

Recipes for Semantic Web Dog Food – The ESWC and ISWC Metadata Projects

Knud Möller¹

Tom Heath²

Siegfried Handschuh¹

John Domingue²

¹DERI, NUI Galway and ²KMi, The Open University



“Eating One’s Own Dog Food”

“using one’s own products to gain first hand experience of their fitness for purpose”

Overview of this Talk

- Previous Servings of Dog Food
 - ESWC2006, ISWC2006, (ESWC2007)
- Lessons Learned
 - Recipes for Semantic Web Dog Food, with focus on ISWC+ASWC 2007

Contributions of this Paper

- Accounts of real deployments
- Analysis of issues encountered
- Lessons learned and recommendations
- Today's special dish: a report on the latest serving

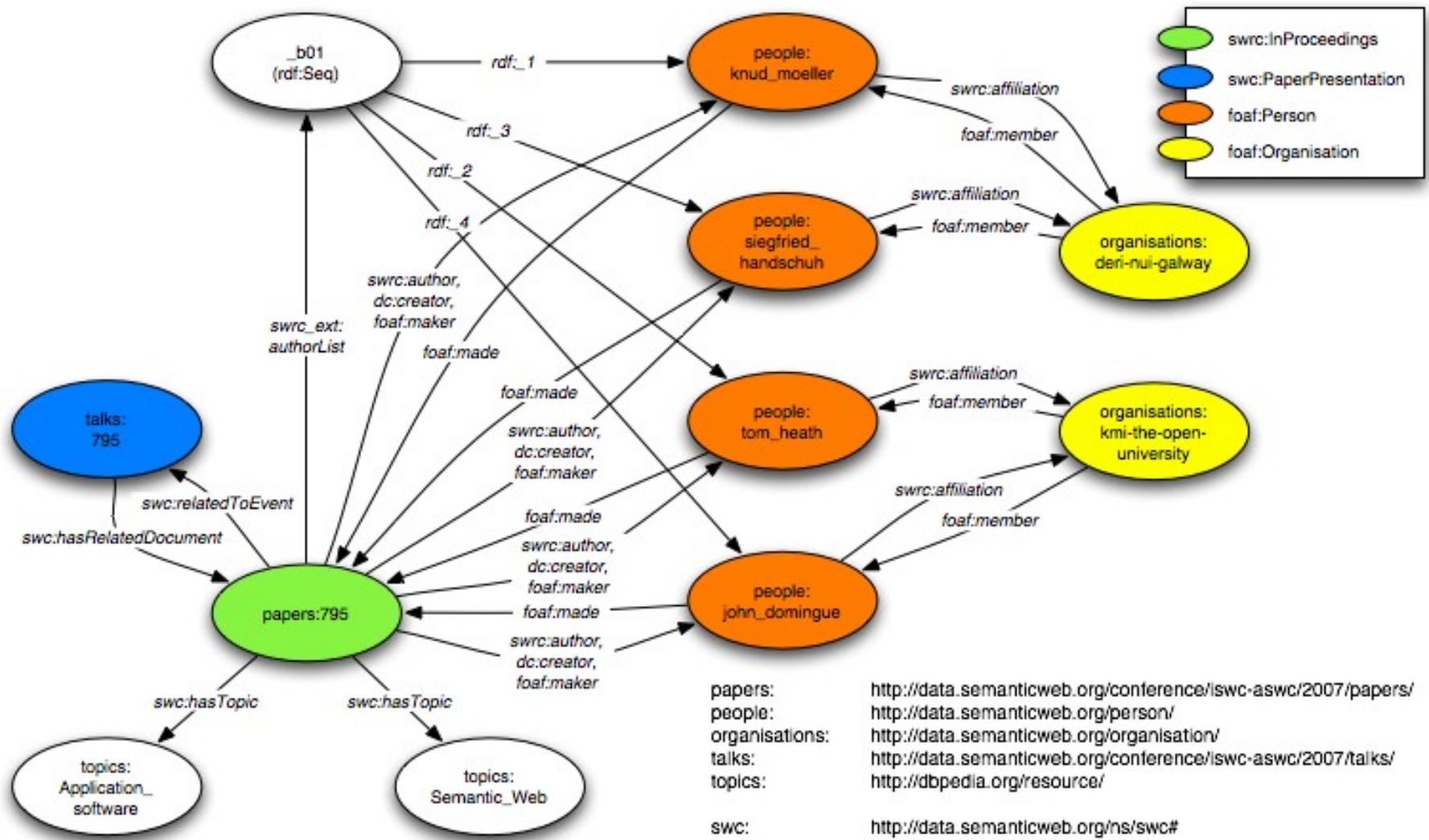
The ESWC2006 and ISWC2006 Metadata Projects

- Goals:
 - showcasing, learning, evaluation, insight
- Outcomes:
 - conference domain modelled to produce a dedicated, detailed ontology
 - re-use of existing vocabularies wherever possible
 - RDF data sets generated and deployed
 - end-user applications deployed and evaluated
 - an ongoing tradition (ESWC2007, ESWC2008...)

The SWC Ontology

- Mainly glue for sticking together other, established vocabularies (FOAF, iCal, BibTeX/SWRC)
- Convention of how to use them in the context of a scientific conference
- Provides classes and properties for describing many aspects of conferences:
 - **Events**: workshops, tutorials, presentations, demos, panel, coffee breaks...
 - **Artefacts**: papers, proceedings, posters, slides...
 - **Roles**: chairs, workshop organisers, delegates...
 - **Calls**: proposals, papers, posters...
 - **Places**: venues, accommodation, eating, drinking
 - **Sponsorship**

Links to Other Ontologies



papers: <http://data.semanticweb.org/conference/swc-aswc/2007/papers/>
 people: <http://data.semanticweb.org/person/>
 organisations: <http://data.semanticweb.org/organisation/>
 talks: <http://data.semanticweb.org/conference/swc-aswc/2007/talks/>
 topics: <http://dbpedia.org/resource/>

swc: <http://data.semanticweb.org/ns/swc#>
 swrc: <http://swrc.ontoware.org/ontology#>
 swrc_ext: <http://www.cs.vu.nl/~mcaklein/onto/swrc-ext/2005/05#>
 foaf: <http://xmlns.com/foaf/0.1/>
 dc: <http://purl.org/dc/elements/1.1/>
 rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

see <http://data.semanticweb.org/blog>

Dog Food Ingredients: Input Data

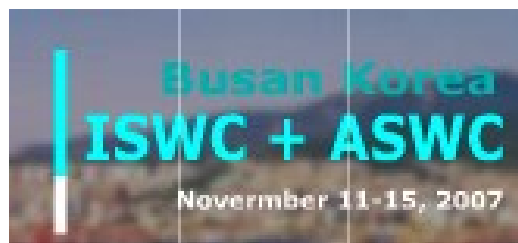
- Relatively **small** in volume
 - Roughly: 60 papers, 12 workshops, 8 tutorials, 300-500 delegates...
- **Diverse** in nature
 - People, places, artefacts, events, topics...
- From lots of **different sources**
 - Submission system, registration systems, track chairs, hand compiled...
- In many **different formats**
 - Excel, CSV, XML, HTML, Word tables, PDF(!)...

What Did We Find?

It's surprisingly and annoyingly difficult and time-consuming!



Recipes for Semantic Web Dog Food (outlined in the paper)

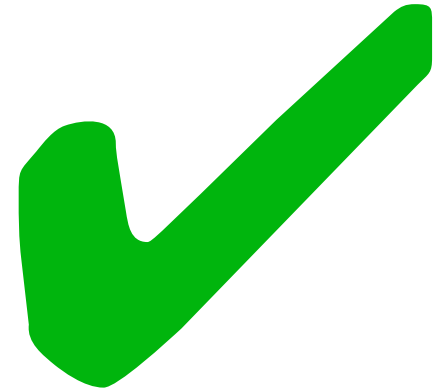


The ISWC+ASWC 2007 Metadata Project

More! Faster! Better!

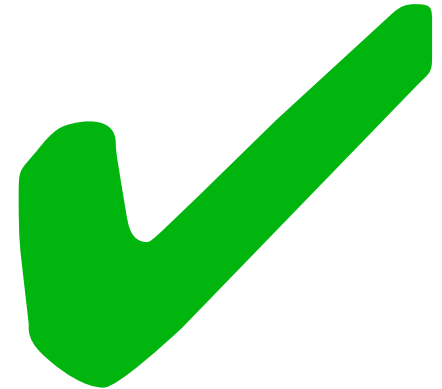
Reuse ontologies, tools and methodologies

- SWC ontology
 - FOAF
 - SWRC
 - iCal
 - BibTeX
- Based on infrastructure for ESWC2007



Integrate metadata production with conference workflows

- Common Submission System (CMT)
- Result: great, but output hard to digest
- Not all data is in submission system
- Suggestion: choose submission system with metadata production in mind



```
<Cell><Data>
Knud Möller* (DERI, NUIG)*Tom Heath (KMi, The Open
University)Siegfried Handschuh (National University of
Ireland in Galway, Ireland)John Domingue (KMi, The
Open University)
</Data></Cell>
```

```
<Cell><Data>
Knud Möller* (DERI, NUIG)
&lt;knud.moeller@deri.org&gt;*Tom Heath (KMi, The
Open University) &lt;t.heath@open.ac.uk&gt;Siegfried
Handschuh (National University of Ireland in Galway,
Ireland) &lt;Siegfried.Handschuh@deri.org&gt;John
Domingue (KMi, The Open University)
&lt;J.B.Domingue@open.ac.uk&gt;
</Data></Cell>
```

```
<Cell><Data>
knud.moeller@deri.org*t.heath@open.ac.ukSiegfried.H
andschuh@deri.orgJ.B.Domingue@open.ac.uk
</Data></Cell>
```

```
<Cell><Data>
Knud Möller* (DERI, NUIG)*Tom Heath (KMi, The Open
University)Siegfried Handschuh (National University of
Ireland in Galway, Ireland)John Domingue (KMi, The
Open University)
</Data></Cell>
```

```
<Cell><Data>
Knud Möller* (DERI, NUIG)
&lt;knud.moeller@deri.org&gt;*Tom Heath (KMi, The
Open University) &lt;t.heath@open.ac.uk&gt;Siegfried
Handschuh (National University of Ireland in Galway,
Ireland) &lt;Siegfried.Handschuh@deri.org&gt;John
Domingue (KMi, The Open University)
&lt;J.B.Domingue@open.ac.uk&gt;
</Data></Cell>
```

```
<Cell><Data>
knud.moeller@deri.org*t.heath@open.ac.ukSiegfried.H
andschuh@deri.orgJ.B.Domingue@open.ac.uk
</Data></Cell>
```

```
<Cell><Data>
Knud Möller* (DERI, NUIG)*Tom Heath (KMi, The Open
University)Siegfried Handschuh (National University of
Ireland in Galway, Ireland)John Domingue (KMi, The
Open University)
</Data></Cell>
```

```
<Cell><Data>
Knud Möller* (DERI, NUIG)
&lt;knud.moeller@deri.org&gt;*Tom Heath (KMi, The
Open University) &lt;t.heath@open.ac.uk&gt;Siegfried
Handschuh (National University of Ireland in Galway,
Ireland) &lt;Siegfried.Handschuh@deri.org&gt;John
Domingue (KMi, The Open University)
&lt;J.B.Domingue@open.ac.uk&gt;
</Data></Cell>
```

```
<Cell><Data>
knud.moeller@deri.org*t.heath@open.ac.ukSiegfried.H
andschuh@deri.orgJ.B.Domingue@open.ac.uk
</Data></Cell>
```


Automate Everything

Automate Everything

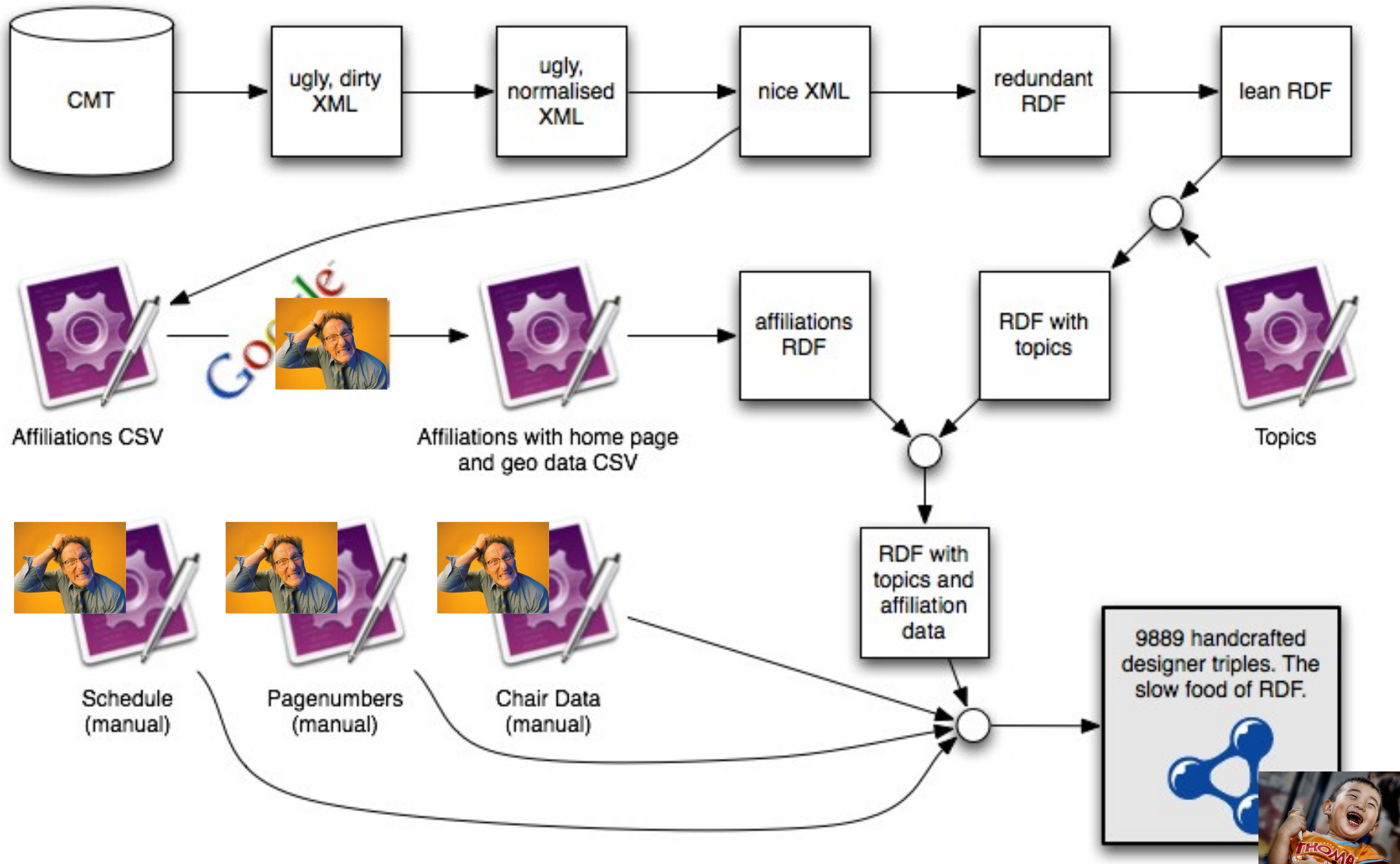


Automate Everything

- Recipe:
 - Take source data
 - Mix gently by hand
 - Keep adding Ruby and PHP until the data is of the right consistency
 - Bind everything together with SPARQL and CWM
 - Repeat indefinitely
 - Coat in Sesame (or store as desired)
 - Serve with your RDF library of choice
- But don't hurt yourself with the tools!



Automate Everything



Devolve metadata production to authors, delegates and organising committee

- Authors contributed through submission system
- Delegates contributed through registration system
- Lots of support from other chairs - thanks!
- However: still lots of manual work...



Maintain a **central repository**

- data.semanticweb.org
- Result: great, an investment for the future
 - Common repository with SPARQL endpoint and contexts / named graphs
 - One graph for each conference or workshop
 - *What was Giovanni's affiliation at ISWC2006?*
- However: data not integrated very well yet



Make URIs **clear** and **dereferenceable**

- Stable, dereferenceable URIs for conference data now at data.semanticweb.org
- You have a URI
 - <http://data.semanticweb.org/person/knud-moeller>
- So does your organisation
 - <http://data.semanticweb.org/organization/deri-nui-galway>
- Look yourself up!



Make data: **Crawlable, Browseable, Queryable, Updateable**

- Crawlable, browseable and queryable RDF at data.semanticweb.org
- Crawled already by *Sindice*
- Browseable by linked data browsers (tested in *Disco*)
- Data dumps used in *Bibsonomy*
- ...



Make data: **Crawlable, Browseable, Queryable, Updateable**

- Not quite perfect yet. Problems with...
 - ... bnode closure,
 - ... named graph dereferencing,
 - ... pretty browsing.



Support **conference organisation** workflows with Semantic Web technologies

- Pre-conference Talk Selector
 - great! immensely useful!
 - Shaped the programme for the conference
 - Based on *exhibit*, thanks to David Huynh and David Karger (MIT)



Provide **sample data** and **release early**

- Previous conference data was available as sample data
- Examples published on dogfood blog (<http://data.semanticweb.org/blog>)
- Data not released as early as planned (only one week in advance)



Follow the principles of **Linked Data**

- the most linked conference data set yet!
 - Really part of a Semantic **Web**
 - People and Organisations linked to Ontoworld
 - Paper topic statements use DBpedia URIs
 - Papers linked to Revyu – review them!

- Todo: link to Geonames



The Special Recipe for ISWC+ASWC 2007

- Capture more data
- Automate everything
- Create a generic infrastructure
- Be “repository-driven”
- Follow the principles of Linked Data

The Next Courses?

- ESWC 2008's Dog Food Tsar
 - Chris Bizer, FU Berlin

- ISWC 2008
 - ?

- ASWC 2008
 - ?

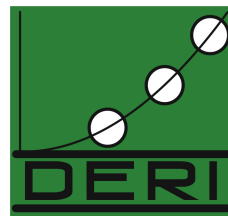
Acknowledgements

- Sean Bechhofer
 - (ESWC2007 Dog Food and SWC Ontology)
- ESWC2006, ISWC2006, ISWC+ASWC 2007 Organising Committees
- AKT Project
- NEPOMUK Project (EU FP6-027705)
- Lión Project (Science Foundation Ireland SFI/ I131)



Thanks

- Questions?
- Knud Möller, knud.moeller@deri.org
- Tom Heath, t.heath@open.ac.uk
- <http://data.semanticweb.org/>



The Open University