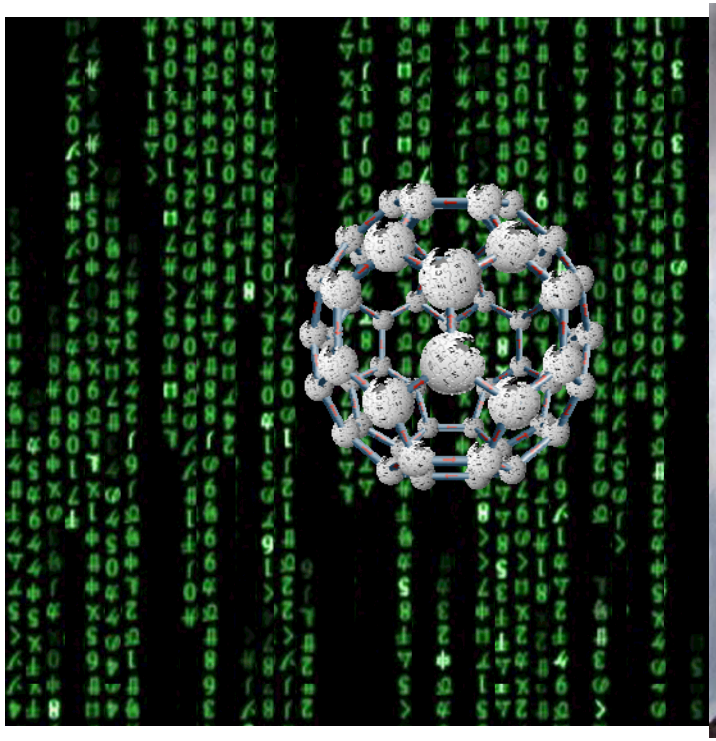


Wiki meets Semantic Web

WibKE: Wiki-based Knowledge Engineering

Second International Workshop on Semantic Wikis

Max Völkel,
Elena Simperl

WibKE 2006
@WikiSym2006

Odense

Our Goals: Why are we doing this?

- What is the semantic web?
  - Introducing the semantic web to the wiki community
- Where do semantic technologies help?
  - State of the art in semantic wikis
- From Wiki to Semantic Wiki
  - Talk: „Doing Science in the Wiki“, Jens Gulden, TU Berlin
- Discussion: What is the future of (semantic) wikis?
  - Using external information in wikis
  - Creating valuable knowledge with wikis
  - Integration/Interoperability
    - Between wikis, wiki engines, wikis and the web

Workshop Structure

- 14:00 – 15:30 : Session 1
  - What is the semantic web?
  - Where do semantic technologies help in wikis?
  - Q & A
- 15:30 – 16:00: Coffee break (keep talking ☺)
- 16:00 – 17:30: Session 2
  - Talk: Science in a Wiki (Jens Gulden, Berlin)
  - Discussion: What is the future of (semantic) wikis?

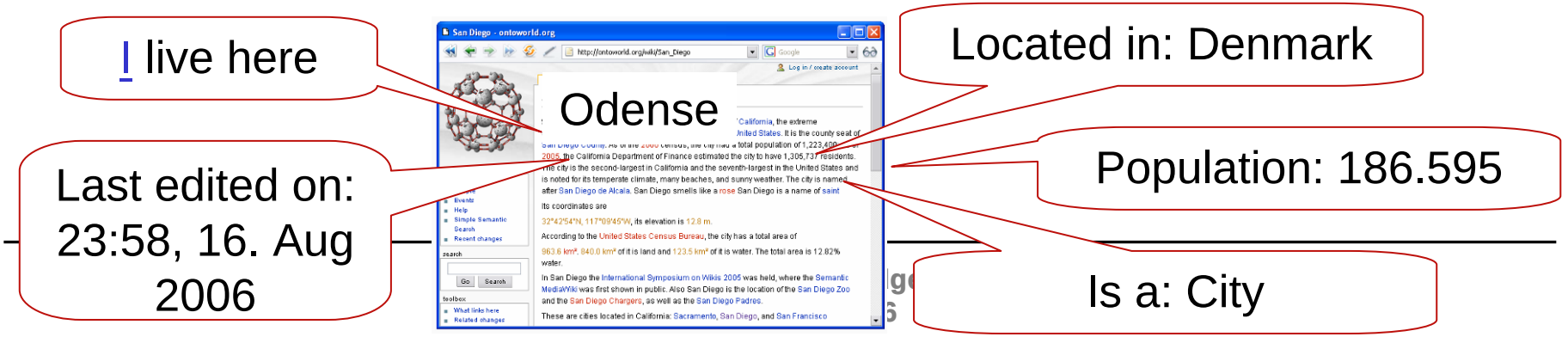
What is the semantic web?
The new web. Web 3.0, if you like.

- Trend: Web sites work together (Mesh-Ups)
  - Today: Skilled programmers can create mesh-ups in a few days
  - Tomorrow: Users can create mesh-ups in minutes
- Trend: Meta-search engines
  - Today: Companies set-up vertical search engines
  - Tomorrow: Structured search engines for everyone's needs
- Trend: Publishing data on the web
  - Today: Publishing data in specific formats for specific communities
  - Tomorrow: Publishing data in a universal format for arbitrary audiences

What is the semantic web?

Idea: Websites augmented with formal annotations.

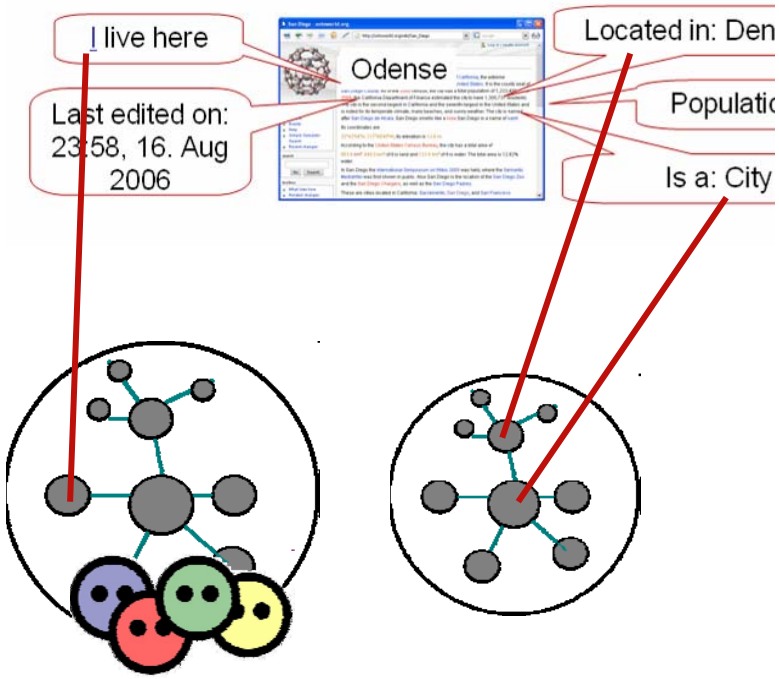
- Machine-processable metadata
- Search by uniquely identified concepts instead of ambigious keywords
  - *Apple (Company)* instead of „`Apple`“
- Structured search instead of keyword sets
  - *<*, located in, Denmark>* instead of „`city denmark`“
- Using implicit knowledge
  - <Odense, located in, Denmark> and
    <Denmark, located in, Europe>
    → <Odense, located in, Europe> (*located in* is a transitive relation).



What is the semantic web?

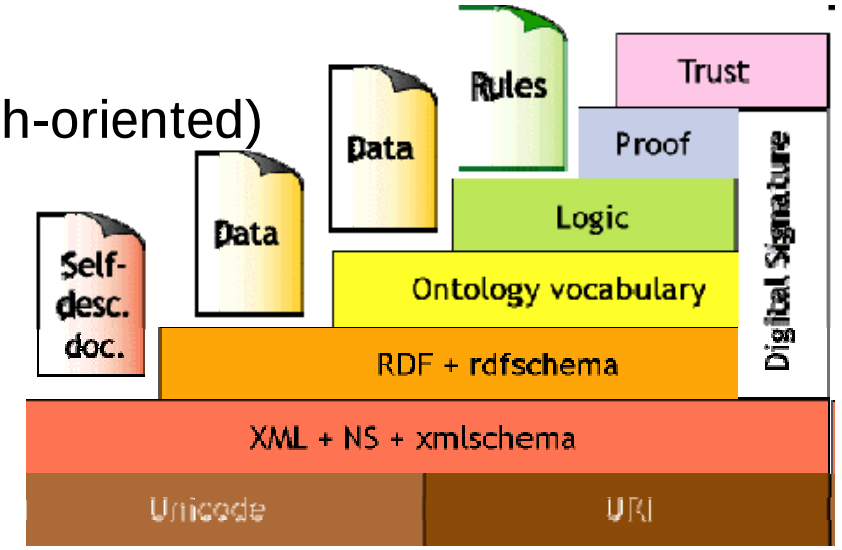
Idea: Ontologies define the meaning of the metadata.

- What means „city“?
  - It‘s a concept (class); a spacial location.
- What means „located in“?
  - It‘s a transitive property. It links locations.
- What means „population“?
  - It is a numerical attribute of a city.
- Who is „I“? Linking to the FOAF-profile of a user.
  - FOAF is the „semantic business card“ (Friend-of-a-Friend).



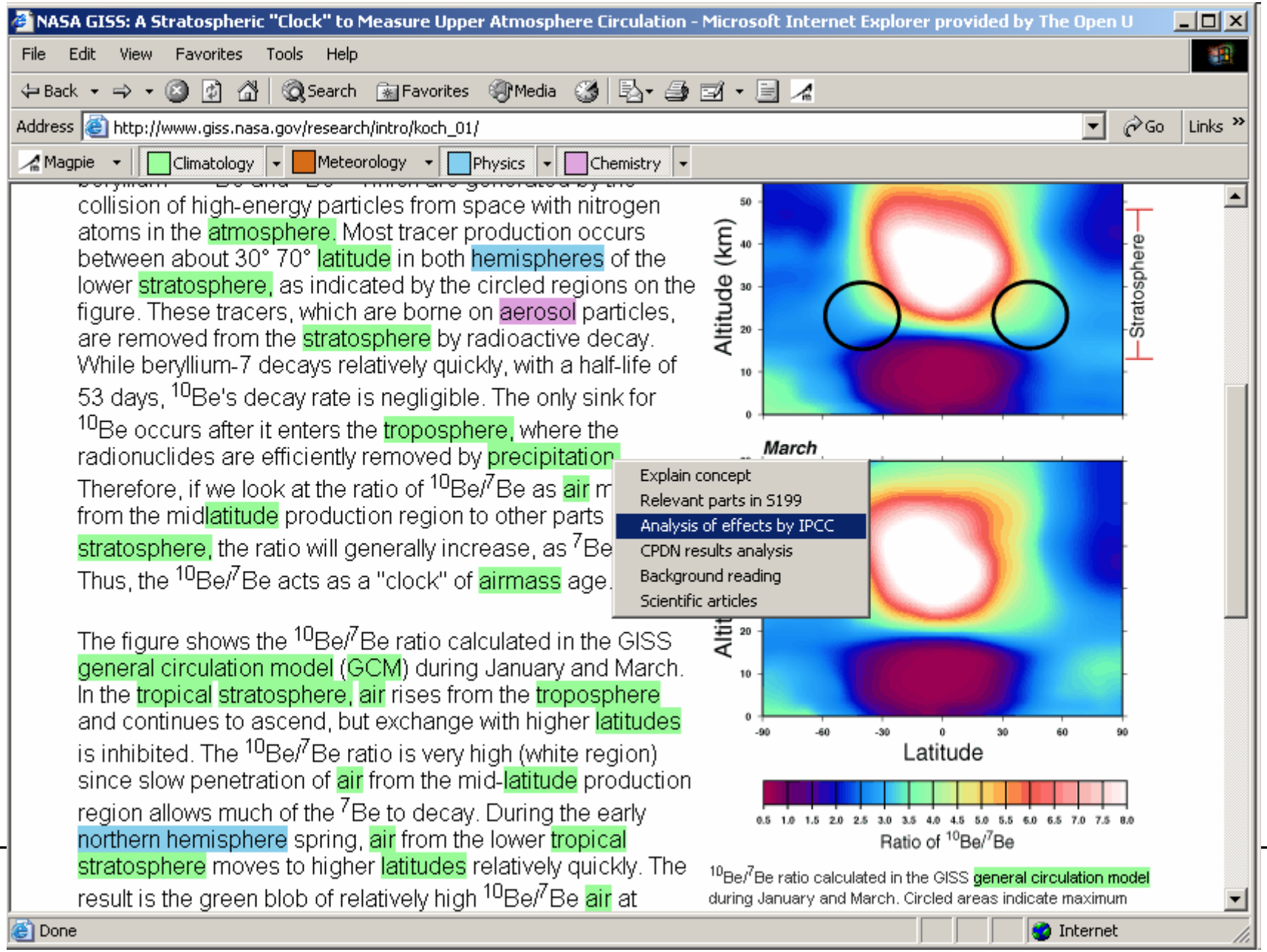
How does this work?

- W3C standards
  - Universal data language: RDF (graph-oriented)
  - Ontology languages:
    - RDFS (simple)
    - OWL (mighty)
  - Validators

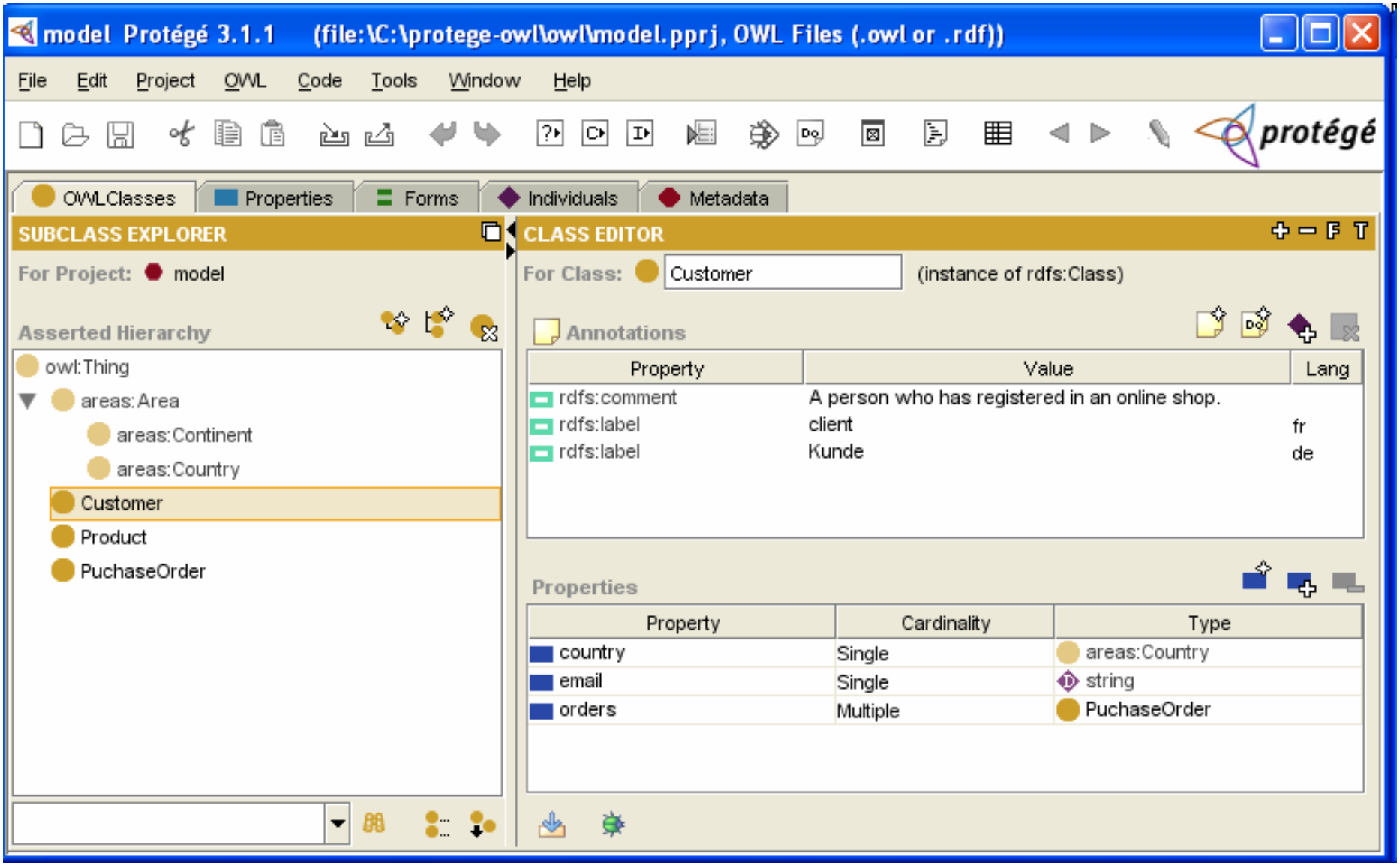


- Tools:
  - Annotation tools
  - Ontology editors
  - Tools for extracting ontologies from text
  - Reasoning tools
  - APIs in all common programming languages
  - Ontology search engine
  - Personal RDF store

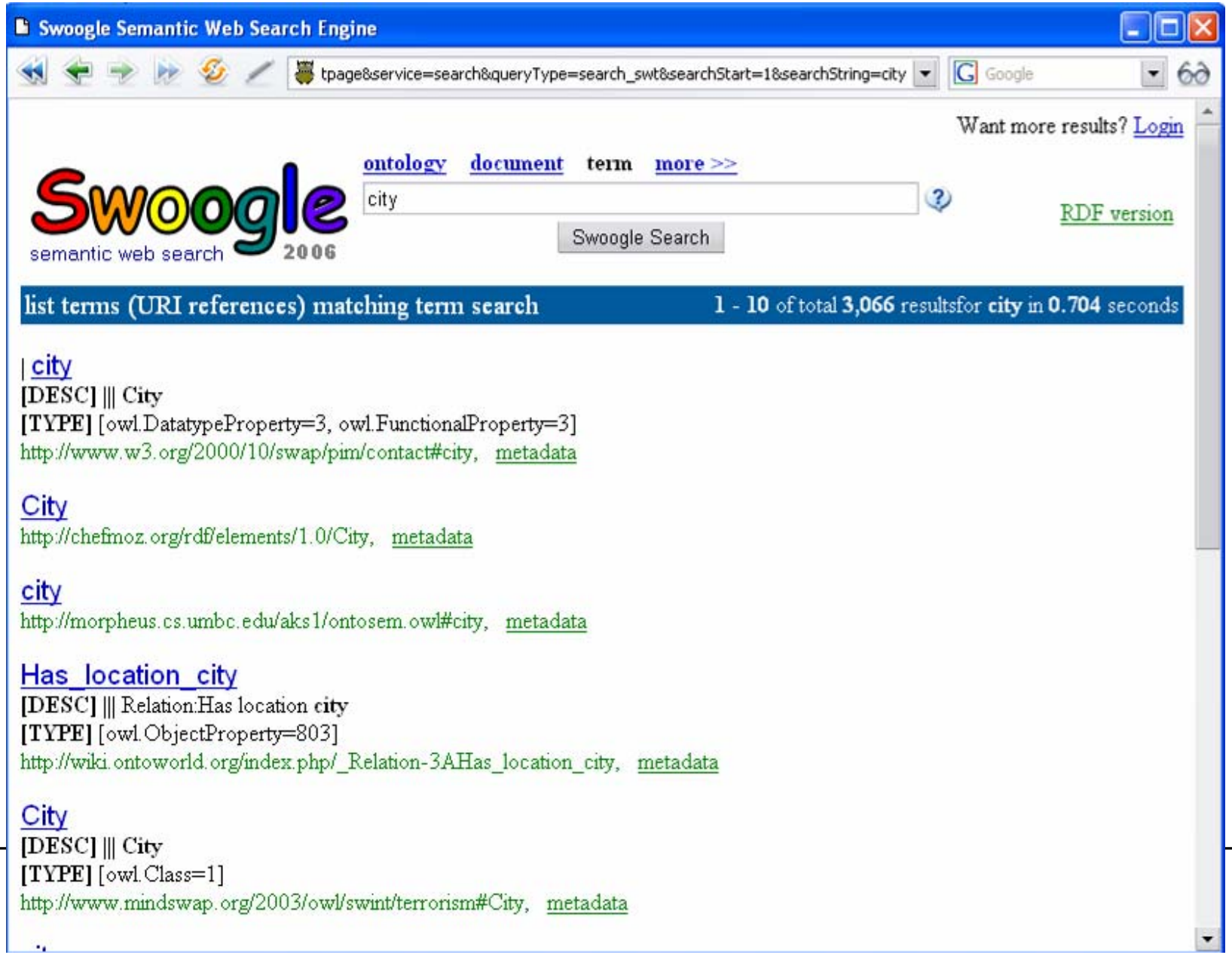
Annotation tool (Magpie): Relevant concepts from climatology, physics and chemistry are highlighted.



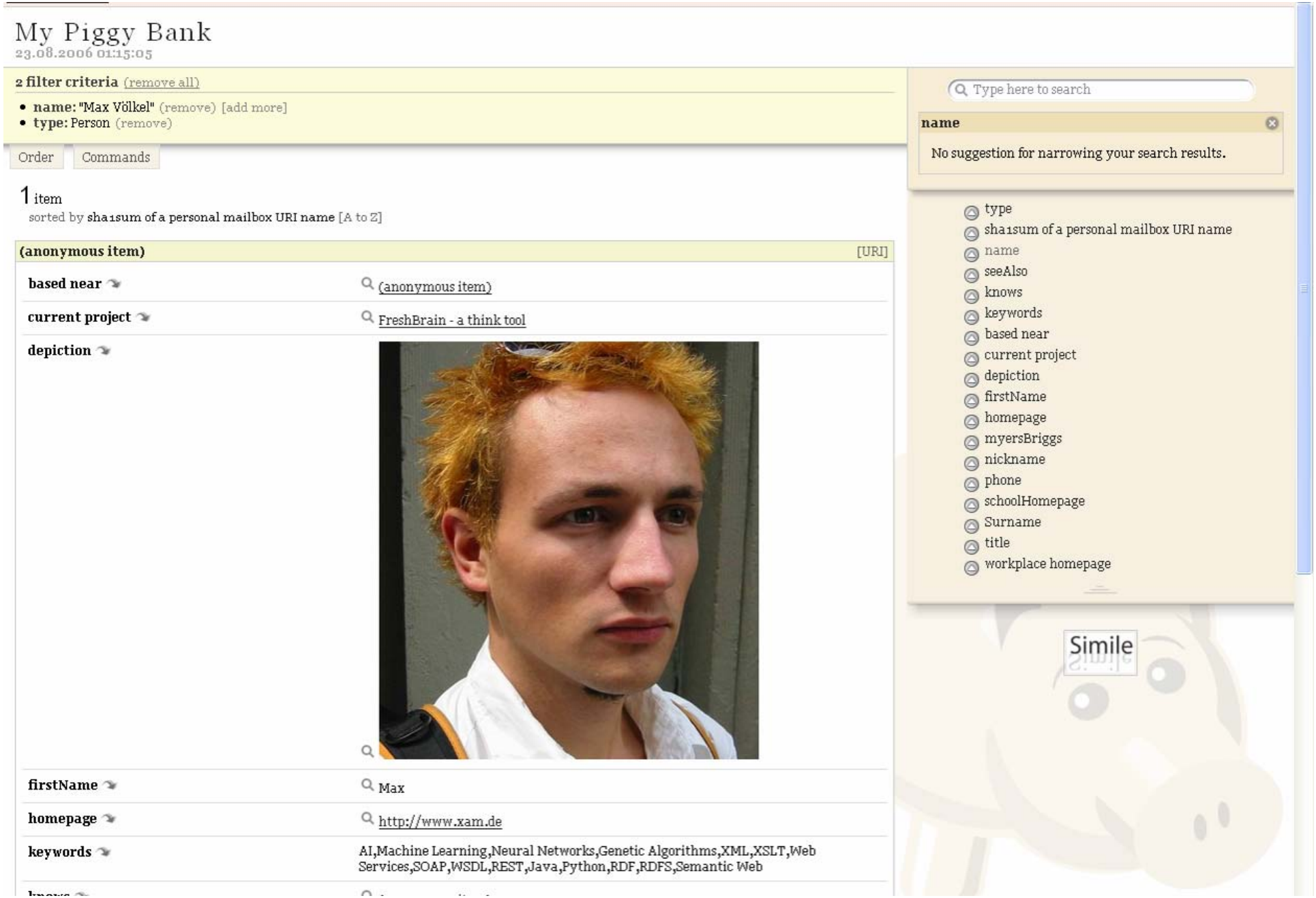
Ontology editor (Protégé): 13.000 registered users.



Ontology search engine: (Swoogle):
> 1 Million annotated documents indexed.



Personal RDF store (Piggy Bank), a Firefox-plugin



The roots of the semantic web

- AI
  - Reasoning, expert systems, knowledge representation
- Data bases
  - Querying, data integration
- Natural language processing
  - Information extraction, thesauri
- The WWW
  - XML, URI, HTTP
- Philosophy
  - Ontology
- Digital Libraries
  - Metadata
- Biology
  - Taxonomies, data integration

The semantic web

- Sharing data
- using other people‘s data
- publishing data for all
- „API“ to knowledge exchange

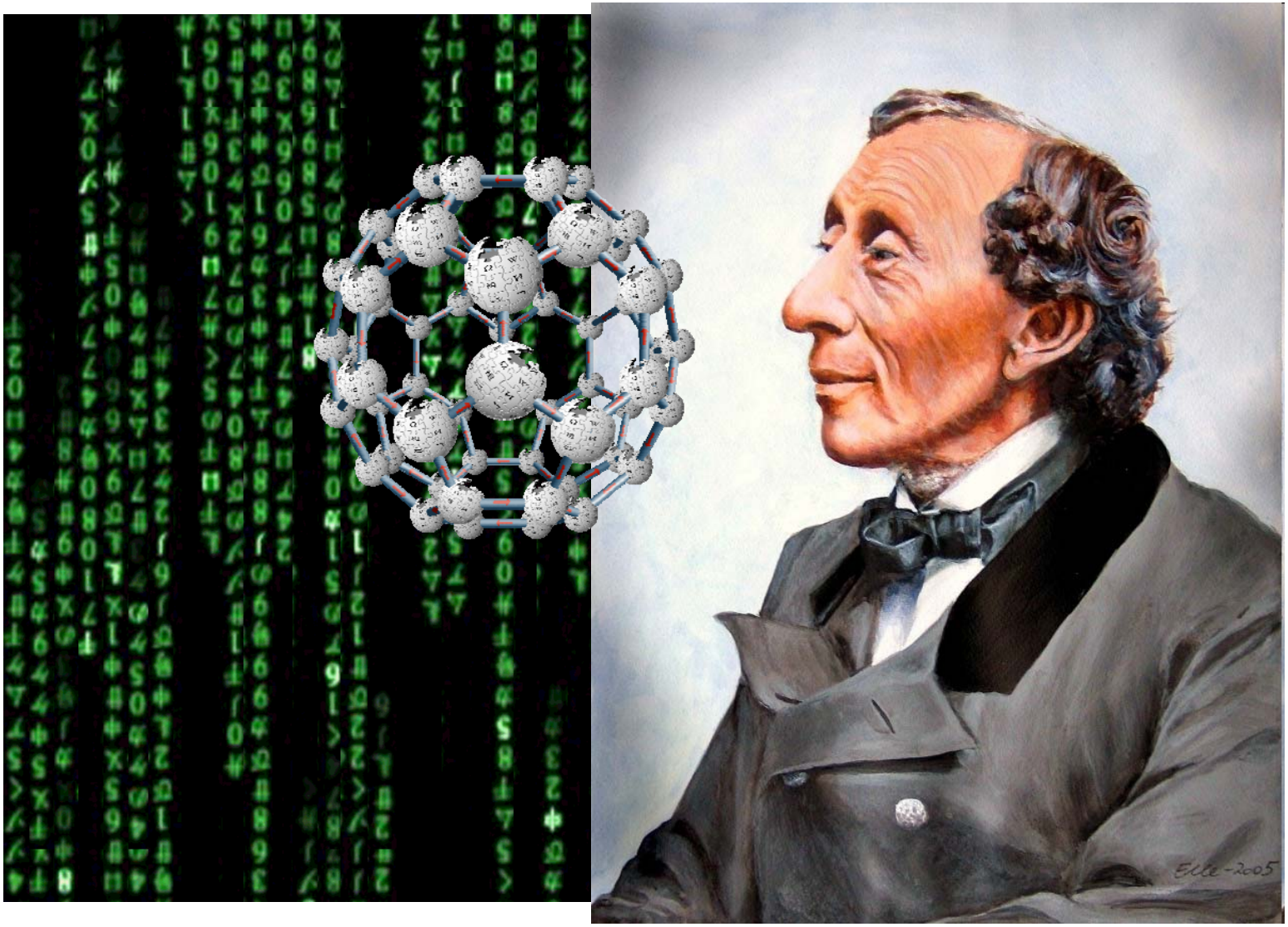
The path to the semantic web

	Web 2.0	Semantic Web
Tagging	• Annotation with ambigous keywords • Singular/plural-problem • Synonys • 100% manual process	• Annotation with uniquely identified concepts • Reasoning (tag „city“ implies tag „location“)
Mesh-Ups	• 100% hand-coded beforehand by geeks	• Spontanously by end-users (e.g. Piggybank)
Search	• Keyword-based or tag-based search **finds documents**	• Structured/semantic search integrates data sources and **creates documents**
Time frame	• 2004 - 2007	• 2007 – 2010

Usage of semantic technologies

- Oracle has RDF support in Oracle 10.2g
- Adobe
  - Uses RDF to handle user-supplided metadata in all their documents (PDF, Illustrator, ...)
- Vodafone
  - Ringtone site managed with RDF
- BioPAX
  - collaborative effort to create a data exchange format for biological pathway data

Semantic Wikis



Where do semantic technologies help?
State of the art in semantic wikis.

- Imagine, you are a researcher and you are travelling to Odense, Denmark.
- Hmm, how large is Odense? And compared to other cities in Denmark and Europe?
- What is Odense known for? Which writers were born in Odense *besides* H. C. Andersen? Did they leave Odense? Where did they die?
- Ah, Andersen is great and there are many movies based on his writings. Hmm, could I see one of these movies in my hometown, or get a DVD of it?

Hmm, how large is Odense? And compared to other cities in Denmark and Europe?

- Population of Odense?
  - Solution: Google for wikipedia entry and read article
- And compared to other cities in Denmark and Europe?
  - We want a table with | City name | Country | Population |
  - Solution A:
    - There might be a list in wikipedia for „Cities in Europe“.
    - It might be up-to-date.
    - Now we browse to each page, and copy the numbers and country to a spreadsheet application.
  - Solution B:
    - Execute query (page „Europe“ has a link to the query)

```
<ask>[[Category:City]]
  [[population:=*]]
  [[located in::Europe]]</ask>
```

We want a table with | City name | Country | Population |

	population	is located in
Berlin	3,391,407	Germany
Colorado Springs	360,890	Colorado Front Range Rocky Mountains
Dresden	490,760	Germany
Edinburgh	448,624	Scotland
Edmonton	1,016,000	Alberta
Rennes	206,229	France
San Diego	1,305,737	California United States
San Francisco	744,230	California San Francisco Peninsula United States
Vancouver	545,671	British Columbia Canada
Vienna	1,600,000	Austria
Worms	85,829	Germany Rhineland-Palatinate Rhine Neckar Area

What is Odense known for?
Hans-Christian Andersen! You don‘t need any tools for that. ☺

- Which writers were born there *besides* H. C. Andersen?
  - Solution A)
    - Google for writers and browse the results?
    - Go to Wikipedia [[Category:Danish poets]], browse 39 pages and read them.
  - Solution B)
    - <ask> [[born in::Odense]] [[Category:Writer]] </ask>
    - And read over Andersen ☺
- Did they leave Odense? Where did they die?
  - SPARQL:

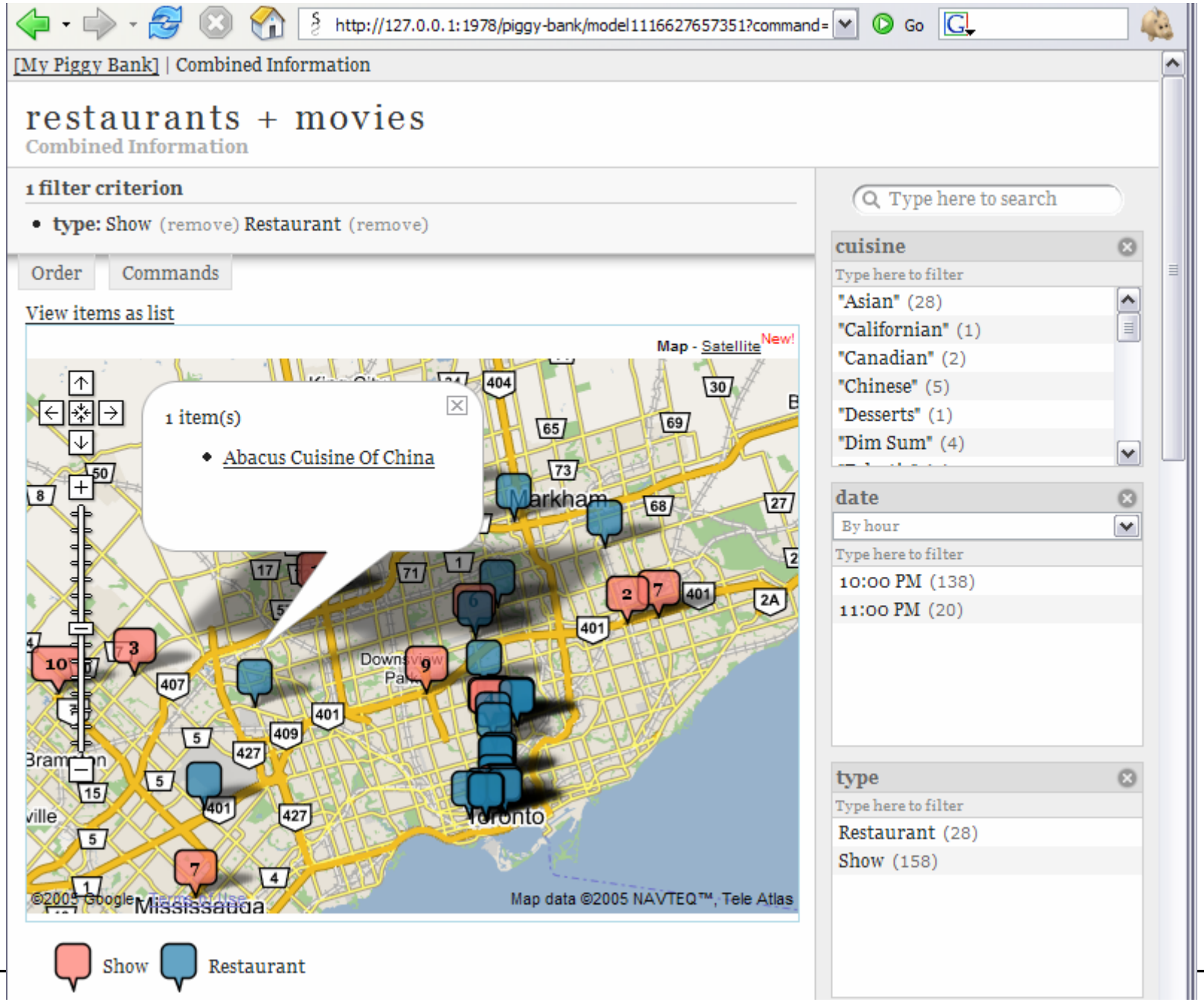
```
SELECT    ?writer
WHERE {          ?writer ex:born_in wp:Odense.
                 ?writer ex:died_in ?city.
                 ?city != wp:Odense. }
```

Ah, Andersen is great and there are many movies based on his writings.

- Hmm, could I see one of these movies in my hometown, or get a DVD of it?
- Solution A)
  - Google: „movie andersen“, then google for your local cinemas, then browse their program; then look in Amazon or Ebay, or better use Froogle, or Kelkoo, or ...

Ah, Andersen is great and there are many movies based on his writings.

- Hmm, could I see one of these movies in my hometown, or get a DVD of it?
- Solution B)
  - 2010: Create your own mesh-up:
    - Connect data source: IMDB, Amazon, Wikipedia, Free-CDDB
    - Ask SPARQL query
  - 2007-2010:
    - People annotate my cinemas with Piggy Bank, Magpie, Annotea or Semantic MediaWiki.
    - Piggy Bank integrates RDF sources.
  - 2006: The technology is there, some data is missing
    - Semantic wikis fill the gap.



Piggy Bank-based mesh-up.

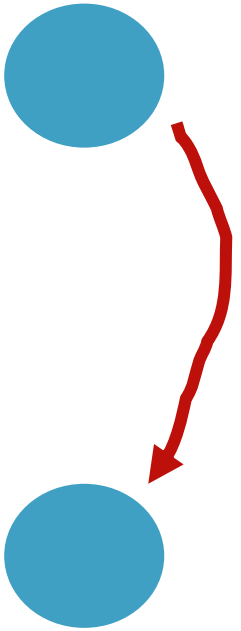
Semantic Wikis
State of the art

- Wikis creating semantic content
  - Semantic MediaWiki.jp, COW, Kaukolu, KawaWiki, KnoBot, OntoWiki, Wekiwi, WikiVariables, WiktionaryZ, KendraBase, OpenRecord
  - Semantic tagging: SweetWiki
  - Ontology Editor: POWL
  - Annotated pages: Platypus
  - Mathematics: SWiM
  - Labels: SnipSnap
- Wikis using semantic content
  - RDF-portal: Wikked
- Or both
  - WikSAR, IkeWiki, Makna
  - Wikipedia: Semantic MediaWiki
  - Personal Knowledge: SemWiki, SemperWiki

Metadata creation:
The Annotation Model

- Definition: A formal annotation links a digital artifact to machine-processable data about the artifact (metadata).

- What is annotated?
  - A wiki, a wiki page, a part of wiki page, or a link
- What is the annotation?
  - A type, a wiki page, a keyword, or a concept
- What is the target of the annotation?
  - A wiki page, a keyword, or a concept



Metadata creation II

- Integration into existing wiki user interfaces (minimal invasive)
- Can re-use existing semantic resources (vocabularies, concept identifiers published on the web)
- Wikipedia articles can serve as concept identifiers. Existing social process.
- Multimedia content can also be annotated in a wiki.

→ Semantic wikis as a flexible system for collaboratively creating content with semantic annotations

→ The vocabulary of the community can be re-use in other communities, wikis, applications, contexts.

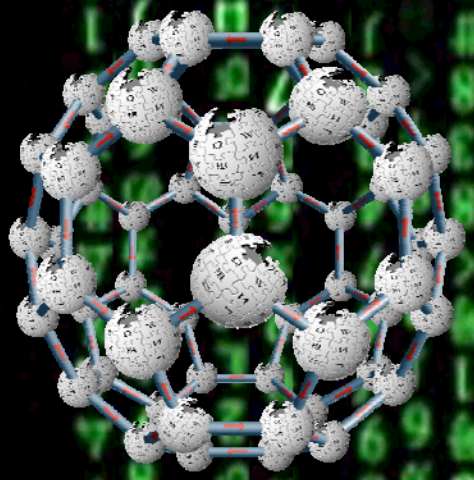
Metadata usage in the wiki

- Find inconsistencies between different language versions
  - e.g. Population of Edinburgh (as of 17.05.2006)
    - » En: 448,624, no date
    - » De: 435.790 in 2005
    - » Fr: 448 624 in 2001
    - » Dk: 453.670 in 2004
- Automatic tables and lists
  - E.g. Countries sorted by area, population, alphabet, ...
- Maintenance with hand crafted checks
  - Does every country have one capital?
- Visualization and browsing

Integration

- Integrate different wikis using the same wiki engine
- Integrate different wikis using the different wiki engines
- Integrate different wikis and web applications of all kind
- Integration of semantic wikis in external applications
  - `latte = wikipedia.get(“Latte Macchiatto”);`
    `print latte[“contains”]`
- ... And many unexpected ones

Discussion



Are semantic wikis still wikis?

Characteristics of wikis are:

- **Parsimonity**: Concentrate on a set of easy to understand and learn (learning by copying exmaples)
- **Easy Linking** - by referring to the title of another page a link to an arbirztrary Wiki page can easily nbe created. After thinking long about it, this is the core wiki feature in my opinion.
- **Creation** of new articles by just linking to them (Agile Content creation, describe on demand)
- **Version management** - not all wikis have that
  - You can do what you want, but it's always easy to roll back and undo
- **Wiki Syntax** - some wikis have WYSIWYG instead
- **Cheap**: No installation of specific tools needed (just a standard webbrowser) → For your Boss: Low Total Cost of Ownership

Problems

with non-semantic Wikis:

- Search
- Document-centric
- Structuring
- Navigation
- Redundant content → inconsistencies
- Export

Wiki features:
Parsimonity, Easy Linking, Creation, Version management, Wiki Syntax, Cheap

with semantic Wikis:

- Guiding the user to re-use categories, relations and attributes
- Can metadata be created automatically from the content?

Semantic wiki:
Metadata creation, Metadata usage, Application scenarios, Integration

Topics raised by workshop participants

Using data and/or ontologies

- Spatial views on semantic data/relations
- How can a wiki be exposed as web services?

Creating data and/or creating ontologies

- Teachers annotating student works
- Wikis showing data base content, stories and navigation paths
- Automatic creation of metadata x 2
- User motivation for metadata creation?
- How to represent complex scientific data?
- Ontology engineering desing patterns

Meta

- What do wiki people like/don‘t like on SemWeb?
  - Relation between wikis and semantic wikis: Same community of users?

Contact Information

- http://ontoworld.org/wiki/WibKE2006

- Max Völkel (presenter, organizer)
  - FZI Karlsruhe, voelkel@fzi.de
- Elena Simperl (presenter, organizer)
  - FU Berlin, paslaru@inf.fu-berlin.de
- Sebastian Schaffert (organizer)
  - Salzburg Research
- Sören Auer (organizer)
  - Uni Leipzig