

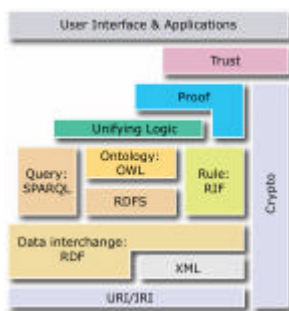
## RDFa & SKOS

Antoine Isaac  
Semantic Web SIKS Course  
25 Sept 2009

## Me

- Vrije Universiteit Amsterdam
- Cultural Heritage-related projects
  - STITCH, TELplus, EuropeanaConnect
- Member of W3C Semantic Web Deployment working group
  - Working on SKOS
- <http://www.few.vu.nl/~aisaac>

## This talk: beyond RDF, OWL and rules?



- Representing data in RDF is good
- Reasoning with it is excellent
- But there is more in the Semantic Web initiative!

## W3C activities beyond RDF, OWL and rules

### *Identifying (Best) Practices*

- Publishing data
  - cf. Open Linked Data initiative
- Identifying relevant ontologies
- Outreach
  - Facilitating technology adoption
- Re-using and linking to existing knowledge
  - Legacy data

## Semantic Web Deployment W3C Working Group

- Best practices for publishing vocabularies
  - *Practices for serving ontologies*
- SKOS
  - *An ontology*
- RDFa
  - *A syntax to express RDF data*

## RDFa

## Problem (1)

- Many data in HTML web pages
  - Esp. those dynamically generated from databases
  - Flickr, IMDb, LinkedIn...
- But it is human-readable only
- Could we publish that data and still make it explicit for re-use by machines?

*Well, it's a bit what the Semantic Web is about, isn't it?  
So let's just publish that data as RDF file or SPARQL endpoint...*

## Problem (2)

- This implies publishing data in two forms:
  - document web (HTML page)
  - data web (RDF data)
- Duplication of efforts
  - HTML presentation structure often mirrors data structure
  - Many data values also occur in HTML pages

The image shows a screenshot of a web page for Antoine Isaac, a member of the Faculty of Sciences at Vrije Universiteit Amsterdam. The page contains a profile picture and contact information. Overlaid on the right side of the page is a block of RDF data in XML format, which encodes the same information as the HTML page but in a machine-readable format.

## RDFa goal

- Avoiding duplication would be great for many cases
  - For information publishers
  - For information consumers
- RDFa's goal is to publishing both human-readable and machine-accessible data along a same channel
- Simple idea: embedding RDF data on normal web pages
  - Using their presentation structure

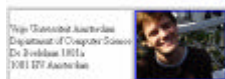
## RDFa basic facts

- RDFa enables embedding RDF facts in (X)HTML pages
- RDFa uses specific XHTML attributes for expressing RDF data as metadata on XHTML elements
- RDFa is another serialization syntax for RDF
  - As RDF/XML, N3 and Turtle are, though RDFa is less complete

## Running Example

- Adding RDFa to simple bio/contact page

### Antoine Isaac



I am a postdoc working for the [EuropeanConnect](#) project. I am attached to the [Web & Media Group of the Vrije Universiteit](#).

## Original XHTML page

```
<div xmlns="http://www.w3.org/1999/xhtml">
<h1>Antoine Isaac</h1>
<table style="text-align: left" borders="1">
<tbody><tr>
<td>
<div>Vrije Universiteit Amsterdam</div>
<div>Department of Computer Science</div>
<div>De Boelelaan 1081a</div>
<div>1081 HV Amsterdam</div>
</td>
<td><a href="http://www.few.vu.nl/~aisaac/Antoine.jpg">

</a>
</td>
</tr></tbody>
</table>
<p>I am a postdoc working for the <a href="http://www.europeanconnect.eu">EuropeanConnect</a>
project. I am attached to the <a href="http://wiki.cs.vu.nl/web-media">Web & Media Group of the
Vrije Universiteit</a>.</p>
</div>
```



### Step 1: flagging page as RDFa-enabled

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML+RDFa 1.0/EN"
"http://www.w3.org/MarkUp/DTD/xhtml-rdfa-1.dtd">

<div xmlns="http://www.w3.org/1999/xhtml">
<h1>Antoine Isaac</h1>
[...]
```



### Step 2: declaring ontologies

```
<div xmlns="http://www.w3.org/1999/xhtml"
xmlns:foaf="http://xmlns.com/foaf/0.1/"
xmlns:v="http://rdf.data-vocabulary.org/#">
<h1>Antoine Isaac</h1>
[...]
```

Ontology declarations are valid throughout the XHTML element where they are declared

### Step 3: subjects of triples to come

```
<div xmlns="http://www.w3.org/1999/xhtml"
xmlns:foaf="http://xmlns.com/foaf/0.1/"
xmlns:v="http://rdf.data-vocabulary.org/#"
about="http://www.few.vu.nl/~aisaac/foaf.rdf#me">
<h1>Antoine Isaac</h1>
[...]
```

The resource introduced by the `about` will be the subject of the RDF statements firstly encountered down the XHTML sub-tree

### Step 4: RDF statements using text values

```
<div [...] about="http://www.few.vu.nl/~aisaac/foaf.rdf#me">
<h1 property="v:name">Antoine Isaac</h1>
<table style="text-align: left" border="1">
<tbody><tr>
<td>
<div property="v:affiliation">Vrije Universiteit Amsterdam</div>
<div>Department of Computer Science</div>
[...]
```

The text content of the node where `property` is used serves as literal object of the RDF statement derived from the page:

```
<http://www.few.vu.nl/~aisaac/foaf.rdf#me
v:name "Antoine Isaac" ;
v:affiliation "Vrije Universiteit Amsterdam" .
```

### Step 5: RDF statements using hyperlinks

```
<div [...] about="http://www.few.vu.nl/~aisaac/foaf.rdf#me">
[...]
```

The `href` of the node where `rel` is used serves as resource object:

```
<http://www.few.vu.nl/~aisaac/foaf.rdf#me v:photo
<http://www.few.vu.nl/~aisaac/Antoine.jpg> .
```

### Nested descriptions

```
<div [...] about="http://www.few.vu.nl/~aisaac/foaf.rdf#me">
[...]
```

The use of property "nested" in a `rel` without hyperlink denotes the existence of a (blank) resource being described:

```
<http://www.few.vu.nl/~aisaac/foaf.rdf#me v:address
[
v:street-address "De Boelelaan 1081a"
] .
```

## Adding XHTML structure when not present

```
<div [...] about="http://www.few.vu.nl/~aisaac/foaf.rdf#me">
[...]
<span rel="v:address"><div property="v:street-address">De Boelelaan 1081a</div>
<div><span property="v:postal-code">1081 HV</span> <span
property="v:locality">Amsterdam</span></div>
</span>
[...]
<p>I am a <span property="v:role">postdoc</span> working [...>
</div>
```

span can be used to create a new structure layer inside the existing one:

```
<http://www.few.vu.nl/~aisaac/foaf.rdf#me>
v:address [
    v:street-address "De Boelelaan 1081a" ;
    v:postal-code "1081 HV"
] ;
v:role "postdoc" .
```

## What does it buy me? – RDF data in my page!

```
sprefix foaf: <http://xmlns.com/foaf/0.1/> .
sprefix v: <http://rdf.data-vocabulary.org/#> .
<http://www.few.vu.nl/~aisaac/foaf.rdf#me> a v:person ;
v:address
[
v:locality "Amsterdam" ;
v:postal-code "1081 HV" ;
v:street-address "De Boelelaan 1081a"
] ;
v:affiliation "Vrije Universiteit Amsterdam" ;
v:name "Antoine Isaac" ;
v:photo <http://www.few.vu.nl/~aisaac/Antoine.jpg> ;
v:role "postdoc" ;
foaf:currentProject <http://www.europeanconnect.eu> ;
foaf:workplacehomepage <http://wiki.cs.vu.nl/web-media> .
```



## Important (coming) implementations

- Consuming RDFa
  - Yahoo! SearchMonkey
  - Google Rich Snippets
- Producing RDFa
  - Drupal
- Publishing sites
  - Government sites
  - Etc.

### Google Rich Snippets

Google webmaster tools

#### Rich Snippets Testing Tool <sup>beta</sup>

Rich Snippet allows you to enhance your Google search results by marking up web pages with Microformats or RDFa.

Test your website

Enter a web page URL to see how it may appear in search results:

Help with:  
[Documentation](#)  
[Tips & Tricks](#)

#### Google search preview

Amsterdam - postdoc at Vrije Universiteit Amsterdam  
Excerpt from the page will show up here. Excerpt from the page will show up here.  
www.few.vu.nl/~aisaac/exampleRDFa.html - [Cached](#) - [Similar pages](#)

Note that there is no guarantee that a Rich Snippet will be shown for this page on actual search results. For more details, see [Rich Snippets](#).

#### Extracted Rich Snippet data from the page

Person  
address  
street-address = De Boelelaan 1081a  
postal-code = 1081 HV  
locality = Amsterdam  
photo = Antoine  
currentProject = EuropeanConnect  
workplacehomepage = Web & Media Group of the Vrije Universiteit  
name = Antoine Isaac  
affiliation = Vrije Universiteit Amsterdam  
role = postdoc

<http://www.google.com/webmasters/tools/richsnippets>

## RDFa take-home message

- Making (X)HTML publication a more efficient and rewarding action
- By publishing structured data in pages, next to the presentation structure

## SKOS

## Knowledge organization out there



- The SW is not everything
- Libraries and many other institutions routinely create metadata

## "Traditional Metadata"

Titel: [The Cambridge companion to Molière](#) / ed. by David Bradby, Andrew Calder  
 Medewerker: [David Bradby; Andrew Calder \(1942-\)](#)  
 Jaar: 2006  
 Uitgever: Cambridge [etc.] : [Cambridge University Press](#)  
 Reeks: [Cambridge companions to literature](#)  
 Annotatie: Met lit. opg. en index  
 Omvang: XIX, 242 p. : ill. ; 24 cm  
 ISBN: [978-0-521-54665-6](#); [978-0-521-54665-6](#); [978-0-521-54665-6](#); [978-0-521-54665-6](#)  
 Trefwoord GOO: [Handboeken \(vorm\)](#) [1600-1700](#)  
 Trefwoord persoon: [Molière, Jean Baptiste Poquelin, van](#)  
 Basisclassificatie: [J8.25 Franse letterkunde](#)

- Description of objects and their content
- Using controlled **Knowledge Organization Systems**
  - Thesauri, classification systems, subject heading lists...
- Knowledge organization has been done for decades (centuries?)
  - Information science

## KOS example: Dutch Basic Classification

WETENSCHAP EN CULTUUR IN HET ALGEMEEN	
02.00	<b>Wetenschap en cultuur in het algemeen</b>
02.01	• Geschiedenis van wetenschap en cultuur Hoof: <i>instructie geschiedenis</i>
02.02	• Filosofie en theorie der wetenschap Is: 08.35 (filosofie, wetenschapsfilosofie) Hoof: <i>wetenschapsfilosofie, algemene semiotiek</i> Verwijz: <i>voor de semiotiek van afzonderlijke wetenschapsgebieden, zie de betreffende vakgebieden</i>
02.10	• Wetenschap en samenleving
02.11	• Wetenschapssociologie
02.12	• Kennissociologie
02.13	• Wetenschapsbeoefening Hoof: <i>normaalsite</i> Tedecl: <i>methodes en technieken van wetenschappelijk onderzoek</i>
02.14	• Organisatie van wetenschap en cultuur Hoof: <i>organisatie en instellingen, academies van wetenschappen, verenigingsinstellingen</i> Verwijz: <i>voor organisaties of instellingen die zich met afzonderlijke vakgebieden bezighouden, zie de betreffende vakgebieden</i>
02.15	• Wetenschapsbeleid, cultuurbeleid Hoof: <i>wetenschappelijke prijzen toegekend door overheden</i>
02.16	• Wetenschappelijke samenwerking, culturele samenwerking
02.20	• Wetenschapsvoorsichting Hoof: <i>wetenschapsjournalistiek</i>
02.30	• Museologie Verwijz: <i>voor algemene musea, zie: 02.14 (organisatie van wetenschap en cultuur)</i> Verwijz: <i>voor musea van afzonderlijke vakgebieden, zie de betreffende vakgebieden</i>

## KOS example: Iconclass

**Browse by subject**  
using the Iconclass classification system

0 Abstract, Non-representational Art

1 Religion and Magic

2 Nature

25 earth, world as celestial body

25F animals [\[show images >25\]](#)

25F3 birds [\[show images >25\]](#)

25F31 groups of birds [\[show images < 5\]](#)

25F32 song-birds [\[show images >25\]](#)

25F33 predatory birds [\[show images >25\]](#)

25F34 owls [\[show images < 25\]](#)

25F35 ornamental birds [\[show images < 25\]](#)

25F36 water-birds [\[show images >25\]](#)

25F37 shore-birds and wading-birds [\[show images >25\]](#)

25F38 walker and runner birds [\[show images < 25\]](#)

25F39 other birds [\[show images >25\]](#)

3 Human Being, Man in General

4 Society, Civilization, Culture

5 Abstract Ideas and Concepts

6 History

7 Bible

8 Literature

9 Classical Mythology and Ancient History

ICONCLASS © Royal Netherlands Academy of Arts and Sciences

Show only notations used in the manuscript database  
 Show full Iconclass hierarchy

**Search by keyword**

## Knowledge Organization systems

- (dozens of) thousands of concepts
  - DDC, AAT, LCSH, GOO, Iconclass...
- Loose semantics – but still, semantics!
  - **Car wheel BT Car** (BT = broader term)
- Proven to be useful for applications
  - Search, description
- Obviously it would be great to be able to represent those legacy/external KOSs on the SW
  - But how??

## SKOS background

- There are many KOS models and formats
- But also common features and application requirements
  - Lexical information, semantic links
- W3C Semantic Web Deployment working group
- Data model to represent KOSs in a *simple* way
- SKOS: Simple Knowledge Organization System



## Use Cases and Requirements

- Gathering use cases for SKOS
  - Existing or anticipated applications
  - E.g., "Semantic search service across mapped multilingual thesauri in the agriculture domain"
- From use cases, requirements were elicited
  - E.g., using generalization links between concepts (can be used for hierarchical browsing)

[SKOS Use Cases and Requirements](#)  
 W3C Working Group Note 18 August 2009  
This version: <http://www.w3.org/TR/2009/NOTE-skos-ucr-20090818/>  
 Latest version: <http://www.w3.org/TR/skos-ucr/>  
 Previous version: <http://www.w3.org/TR/2007/NOTE-skos-ucr-20070516/>  
 Editors: Ardoine Isaac, Vrije Universiteit Amsterdam, [aisa@few.vu.nl](mailto:aisa@few.vu.nl)  
 Jon Phipps, Cornell University, [jphipps@mac.cornell.com](mailto:jphipps@mac.cornell.com)  
 Daniel Rubin, Stanford Medical Informatics, [drubin@stanford.edu](mailto:drubin@stanford.edu)

## SKOS basis

- SKOS offers a vocabulary to create RDF data about:
  - **Concepts** and **ConceptSchemes**
  - Lexical properties
  - Semantic relations
  - Notes

## Thesaurus example

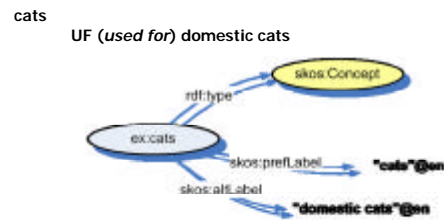
**animals**  
 NT (*narrower term*) cats

**cats**  
 UF (*used for*) domestic cats  
 RT (*related term*) wildcats  
 BT (*broader term*) animals  
 SN (*scope note*) used only for domestic cats

**domestic cats**  
 USE cats

**wildcats**

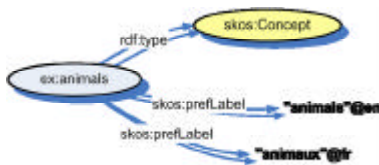
## Concepts and labels



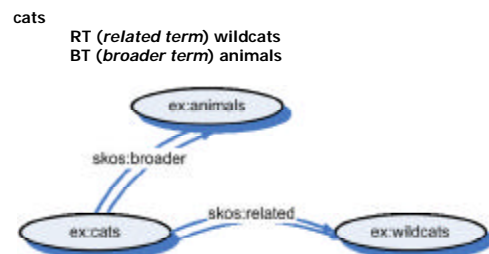
skos: = <http://www.w3.org/2004/02/skos/core#>  
 rdf: = <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

- SKOS is concept-oriented
  - Concepts are first-order resources

## Multilingual labels



## Semantic relations

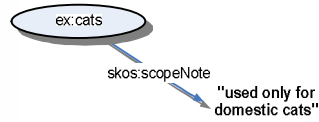


## Documenting concepts

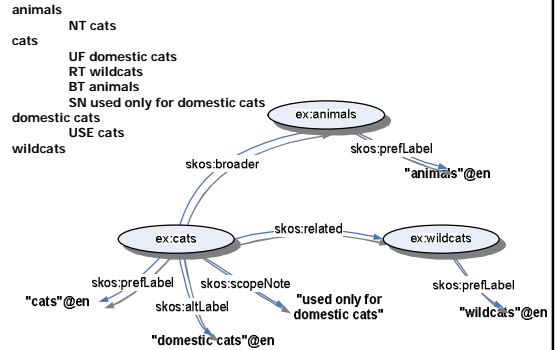
`skos:note`

```

|-- skos:definition
|-- skos:scopeNote
|-- skos:example
|-- skos:historyNote
|-- skos:editorialNote
|-- skos:changeNote
  
```



## Putting it together

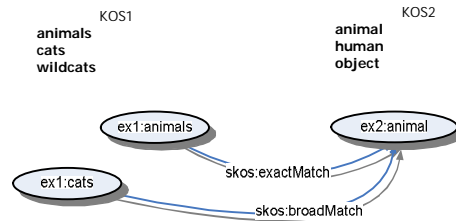


## Other features

- Explicit representation of Concept Schemes
  - `skos:ConceptScheme`, `skos:inScheme`
- Concept grouping
  - `skos:Collection`, `skos:member...`

## Networking controlled vocabularies in SKOS

- Matching properties as specific conceptual relations



## The relevance of networked KOSs



Johan Stapel, Koninklijke Bibliotheek, Den Haag

## SKOS semantics

- SKOS is an OWL ontology
- It has formal axioms
  - Domains and ranges
  - **broader** and **narrower** are inverse of each other
  - **related** is symmetric
  - ...
- But the semantics of KOSs are quite loose
- Axioms may be less rich than expected (from an OWL perspective!)

## SKOS and OWL?

- “OWL is a Harley-Davidson, SKOS is a mountain bike” — (Tom Baker, co-chair of the SW Deployment WG)
- SKOS and OWL are meant for quite different things
- SKOS = Model to represent KOSs in a *simple* way
  - Ontology for **concepts** – the elements in (CH) vocabularies
  - Not an ontology for OWL classes!
  - *Model vs. meta-model*

## Deployment Examples

- FAO Agrovoc
  - <http://www.fao.org/agrovoc>
  - Application: cf. <http://www.fao.org/ag/agn/publications/fna>
- NASA:
  - <http://nasataxonomy.jpl.nasa.gov/fordevelopers/>
- Etc.
- See SKOS implementation report
  - <http://www.w3.org/2006/07/SWD/SKOS/reference/20090315/implementation.html>

Serving interlinked KOSs:  
From RAMEAU  
(French Library)  
to LCSH  
(US Library)  
and back

**Concept information for Oiseau**

URI: <http://id.loc.gov/authorities/ah8014310#concept>

Label: Oiseau

Alternate labels: Avifaune, Oiseau

Broader terms: Animalia, Vertébrés, Oiseau, Oiseau incapable de voler

<http://stitch.cs.vu.nl/rameau>

Mappings (simple SKOS statements)

Mapping Relation	Concept
closeMatch	<a href="http://id.loc.gov/authorities/ah8014310#concept">http://id.loc.gov/authorities/ah8014310#concept</a>

Serving interlinked KOSs:  
From RAMEAU  
(French Library)  
to LCSH  
(US Library)  
and back

**Authorities & Vocabularies**

URI: <http://id.loc.gov/authorities/ah8014310#concept>

Type: Topical Term

Alternate Labels: Aves; Avian fauna; Avifauna

Broader Terms: Animalia, Vertébrés

Narrower Terms: Albatros birds, Galliformes, Accipitriformes, Scolopaciformes, Falconiformes, The American Heritage dict. of the Engl. lang., via WWW, Aug. 31, 2001 (a specific region or period), LC Database, Aug. 31, 2001 (avifauna; Aves; avian fauna)

LC Classification: QL671

Created: 2003-08-22

Last Modified: 2009-10-06 13:01:40

Similar concepts from other vocabularies:

- <http://stitch.cs.vu.nl/vocabularies/rameau/ah8014310#concept>

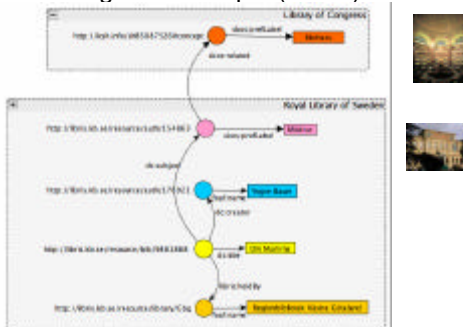
<http://id.loc.gov>

## Library case: Libris

- <http://libris.kb.se/>
- Swedish Library Catalogue as linked data

LIBRIS: The difference engine / William Gibson & Bruce Sterling

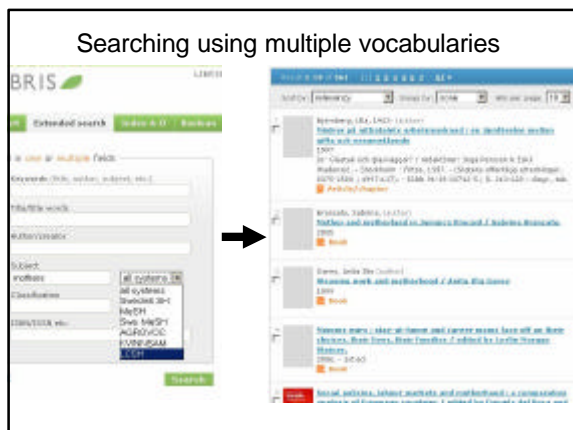
## Links between Libris and Library of Congress concepts (LCSH)



Ed Summers et. al., Dublin Core 2008  
<http://dc2008.de/wp-content/uploads/2008/09/summers-isaac-redding-krech.pdf>



## Searching using multiple vocabularies



The image shows a screenshot of the BRIS library search interface. On the left, there is a search form with various filters: 'Multiple fields', 'Multiple words', 'All languages', 'Classification', 'Subject', 'Classification', 'Date range', and 'Country'. A black arrow points from the search form to the search results on the right. The search results list several books, including 'Nürnberg, Uta. Land - (Aktion)' and 'Nürnberg, Uta. Land - (Aktion)'. The interface is in German and includes a search bar at the top with the text 'Suche in BRIS'.

## SKOS take-home message

Porting, linking and exploiting on the SW a basic but massive amount of organized knowledge



Thanks!

## References

- RDFa Specification  
<http://www.w3.org/TR/rdfa-syntax/>
- RDFa community  
<http://rdfa.info/wiki/>
- RDFa distiller  
<http://www.w3.org/2007/08/pyRdfa>
- SKOS site  
<http://www.w3.org/2004/02/skos>
- SKOS Reference  
<http://www.w3.org/TR/skos-reference>
- SKOS Primer  
<http://www.w3.org/TR/skos-primer>