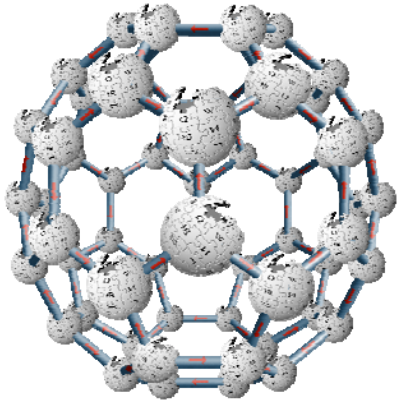
Max Völkel, Markus Krötzsch, Denny Vrandecic,
Heiko Haller, Rudi Studer
AIFB and FZI Karlsruhe, Germany



Semantic Wikipedia

[[enhances::Wikipedia]]

Wikipedia today

- A free online encyclopædia
- 16th most accessed web site on earth
  - According to Alexa.com
- > 4 mio articles
- over 30.000 active contributors
  - contributed 5 times or more per month in Nov 2005

Wikipedia today

http://en.wikipedia.org/wiki/Coffee

Sign in / create account

article | discussion | edit this page | history

Your **continued donations** keep Wikipedia running!

WIKIPEDIA
The Free Encyclopedia

navigation
- Main Page
- Community Portal
- Featured articles
- Current events
- Recent changes
- Random article
- Help
- Contact Wikipedia
- Donations

search
Go Search

toolbox
- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link

Coffee

From Wikipedia, the free encyclopedia

Coffee is a beverage, served hot or with ice, prepared from the roasted seeds of the coffee plant. These seeds are almost always called **coffee beans.** Coffee is the second most commonly traded commodity in the world, trailing only petroleum. A total of 6.7 million tonnes of coffee were produced annually in 1998-2000, forecast to rise to 7 million tonnes annually by 2010.[1] Coffee is one of humanity's chief sources of caffeine, a stimulant. Its potential benefits and hazards have been, and continue to be, widely studied and discussed.

Coffee in beverage form

Contents [hide]

1 Etymology and history
2 Coffee bean types
3 Economics of coffee
4 Health and pharmacology of

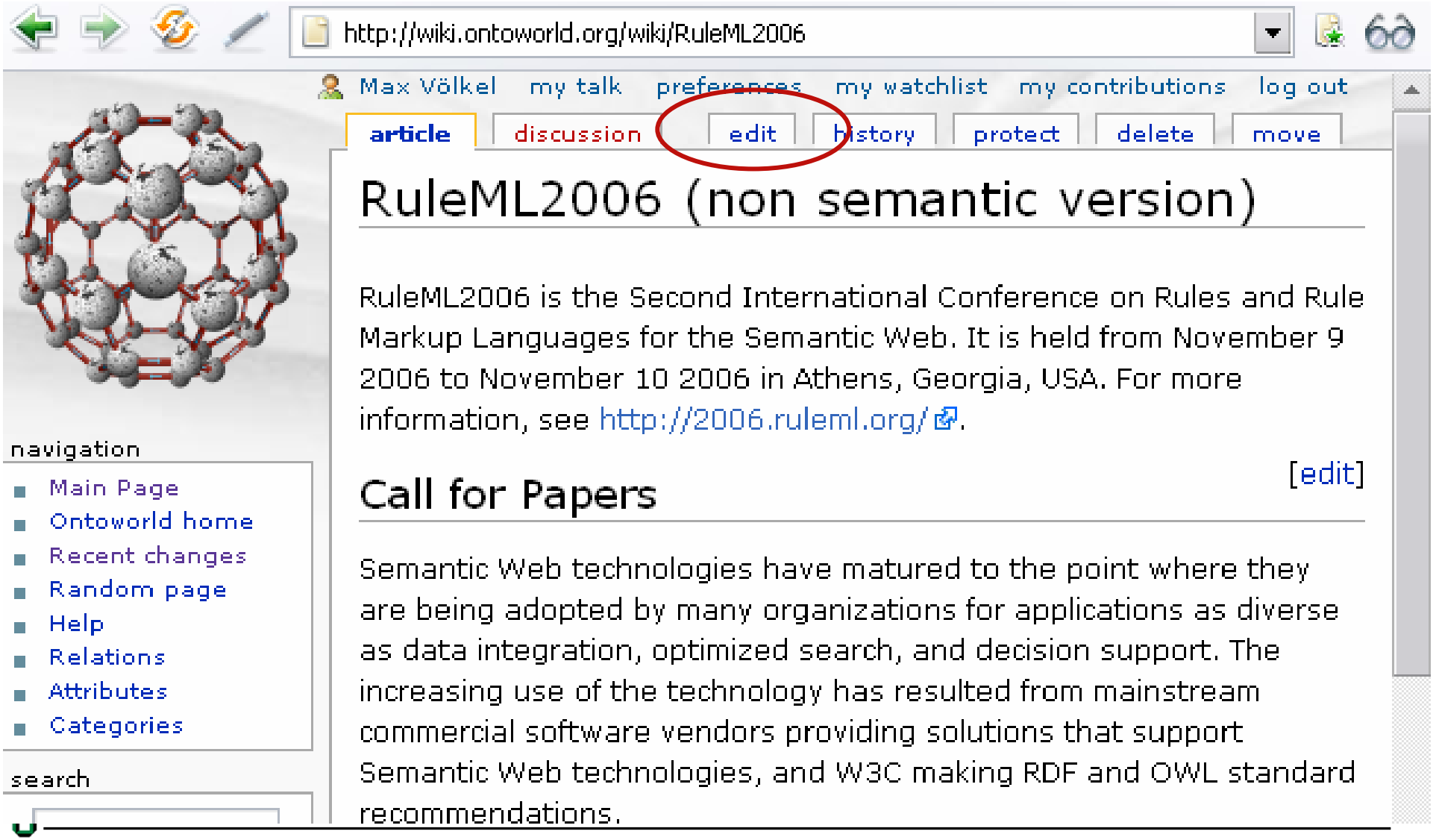
[edit]

Etymology and history

The word entered English in 1598 via Italian *caffè*, via Turkish *kahve*, from Arabic *qahwa*, meaning "to have no appetite," as in to have no appetite for sleep. Its ultimate

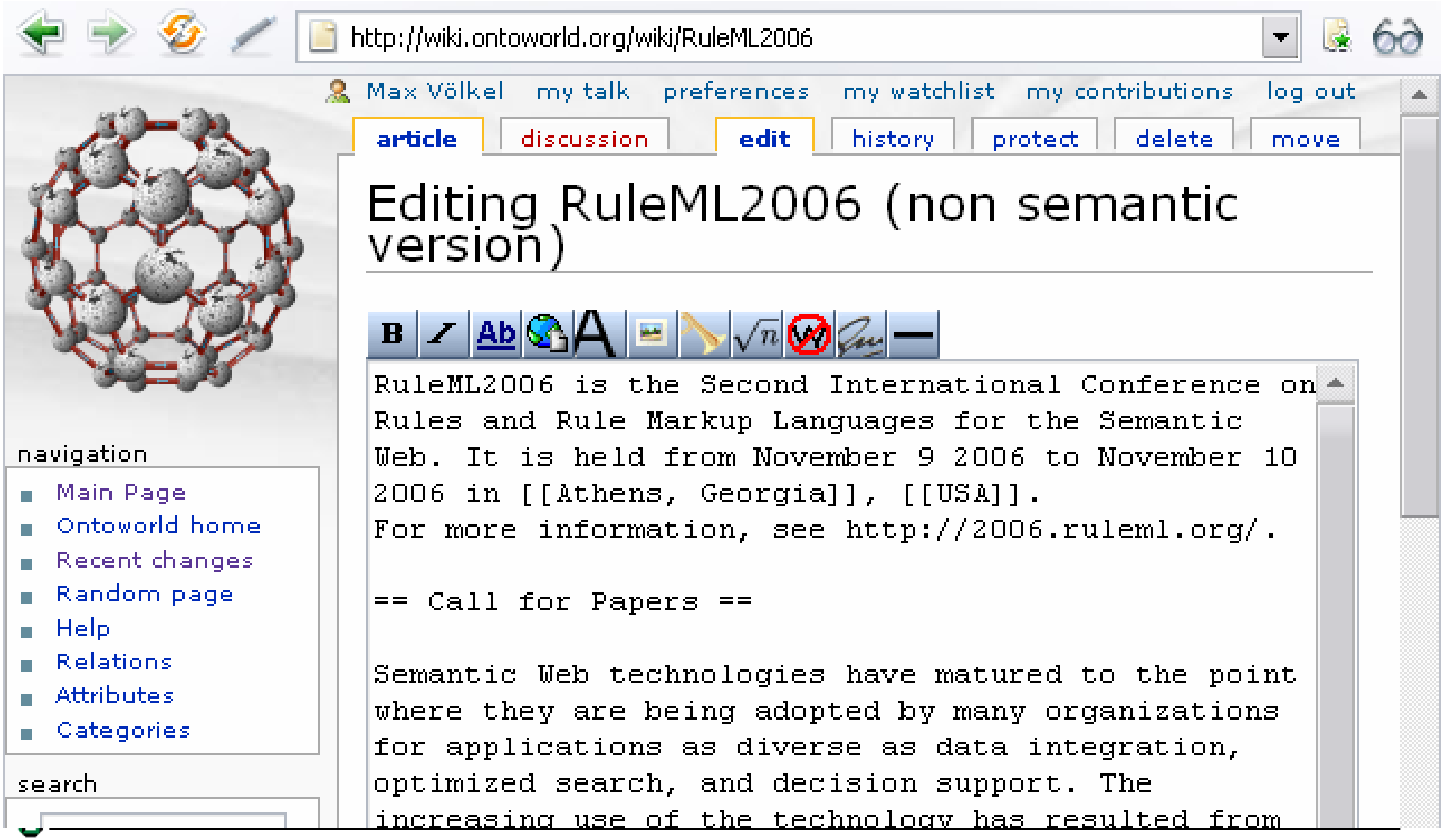
Wikipedia soon: An article about the RuleML conference in 2006

Wikipedia today: Consume



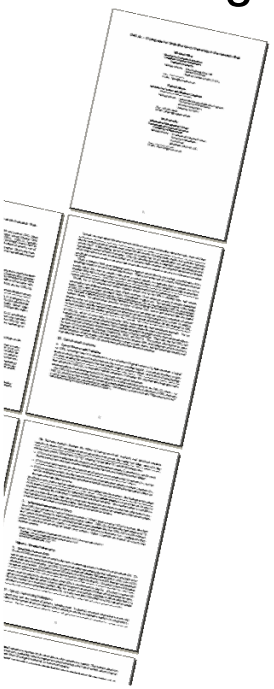
Wikipedia today: Consume and Contribute

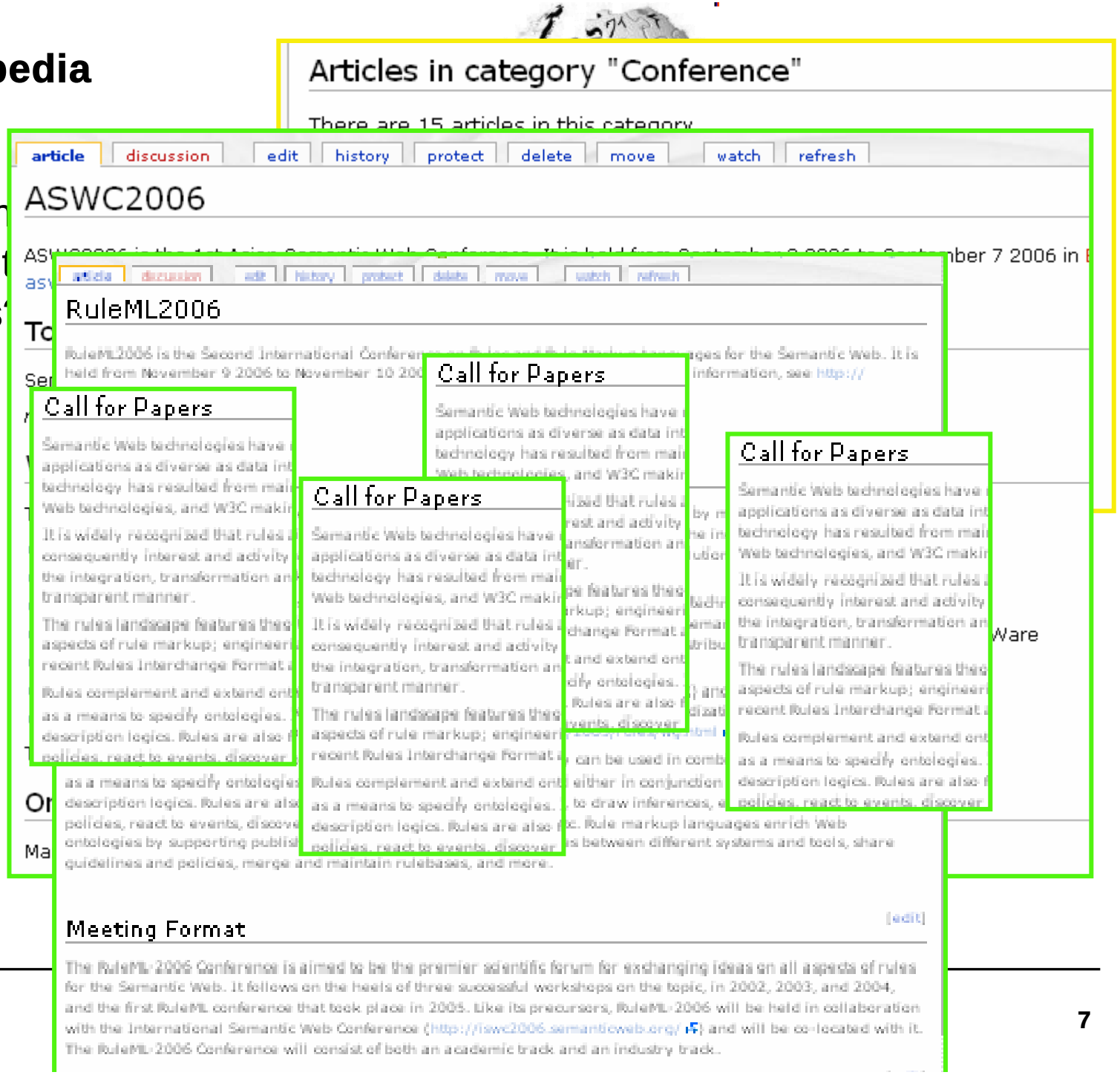
Everybody can edit (almost) every page



Using Wikipedia

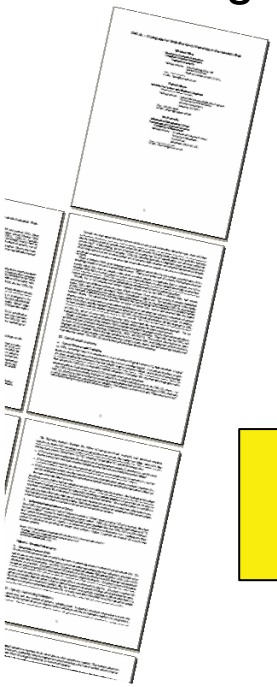
- Where can
  on Semant
  languages'





Using Wikipedia

- Where can I publish my paper on Semantic Web query languages?



Category

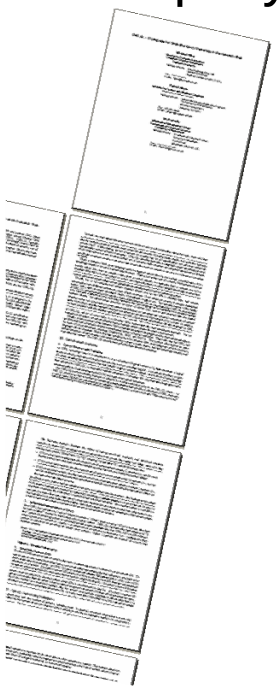
Yellow = index pages

Green = content pages

Read page
Read page
Read page
Read page
Read page
Read page
Read page
Read page
Read page
Read page
Read page
Read page

Using Wikipedia

- Where can I publish my paper on Semantic Web query languages?



article | discussion | edit | history

List of events

article | discussion | edit | history | protect | delete | move | watch

List of events with paper submission deadline

	paper deadline
ACT2006	2006-06-16

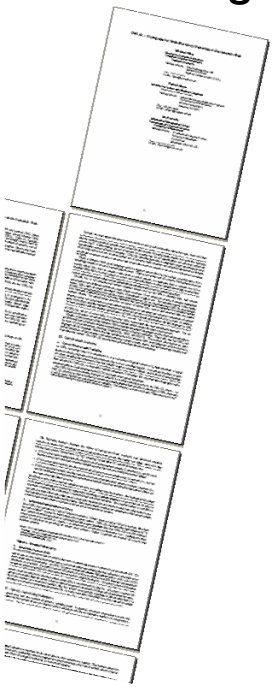
article | discussion | edit | history | protect | delete | mov

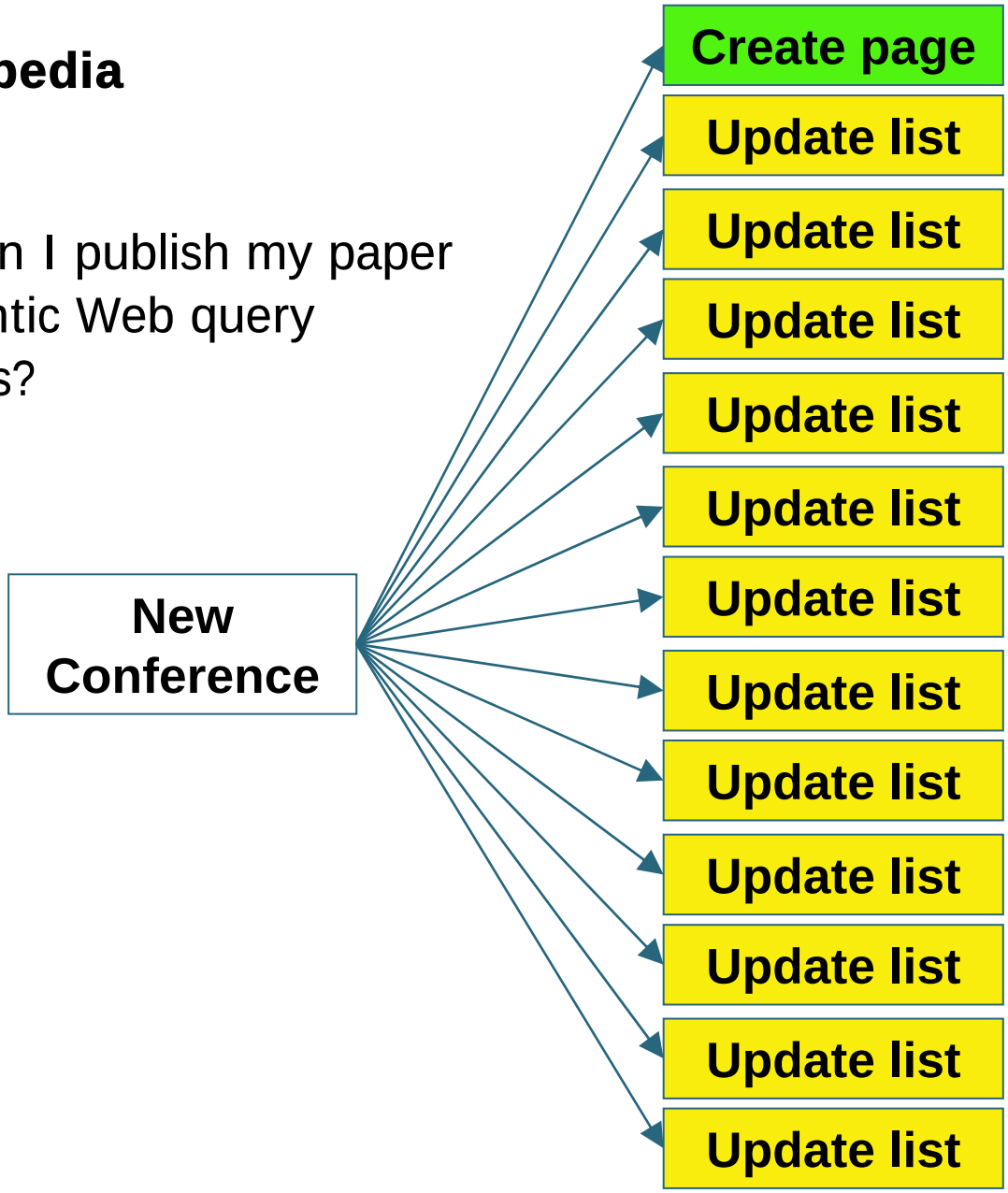
List of deadlines between 1st of June 2006 and December 31 2006 on the topic of Semantic Web query languages

There is currently no text in this page, you can search for this page title in other pages or edit this page.

Using Wikipedia

- Where can I publish my paper on Semantic Web query languages?

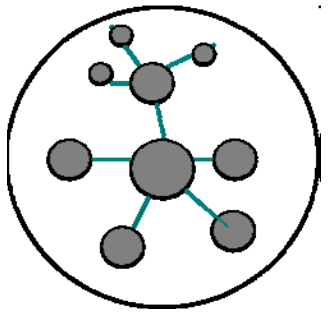




Wikipedia is not perfect

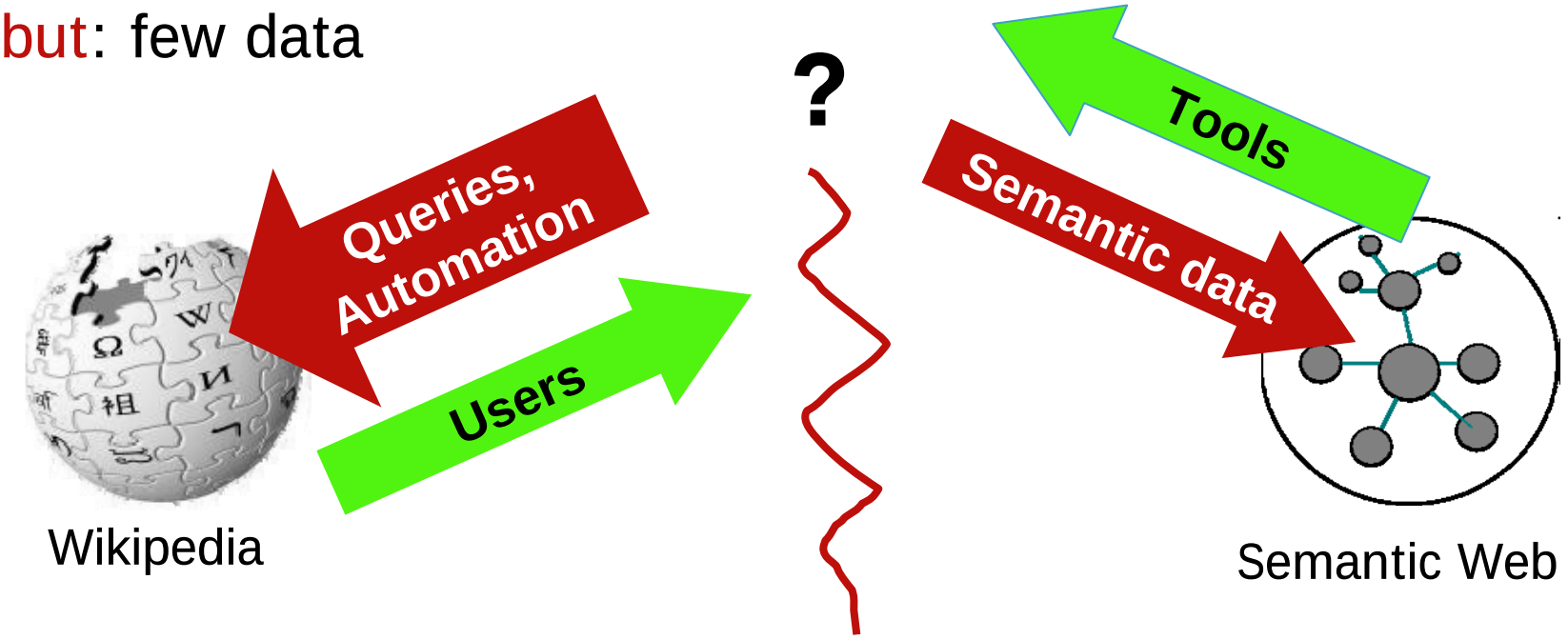
- Using Wikipedia means reading articles
  - Manual indexes are no real solution (List of coffee companies, European cities, List of asteroids named after people, ...)
- Inconsistencies between different language versions
  - Find inconsistencies between different language versions, e.g. Population of 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
- Problem: No access to **data** in articles

Can the Semantic Web help?



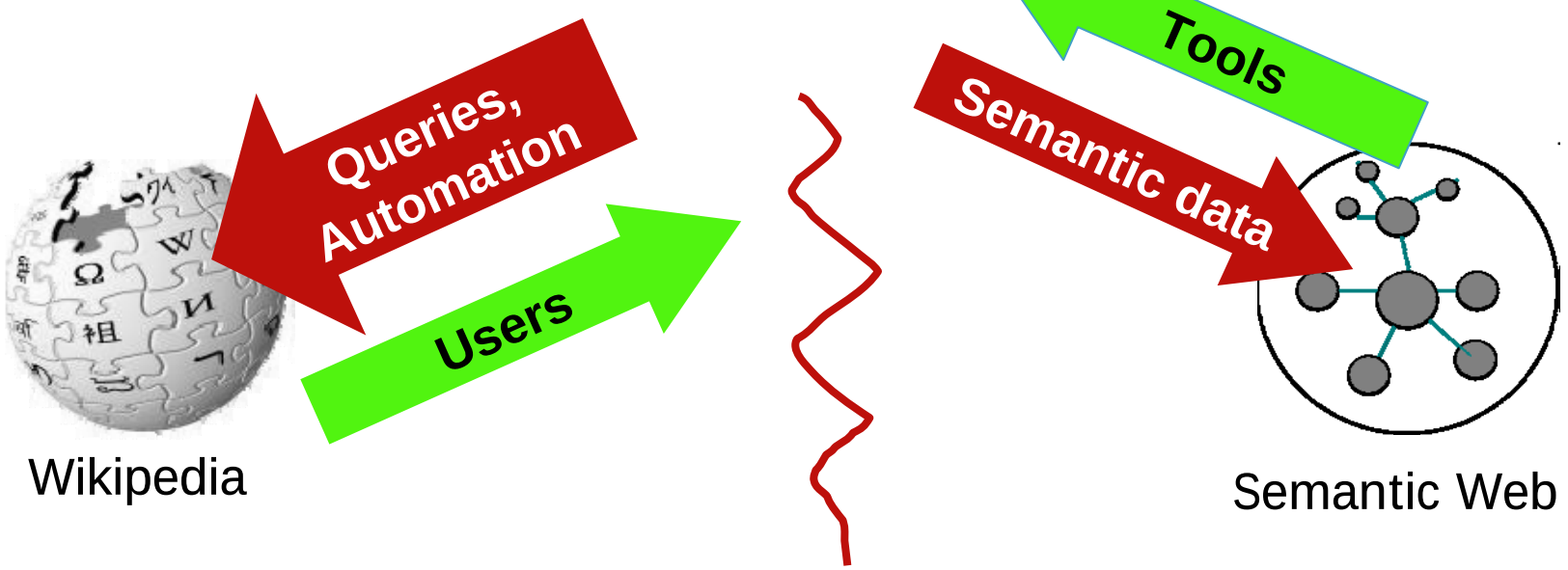
Analysis

- Wikipedia has many users and data,
  but: many manual processes, only text-based search
- Semantic Web has tools for information processing,
  sophisticated queries,
  but: few data



Goal: Marrying Wikipedia and the Semantic Web

- Wikipedia has many s and ,
  but: many manual pr text-based search
- Semantic W ation processing,
  sophisticated
  but: few data

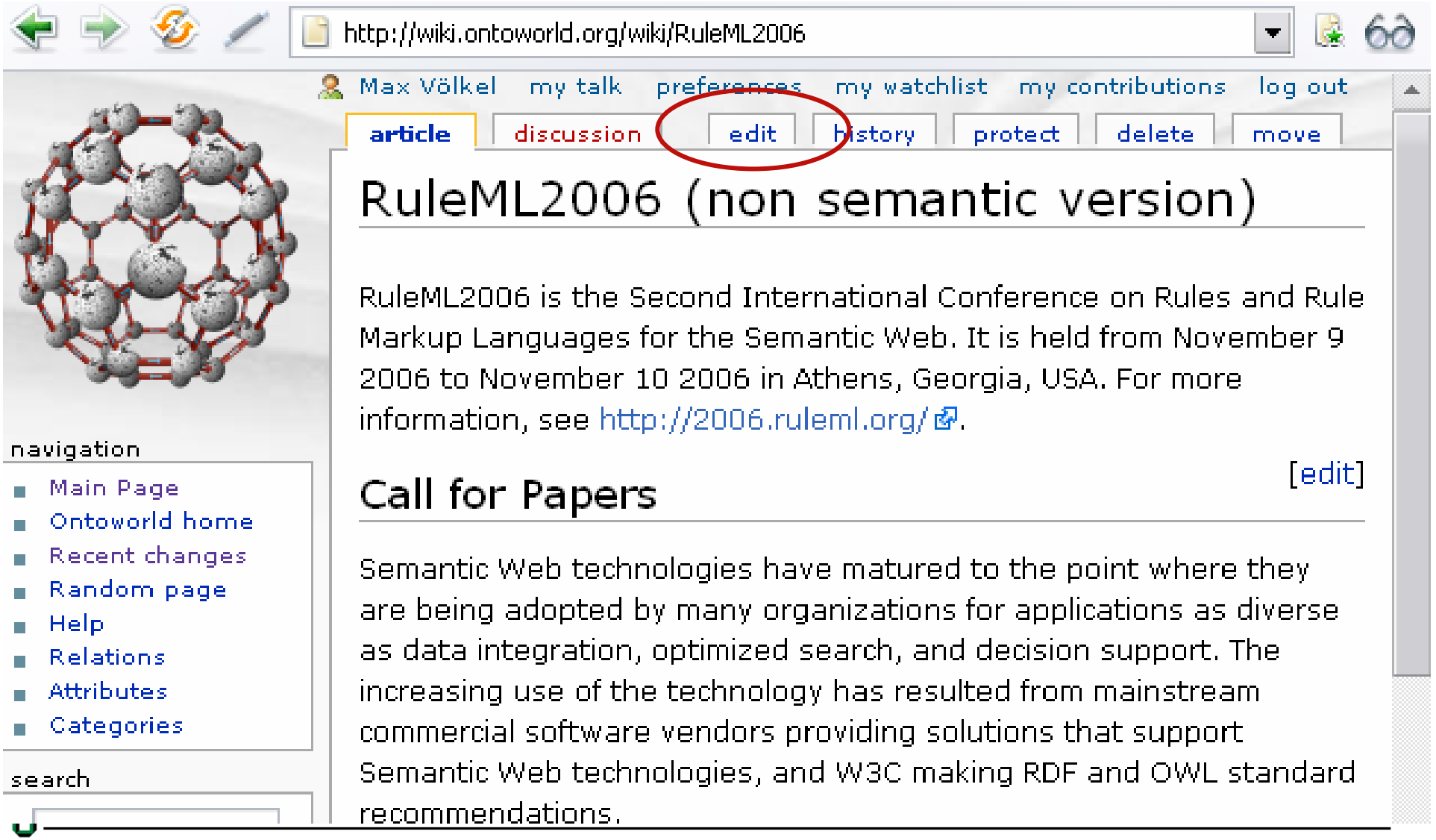


Requirements for

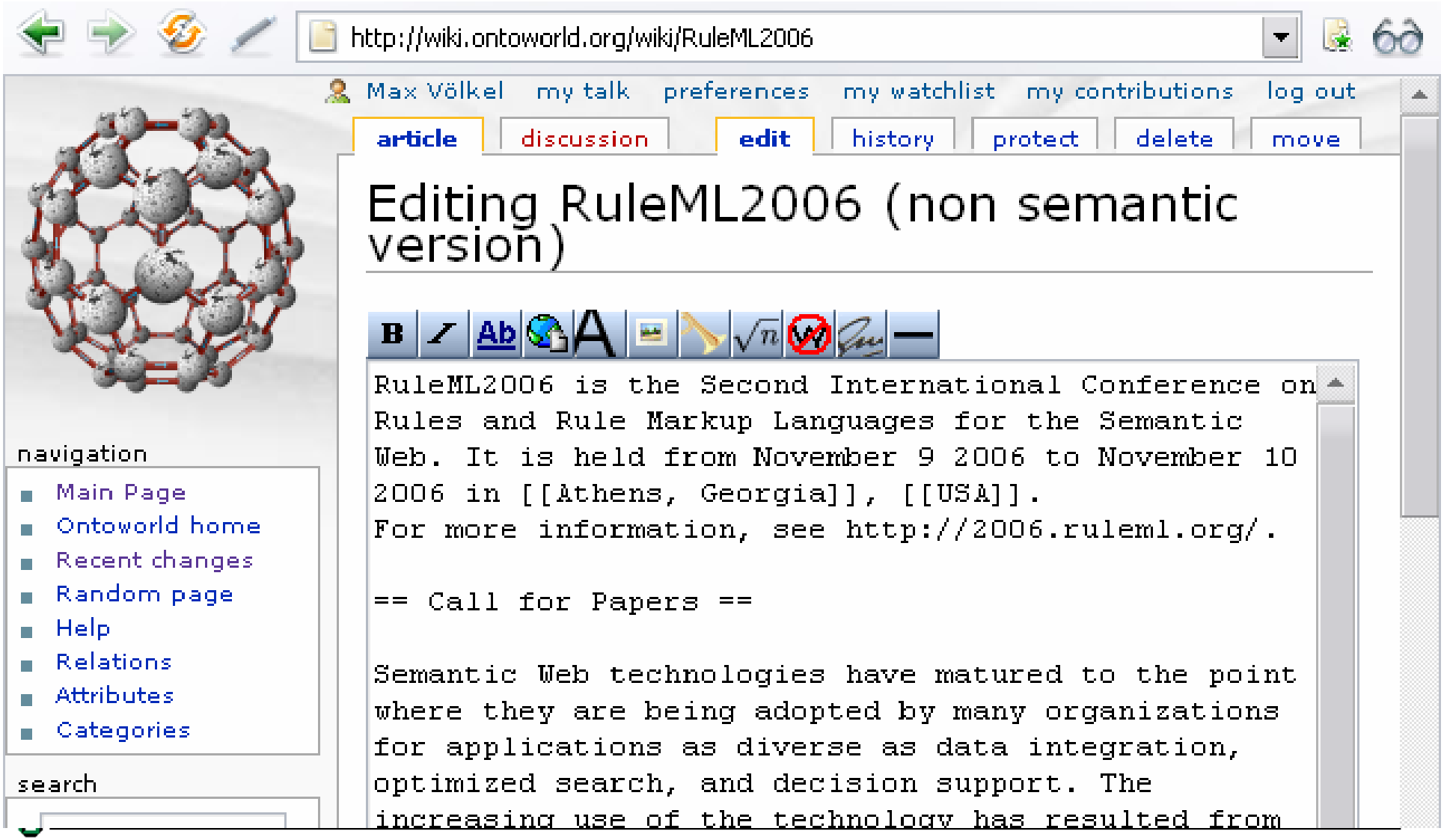
- Wikipedia:
  - Must be very easy to use
  - Must have immediate benefit for the users
    - Lists, Inconsistencies, Better Search
  - Must be efficiently implemented
    - Currently 12,000 hits/second

- Semantic Web:
  - Must have export of semantic data
  - Nice to have SPARQL access
  - Must integrate with existing vocabularies

We go to the article on the RuleML2006 conference ...



... and edit it


http://wiki.ontoworld.org/wiki/RuleML2006
Max Völkel
my talk
preferences
my watchlist
my contributions
log out
article
discussion
edit
history
protect
delete
move
Editing RuleML2006 (non semantic version)
navigation
Main Page
Ontoworld home
Recent changes
Random page
Help
Relations
Attributes
Categories
search
RuleML2006 is the Second International Conference on Rules and Rule Markup Languages for the Semantic Web. It is held from November 9 2006 to November 10 2006 in [[Athens, Georgia]], [[USA]].
For more information, see http://2006.ruleml.org/.
== Call for Papers ==
Semantic Web technologies have matured to the point where they are being adopted by many organizations for applications as diverse as data integration, optimized search, and decision support. The increasing use of the technology has resulted from

Editing RuleML2006 (non semantic version)

article | discussion | edit | history | protect | delete | move

RuleML2006 is the Second Internationa
Rules and Ru ges
for the Semant eb. It is held
from November 9 2006 to November
10 2006 in [[Athens, Georgia]],
[[USA]].
For more information, see
http://2006.ruleml.org/.

There is already an ordinariy link to the article of „Athens, Georgia“

Editing RuleML2006 (semantic version)

```
RuleML2006 is the Second
International
Rules and Rul
for the Semantic
from November 9 2006 to November
10 2006 in [[located in::Athens,
Georgia]], [[USA]].
For more information, see
http://2006.ruleml.org/.
```

Just say what the relation between this page (RuleML2006) and „Athens, Georgia“ is.

From links …

… in [[Athens, Georgia]], [[USA]]. …

… to typed links

… in [[located in::Athens, Georgia]], [[USA]]. …

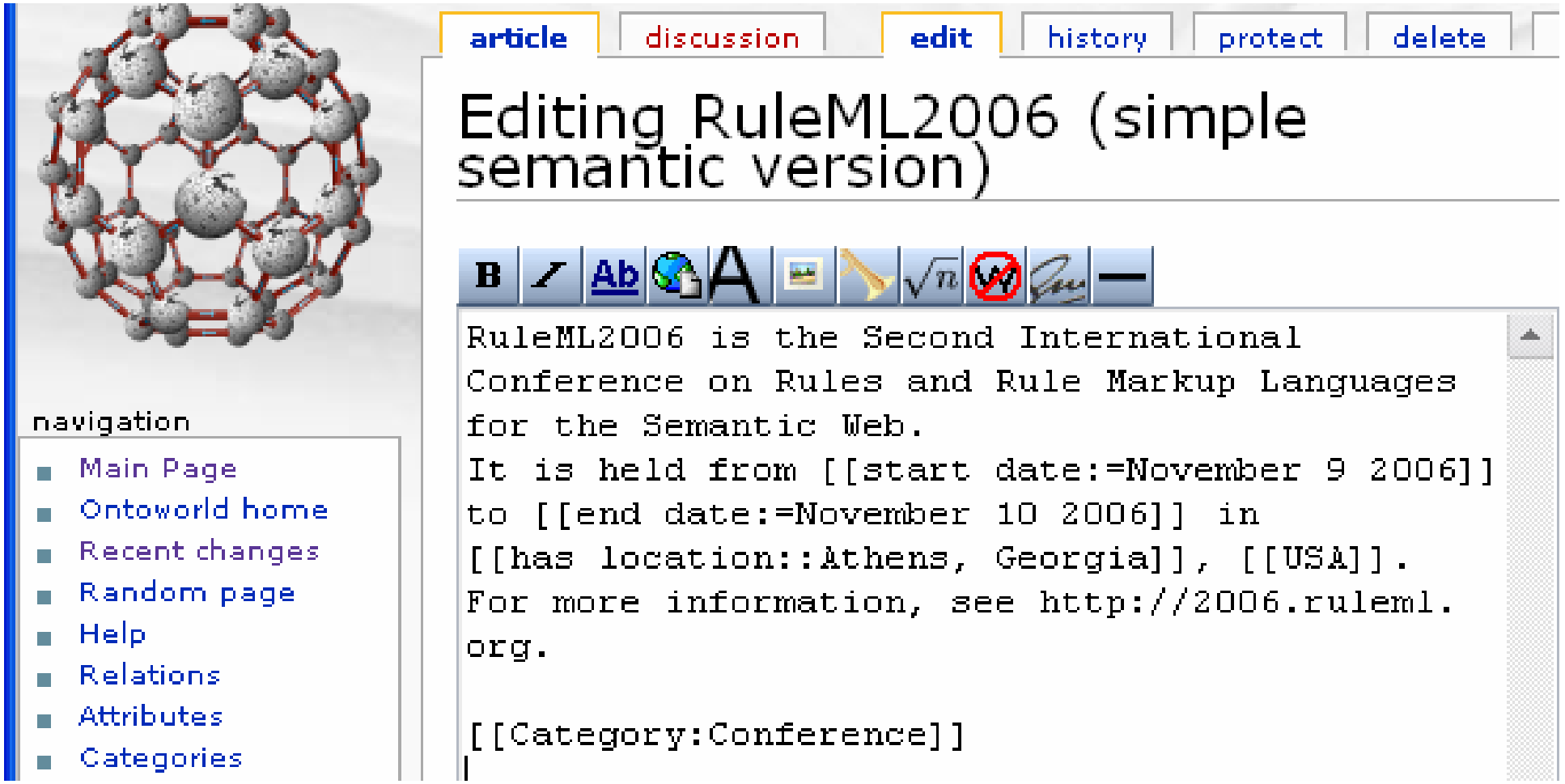
From values …

… It is held from November 9 2006 to November 10 2006 in…

… to attributes

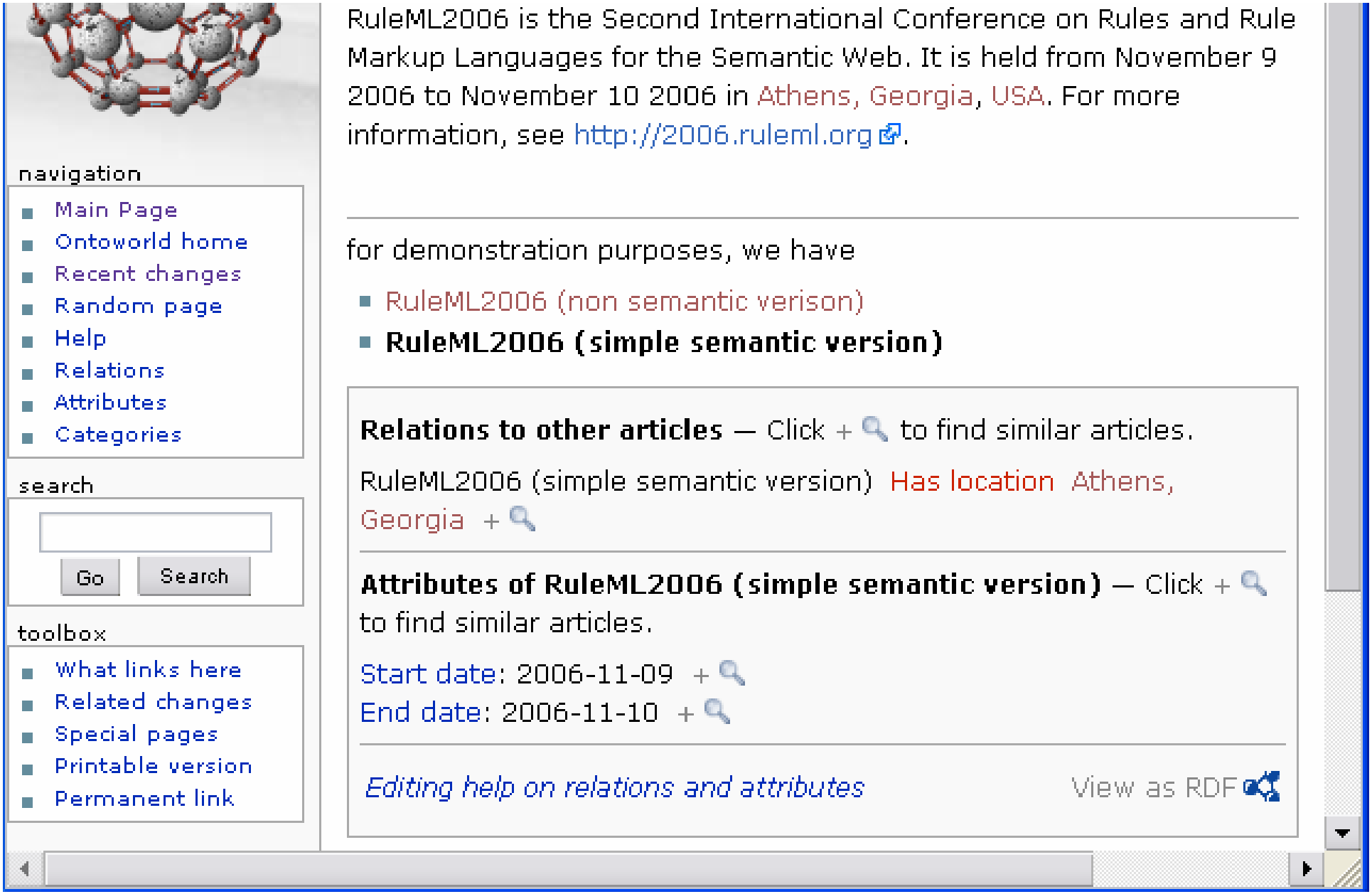
… It is held from [[start date:=November 9 2006]] to [[end date:=November 10 2006]] in…

Save.


article
discussion
edit
history
protect
delete
Editing RuleML2006 (simple semantic version)
RuleML2006 is the Second International Conference on Rules and Rule Markup Languages for the Semantic Web.
It is held from [[start date:=November 9 2006]] to [[end date:=November 10 2006]] in [[has location::Athens, Georgia]], [[USA]].
For more information, see http://2006.ruleml.org.
[[Category:Conference]]
navigation
Main Page
Ontoworld home
Recent changes
Random page
Help
Relations
Attributes
Categories

It looks exactly the same as before

What the humans see, when they scroll down



What the humans see, when they scroll down

Relations to other articles — Click + to find similar articles.

RuleML2006 (simple semantic version) Has location Athens, Georgia +

Attributes of RuleML2006 (simple semantic version) — Click + to find similar articles.

Start date: 2006-11-09 +

End date: 2006-11-10 +

Editing help on relations and attributes

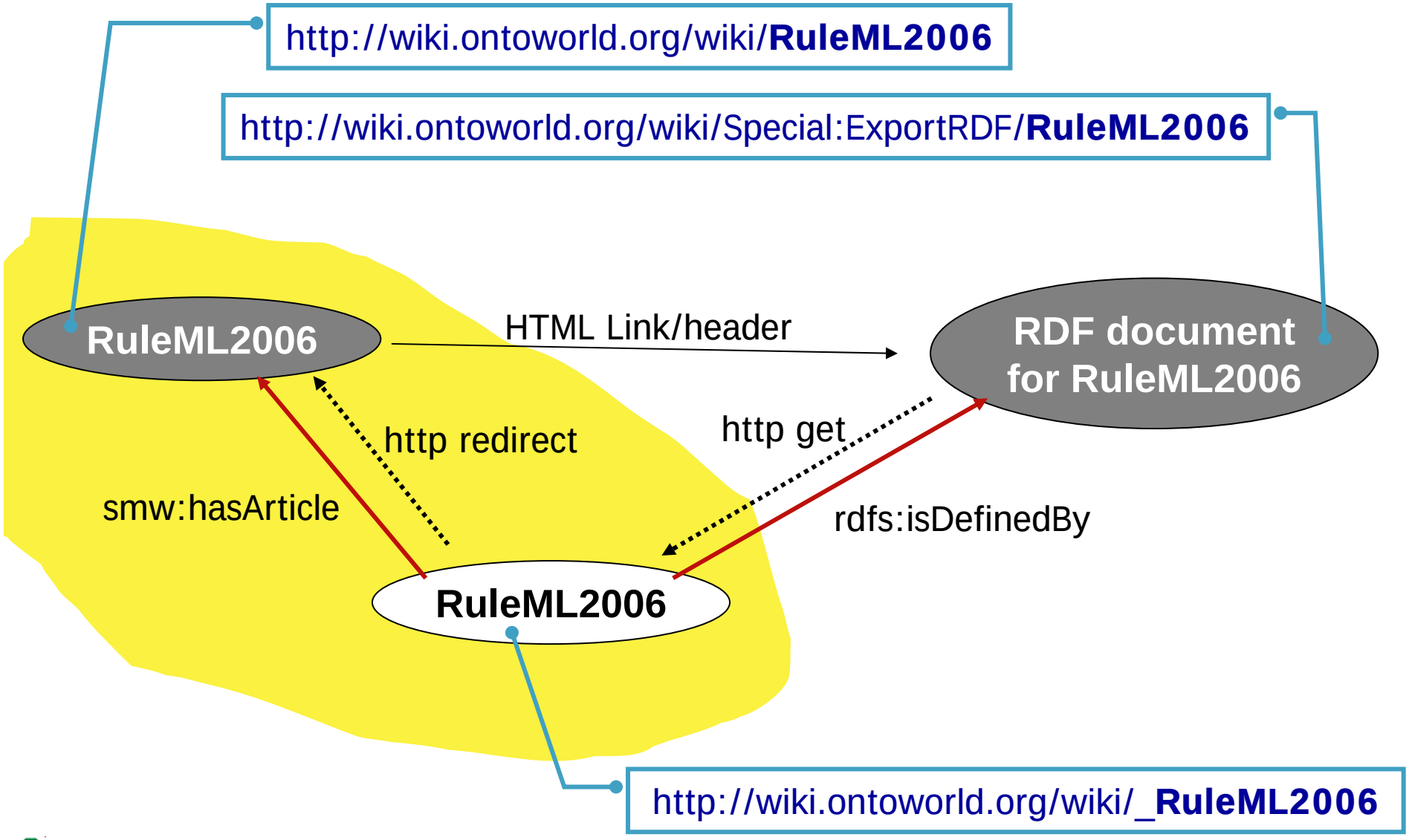
View as RDF

What the machines see

View as RDF http://wiki.ontoworld.org/index.php/Special:ExportRDF/RuleML2006

```
- <rdf:RDF>
    <!-- reference to the Semantic MediaWiki schema -->
  - <rdf:Description rdf:about="http://smw.ontoware.org/2005/smw#hasArticle">
      <rdfs:isDefinedBy rdf:resource="http://smw.ontoware.org/2005/smw"/>
    </rdf:Description>
  - <rdf:Description rdf:about="http://smw.ontoware.org/2005/smw#Thing">
      <rdfs:isDefinedBy rdf:resource="http://smw.ontoware.org/2005/smw"/>
    </rdf:Description>
    <!-- exported page data -->
  - <smw:Thing rdf:about="http://wiki.ontoworld.org/index.php/_RuleML2006">
      <rdfs:label>RuleML2006</rdfs:label>
      <smw:hasArticle rdf:resource="http://wiki.ontoworld.org/index.php/RuleML2006"/>
      <rdfs:isDefinedBy rdf:resource="http://wiki.ontoworld.org/index.php/Special:ExportRDF/RuleML2006"/>
      <rdf:type rdf:resource="http://wiki.ontoworld.org/index.php/_Category-3AConference"/>
      <rdf:type rdf:resource="http://wiki.ontoworld.org/index.php/_Category-3ATopic_Semantic_Web_query_languages"/>
      <rdf:type rdf:resource="http://wiki.ontoworld.org/index.php/_Category-3ATopic_Semantic_Web_rule_languages"/>
      <rdf:type rdf:resource="http://wiki.ontoworld.org/index.php/_Category-3ATopic_semantic_interoperability"/>
      <rdf:type rdf:resource="http://wiki.ontoworld.org/index.php/_Category-3ATopics_missing"/>
      <relation:Has_PC_member rdf:resource="http://wiki.ontoworld.org/index.php/_Sidney_Bailin"/>
      <relation:Has_PC_member rdf:resource="http://wiki.ontoworld.org/index.php/_Ruediger_Klein"/>
      <relation:Has_PC_member rdf:resource="http://wiki.ontoworld.org/index.php/_Ralph_Traphoener"/>
```

Information resources vs. abstract concepts (http range-14)



Did I get that right?

- Everybody can create any relation or attribute?
  - Yes
- Each relation and each attribute have their own wiki-article
  - Relation:located in
  - Attribute:population
- Can I have class hierachies, datatypes and all this?

Mapping Wiki Concepts to OWL

Semantic MediaWiki	Syntax	OWL
Link	[[Scotland]]	-
Relation	[[located in::Scotland]] [[contains::Caffeine]]	owl:ObjectProperty
Attribute	[[population:=5,062,011]] [[length:=17 km]]	owl:DatatypeProperty
Category	[[Category:Monarchies]]	owl:Class

Why typed links?

- Cheap annotations
  - Content (links, values) is already there, just mark them up
- Content and annotations are one
  - DRY – don‘t repeat yourself
  - Metadata doens‘t get out of sync
- Defined annotation process
  - Authors are annotators
- Locality of annotation
- User interface
  - No new tool to learn
  - Total control over annotations
- Wiki-style: No structure imposed
  - All relations and attributes can be used

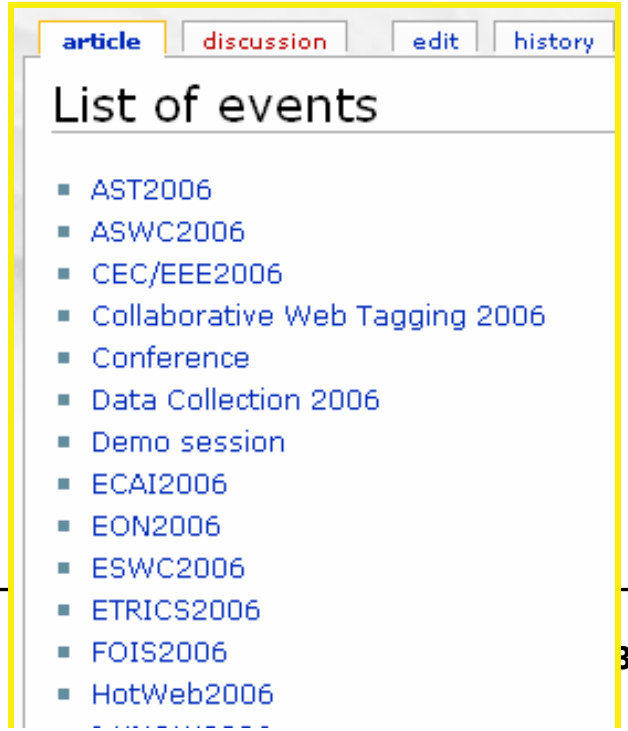
Benefits for the Semantic Web

- Wikipedia as URI source (see Workshop IRW2006)
  - Simple social process: Prefix article name with an underscore
- RDF/OWL Data export
  - Header links to RDF file (PiggyBank, OINK, Tabulator)
  - Per page or bulk
  - Everything has an rdf:label
  - All properties have rdfs:isDefinedBy
  - XSD datatypes fully supported
- Semantic Web food chain
  - SPARQL endpoint
  - SPARQL tools (SNORQL)

Benefits for Wikipedians: <ask> for your data

- Inline queries allow for questions like ...
  - ... movies from the 70s starring Sean Connery
  - ... list of events (all conferences and workshops)

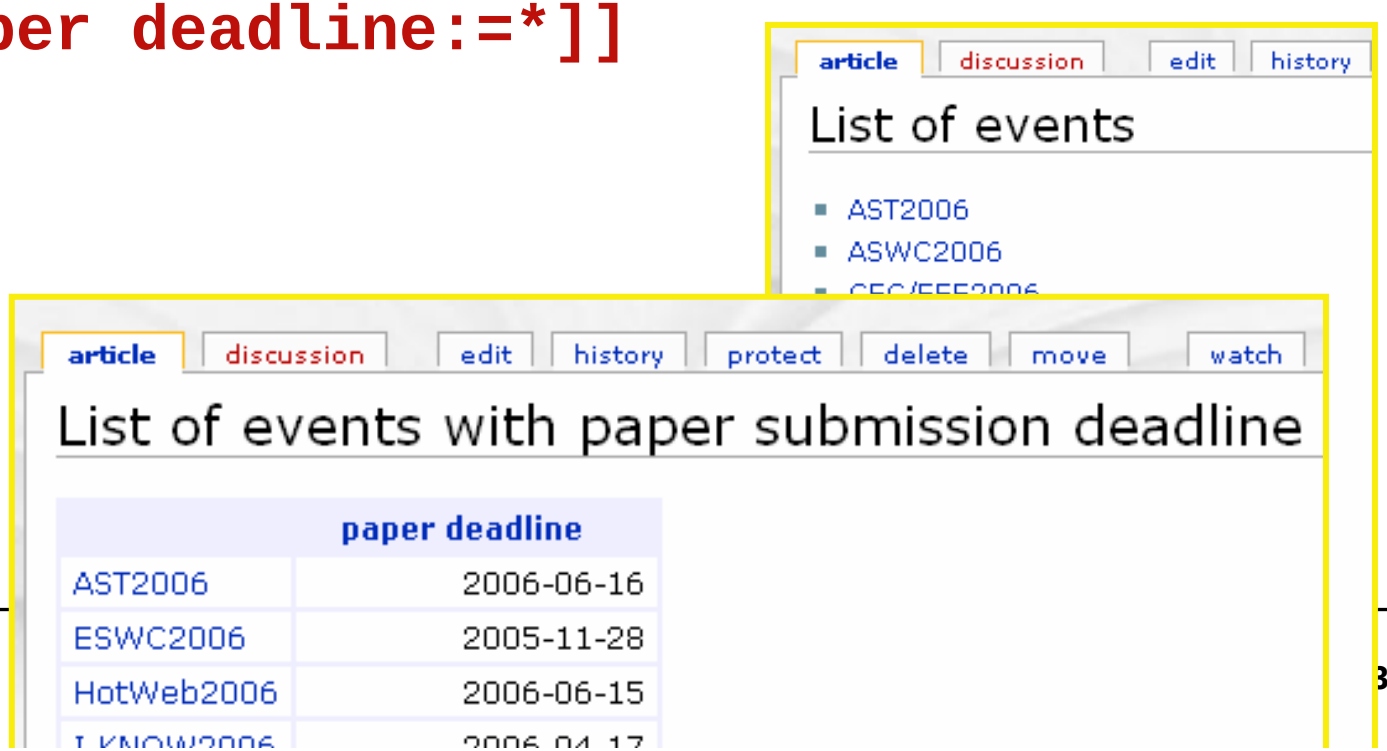
```
<ask format="ul" link="all">
     [[Category:Event]]
</ask>
```



Benefits for Wikipedians: <ask> for your data

- Inline queries allow for questions like ...
  - ... movies from the 70s starring Sean Connery
  - ... list of events with their deadline

```
<ask format="ul" link="all">
      [[Category:Event]]
      [[paper deadline:=*]]
</ask>
```



Benefits for Wikipedians: <ask> for your data

```
<ask format="ul" link="all">
 [[Category:Event]]
 [[paper deadline:=>June 1 2006]]
 [[paper deadline:=<December 31 2006]]
 [[title:=*]]
 [[paper deadline:=*]]
 [[Category:Topic Semantic Web query
languages]]
</ask>
```

	title	paper deadline
RuleML2006	Second International Conference on Rules and Rule Markup Languages for the Semantic Web	2006-06-05

Applications

- Automatic tables and lists
  - E.g. Countries sorted by area, population, alphabet, ...
- **Maintenance** with hand crafted checks
  - Does every country have one capital?
- Integration in applications
  - ```
    latte = wikipedia.get("Latte Macchiatto");
    print latte["contains"]
    ```
- Visualization and browsing
- ... And many unexpected ones
    ```

Who is using Semantic MediaWiki?

- Our research group
  - Wikicompany ~ 9.000 pages, 227 user accounts
  - The Bible Wiki ~ 40.000 pages, 156 user accounts
  - WWW2006 Social Wiki ~ 40 pages, 34 user accounts
  - Semantic eGovernment Community Portal, ~ 30 pages
  - ScubaWiki, ~ 170 pages, 33 user accounts
  - NeOn Wiki ~10 pages
  - this wiki = Wiki.Ontoworld.org ~400 pages, 158 user accounts
  - SourceryForge ~1864 pages, 272 user accounts

Features

- Linking of concepts with external URIs
  - **`[[equivalent URI:= … ]]`**
  - → owl:equivalentClass, owl:equivalentProperty, owl:sameAs
- OWL import (generates pages from ontology)
- All relations and attributes can be documented
- Re-use of existing Category-system
- Annotating templates for a quick win
- Correct RDF export

→ Flexible system for collaboratively creating content with semantic annotations

Future Work

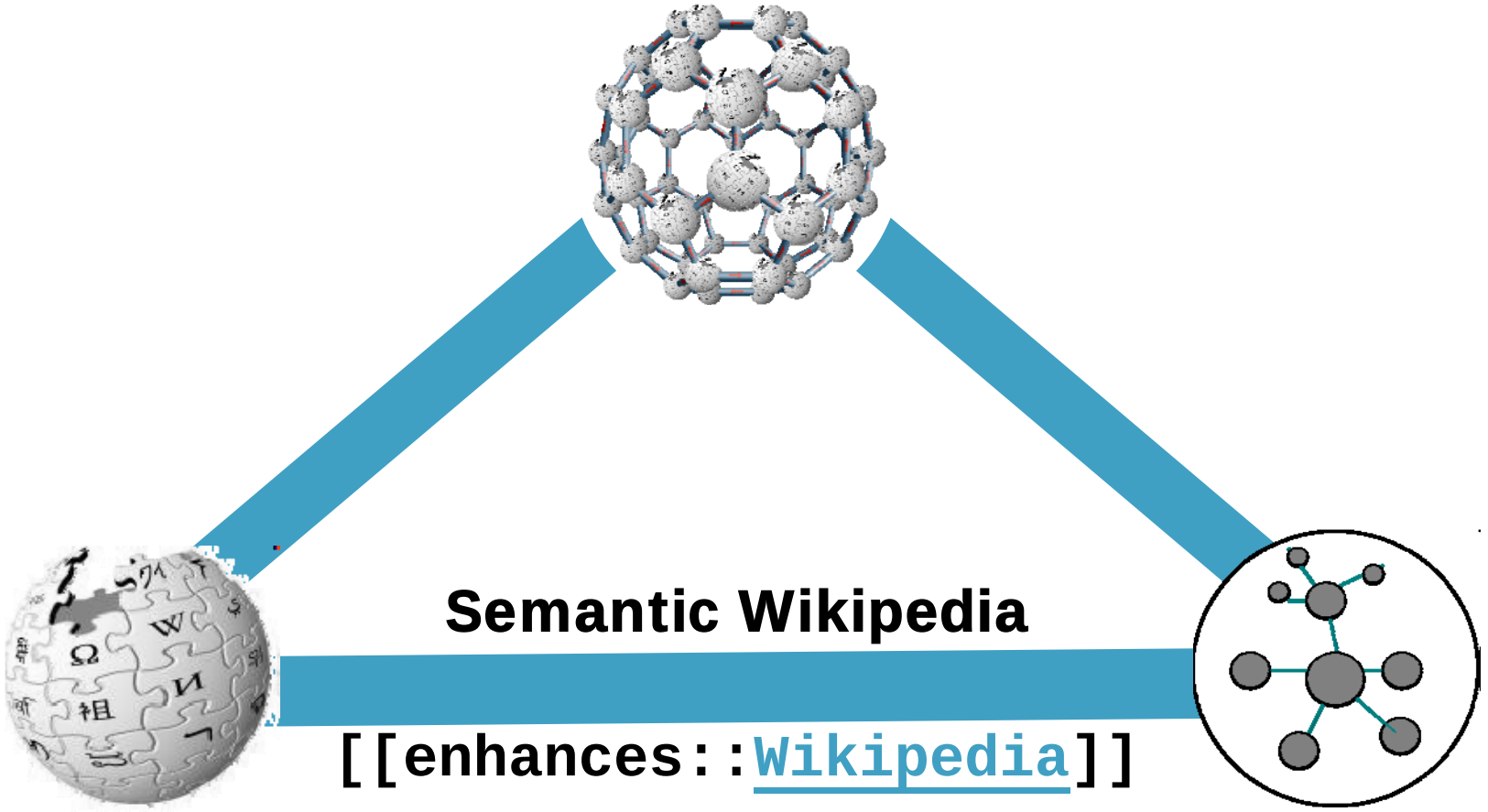
- Performance
- Performance
- Scalability

- More Expressiveness
  - Transitive, symmetric, and inverse relations
- User Interface
  - Suggested relations and attributes
- Evaluation

Conclusions

- **Annotation for the masses**
- **“Soft” introduction**
  - People are often scared about Semantic Web
  - Semantic features can be ignored
- **Content and metadata are one**
- **Linking to existing ontologies**
- ***Create the egg of the chicken-and-egg problem***

Thank You!



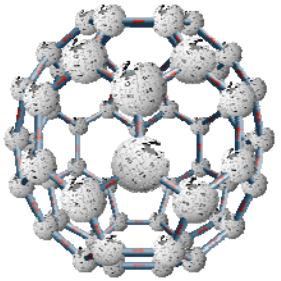
Conclusions

- Annotation for the masses
- “Soft” introduction
- Content and metadata are one
- Linking to existing ontologies
- *Create the egg of the chicken-and-egg problem*

Semantic Wikipedia

`[[enhances::Wikipedia]]`

Comments or questions?

Questions and Answers

- Everybody can create any relation or attribute?
  - Yes
- How are Relations created?
  - You just use them!
- Can people agree on relation types?
  - They did already with categories. We hope it works again.
  - Every Relation has it‘s own wiki page for documentation.
    - Relation:located in, Attribute:population
- Can I make statements about relations?
  - Yes, but no OWL-Metamodelling