ (9) [08 wetenschap](#)
- ⊕ (31) [09 techniek](#)
- ⊕ (6) [10 verkeer en vervoer](#)
- ⊖ (406) [11 communicatie en media](#)
 - (96) [11.00 communicatie en media - a](#)
 - (37) [11.01 communicatietransport](#)
 - (168) [11.02 massacommunicatie en -](#)
 - (31) [11.03 informatievoorziening ove](#)
 - (74) [11.04 audiovisuele en elektronis](#)
- ⊕ (90) [12 kunst en cultuur](#)
- ⊕ (9) [13 sport en vrijetijdsbesteding](#)
- ⊕ (10) [14 natuur en milieu](#)
- ⊕ (4) [15 aardrijkskunde](#)
- ⊕ (19) [03 bestuur en geschiedenis](#)

Termen in Rubrieken

ICT
 IT
[aanplakbiljetten](#)
[abonneeradio](#)
[abonneetelevisie](#)
[actualiteitenprogrammas](#)
[advertenties](#)
[affiches](#)
[akkoorden](#)
[aktes](#)
[alarmnummers](#)
[alfabetten](#)
[amateurfilms](#)
[amateuropnamen](#)

◀ 1 2 3 4 5 6 7 8 9 10 ▶

Term Info

affiches

Used for:
 aanplakbiljetten
 posters

Related terms:
[reclame](#)
[grafische kunsten](#)
[pamfletten](#)

Boom

- ⊖ publicaties
 - [advertenties](#)
 - [affiches](#)
 - [boeken](#)
 - [columns](#)
 - [kranten](#)
 - ⊕ [naslaqwerken](#)
 - [recensies](#)
 - [strips](#)
 - ⊕ [tijdschriften](#)
 - [websites](#)