

# Information Modelling With a Semantic MediaWiki

Karin Haenelt

Shanghai, December 2008



# Contents

- Application
- Wiki
  - Description
  - Hypertext Functionality
  - Knowledge Representation Functionality
  - Database Functionality
- Modelling of the Catalogue of Web Services
- Discussion of the Modelling
- Desired Extensions
- Conclusion and Future Research

# Application

- Application
- Wiki
  - Description
  - Hypertext Functionality
  - Knowledge Representation Functionality
  - Database Functionality
- Modelling of the Catalogue of Web Services
- Discussion of the Modelling
- Desired Extensions
- Conclusion and Future Research

# Application: Catalogue of Web Services

- Online Shops
- Customer self services (management of data or machines)
- Courses
- Library Catalogues
- Download pages
- ...

# Task and Requirements

- Task
    - descriptions of the services
    - providing means for offering and finding the services
      - structured presentations
      - enhanced search capabilities
    - enabling access to the services
  
  - Requirements: handling of
    - unstructured, semi-structured and fully structured data
    - high volume of contributions
    - high volatility of contributions (frequent updates)
-

# Approach: Reasons for Choosing a Wiki

- Versatility of data structures
  - hypertext functionality
  - knowledge representation functionality
  - database functionality
- various possibilities of data search and presentation
- Collaboration
  - management of distributed contributions
  - sharing of workload between lots of contributors

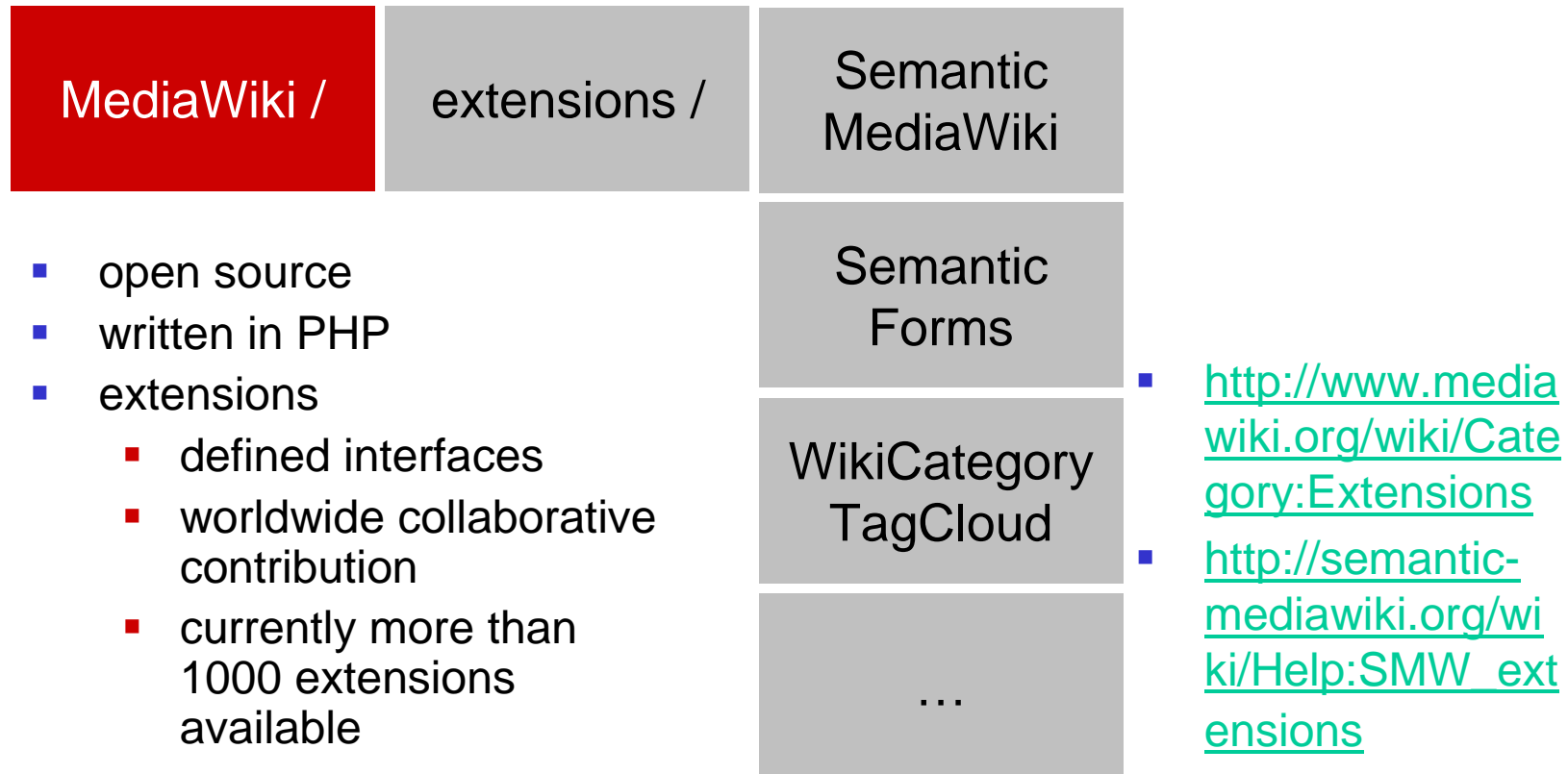
# Wiki

- Application
- Wiki
  - Description
  - Hypertext Functionality
  - Knowledge Representation Functionality
  - Database Functionality
- Modelling of the Catalogue of Web Services
- Discussion of the Modelling
- Desired Extensions
- Conclusion and Future Research

# Wiki, MediaWiki and Semantic MediaWiki

- **What ist a Wiki?**  
a web-based system which allows for the collaborative creation and change of web pages
- **MediaWiki**
  - free Wiki-software
  - originally written for Wikipedia
  - <http://www.mediawiki.org>
- **Semantic MediaWiki**
  - extension of MediaWiki
  - delepment coordination: AIFB institute of Karlsruhe University
  - [http://www.semantic-mediawiki.org/wiki/Semantic\\_MediaWiki](http://www.semantic-mediawiki.org/wiki/Semantic_MediaWiki)

# System Architecture of MediaWiki



# Semantic MediaWiki: Modelling Example

## MediaWiki

- simple link

### Product X

Product X is an e-learning web service offered by the **[[Fraunhofer Gesellschaft]]**

Product X → Fraunhofer Gesellschaft

## Semantic MediaWiki

- semantic annotation
- semantic link

### Product X

Product X is an e-learning web service offered by the **[[Service Provider:: Fraunhofer Gesellschaft]]**

Product X  $\xrightarrow{\text{Service Provider}}$  Fraunhofer Gesellschaft



# Semantic Annotation

- format

Page  
[[Property:: Value]]

data types  
of values:

page,  
string,  
number

- example

Product X  
Product X is an  
[[service domain:: e-learning]] web service  
offered by the [[service provider:: FhG]]

...

(Product X, service domain, e-learning)

(Product X, service provider, FhG)

# Data Types for Property Values

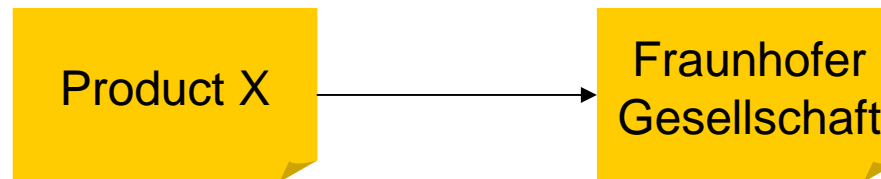
- [Type:String](#) (text strings)
- [Type:Number](#) (integer and decimal numbers with optional exponent)
- [Type:Page](#) (links to pages, the default)
- [Type:Boolean](#) restricts the value of a property to true/false (also 1/0 and yes/no).
- [Type:Text](#) is like Type:String but can have unlimited length;
- [Type:Temperature](#) can't be user-defined since converting temperature units is more complicated than multiplying by a conversion factor.
- [Type:Geographic coordinate](#) describes geographic locations. It recognizes different forms of geographic coordinates. Using [service links](#) it can dynamically provides links to online map services.
- [Type:Date](#) specifies particular points in time. *This type is still somewhat experimental, but may feature complex conversions between (historic) calendar models in the future.*
- [Type:URL](#) displays an external link to its URL object.
- [Type:Annotation URI](#): properties of this type are interpreted as relations to external objects, denoted by the URI. They are special since they are interpreted as *annotation properties* on export. See the type's page for documentation.
- [Type:Email](#) displays an e-mail address as a link (with mailto:).

# Interpretations of the Semantic Annotations

- Hypertext Link Functionality

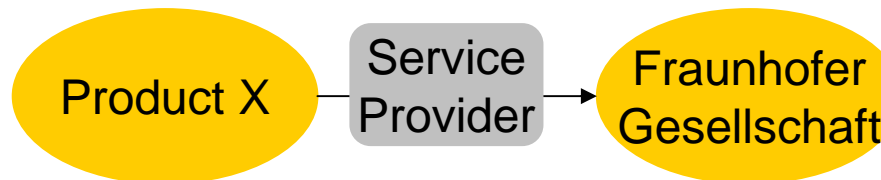
- linked pages

(if data type of property value is „page“)



- Knowledge Representation Functionality

- semantic net



- Data Base Functionality

- n-tuples



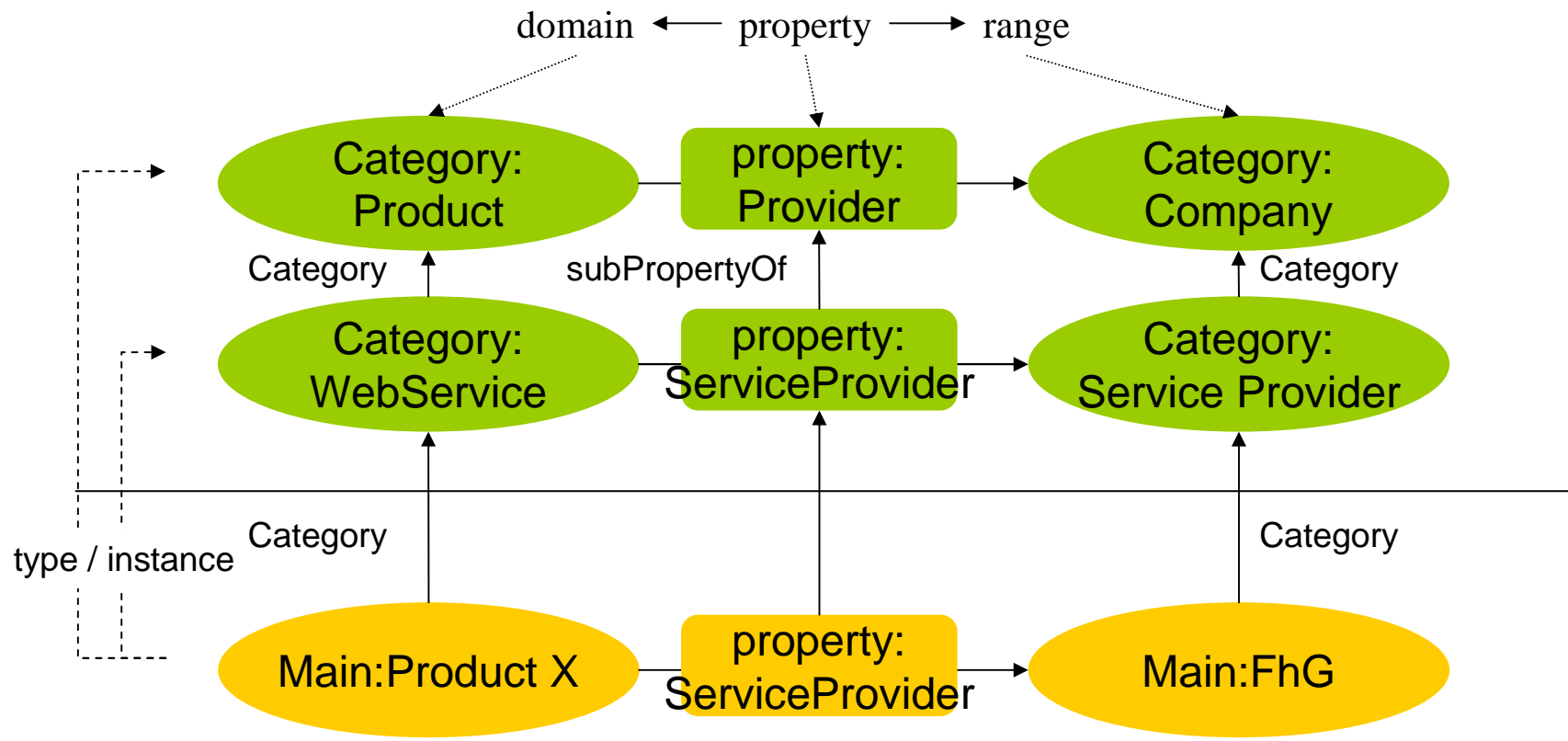
# Hypertext Functionality

- web-based text editor
- simplified notation for generation of HTML-pages
- simplified notation for links between pages
  
- version management

# Knowledge Representation Functionality

- Data model: **RDF**
    - classification and inheritance
    - no inferences
    - allows for the construction of **inheritance networks**
    - components
      - categories and category hierarchies
      - properties and property hierarchies
      - values and value types
  - notation language: MediaWiki-specific (page, property::value)
  - RDF and OWL interfaces
-

# Knowledge Representation: Example

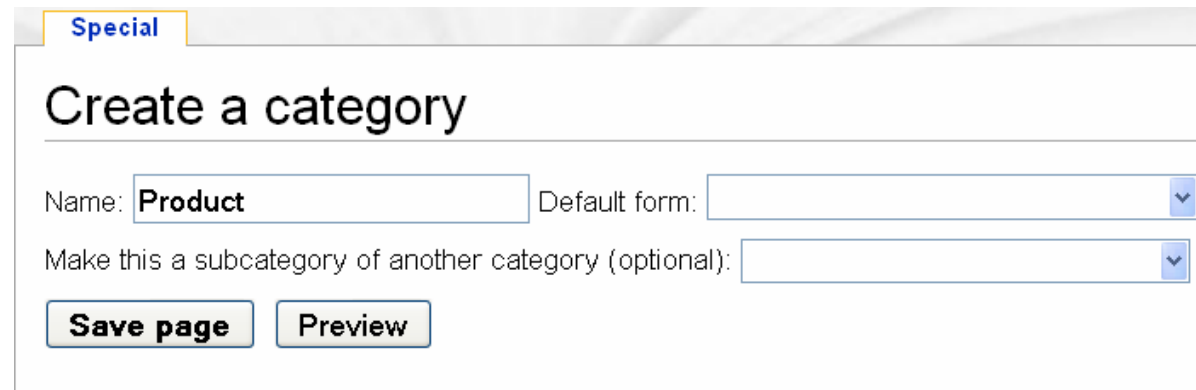


# Creation of Structured Inheritance Networks

- similar to creation of text pages
  - special namespaces
    - main: for text pages
    - category: for categories
    - property: for properties
  - **concepts** and **properties** are defined by creating a page in the respective namespace with the name of the concept or property
  - **subcategories** are defined by defining a category as belonging to another category
  - **subproperties** are defined with the special attribute „subproperty“
  - definition of concepts and properties not compulsory in Semantic MediaWiki; undefined concepts and properties can also be used in annotations
-

# Categories: Example: Category: Product

- Creation of Category Page using the extension „Semantic Forms“



Special

## Create a category

Name:  Default form:

Make this a subcategory of another category (optional):

- Category Page



Category Discussion Edit History Delete

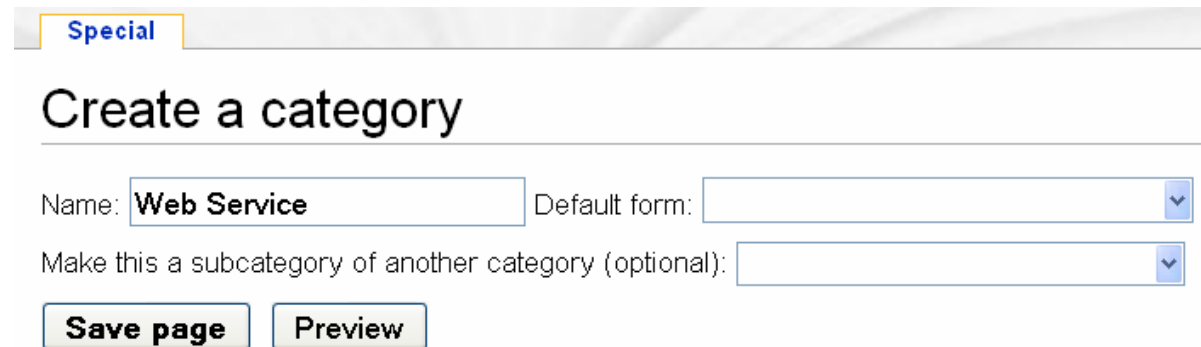
## Category:Product

This is the category Product.

*This category currently contains no pages or media.*

# Categories: Example: Category: Web Service

- Creation of Category Page using the extension „Semantic Forms“



**Special**

## Create a category

Name:  Default form:

Make this a subcategory of another category (optional):

**Save page** **Preview**

- Category Page



**Category** Discussion Edit History Delete

## Category:Web Service

This is the category Web Service

*This category currently contains no pages or media.*

## Subcategories: Example:

- Entering a super-category in a category page

Category: Product

Category↑

Category: WebService

Category Discussion Edit History Delete

Editing Category:Web Service

B / Ab G A [ ] √ π W [ ] -

[[Category:Product]]

Web Service (0)'. A red arrow points from the text 'The super-category page then lists the subcategories' to the 'Web Service' link."/>

Category Discussion Edit History Delete

### Category:Product

This is the category Product.

#### Subcategories

This category has only the following subcategory.

#### W

- [+][Web Service](#) (0)

- The super-category page then lists the subcategories

# RDF and OWL export and import

- `<rdf:RDF xmlns:rdf=http://www.w3.org/1999/02/22-rdf-syntax-ns#  
xmlns:ex="http://.../myExample.de/">  
<rdf:Description rdf:about="http://.../Product X">  
<ex:Service Provider>  
<rdf:Description rdf:about="http://.../Fraunhofer Gesellschaft">  
</rdf : Description>  
</ex : Service Provider>  
</rdf : Description>  
<rdf :RDF/>`

# Database Functionality

- Interpretation of semantic links as n-tuples
- Queries
- Presentation of query results as tables
- Aggregation of attributes  
(with templates and extension „Semantic Forms“)

# Semantic Search: Query Language

- Query

- ① query description
- ② printout statements
- ③ appearance parameters

Page Discussion Edit History Delete

## Editing AskServiceAttributes



**B** / **Ab**  **A**      

```
① {{ #ask: [[Categorie:Web Service]]  
② |? Service Provider  
}}
```

Page Discussion Edit History Delete Move Pro

- Result: sortable table

## AskServiceAttributes

	 Service Provider
Product X	Fraunhofer Gesellschaft
...	...

# Semantic Search: Query Language

SelectCondition	[[Service Provider::Fraunhofer Gesellschaft]]
Conjunction	[[Service Provider::Fraunhofer Gesellschaft]] [[Service URL::*]] a set of conditions is AND-related
Disjunction	[[Service Provider::Fraunhofer Gesellschaft  FhG]].
Star-Operator	* for arbitrary values
Subqueries	<ul style="list-style-type: none"><li>- Definition of a Query [[Category:City]] [[located in::Italy]]</li><li>- Use of a Query in a further Query [[Category:Actor]] [[born in::&lt;q&gt;[[Category:City]] [[located in::Italy]]&lt;/q&gt;]]</li></ul>

# Data Aggregation with Templates and Semantic Forms

- **Attributes**
  - attribute: `ServiceName` | hasType: String
  - attribute: `ServiceURL` | hasType: URL
- **Templates:**  
Aggregations of Attributes
  - `{{T-ServiceID`  
`| ServiceName=`  
`| ServiceURL=`  
`}}`
- **Forms:**  
Aggregations of Templates
  - `{{#forminput:F-ServiceEntry`  
`{{{for template T-ServiceID`  
`{| class="formtable"`  
`! Service Name:`  
`| {{{field|ServiceName}}}`  
`...`  
`{{{end template}}}`  
`}}`
- useful for structured input
- no query mechanisms for aggregations

# 3 Modelling of the Catalogue of Web Services

- Application
- Wiki
  - Description
  - Hypertext Functionality
  - Knowledge Representation Functionality
  - Database Functionality
- **Modelling of the Catalogue of Web Services**
- Discussion of the Modelling
- Desired Extensions
- Conclusion and Future Research

# Modelling of the Catalogue of Web Services

- Creation of Pages with Semantic Forms
  - fully structured information
  - semi-structured information
  - free text
  - semantic tagging

Information	Contact	Description	Categories / Tags
Service Name:	<input type="text"/>		
Service Function:	<input type="text"/>		
Service URL:	<input type="text"/>		

- Access to Services (via Service URL)
- Search Facilities

# Sample Page

Page Discussion Edit with form Edit History Delete Move Protect

## Product X

Product Factbox



**Contents** [hide]

- 1 Product
  - 1.1 Description

**Description** [edit]

Product X is an e-learning web service offered by the **Fraunhofer Gesellschaft**

**Brief Information**

Service Name	Product X
Service Function	E-learning Web Service
Service URL	<a href="http://example.org">http://example.org</a> 
alternative Names	
Service Provider	

Category: Web Service

# Search Facilities

- Information in Wikis: easy to create – difficult to find?
- Search facilities:
  - full text search (text view)
  - semantic navigation (knowledge representation view)
  - semantic search (database view)

# Search Facility: Full Text Search

Search

- 1. step:  
full title match

- 2. step:  
partial title match

matching string +  
semantic information

- 3. step:  
text match

Special

## Search results

You searched for **product** (all pages starting with "product" | all pages that link to "product")

### No page title matches

There is no page titled "**product**". You can [create this page](#).

For more information about searching Wiki01, see [Help](#).

Showing below **2** results starting with #1.

View (previous 20) (next 20) ([20](#) | [50](#) | [100](#) | [250](#) | [500](#))

### Page title matches

- Product X  
|DienstName=**Product X** **Product X** is an e-learning web service offered by the  
[[Service Provider:: Fraunho ...  
286 B (32 words) - 15:13, 24 October 2008

### Page text matches


- Product X  
|DienstName=**Product X** **Product X** is an e-learning web service offered by the  
[[Service Provider:: Fraunho ...  
286 B (32 words) - 15:13, 24 October 2008

# Search Facility: Semantic Navigation: Sidebar

- Sidebar: Direct Access to preselected pages
- Function:
  - offering a particular view on the site
  - guiding users to different views on the services
  - guiding collaborators

<b>Content</b>
<ul style="list-style-type: none"><li>■ Main Page</li><li>■ Services A-Z, List</li><li>■ Services A-Z, Table</li><li>■ Special Tables</li></ul>
<b>Navigation</b>
<ul style="list-style-type: none"><li>■ TagCloud</li><li>■ CategoryTree</li></ul>
<b>Collaborate</b>
<ul style="list-style-type: none"><li>■ Enter Web Service</li></ul>
<b>Create</b>
<ul style="list-style-type: none"><li>■ Category</li><li>■ Property</li><li>■ Template</li><li>■ Form</li></ul>

# Search Facility: Semantic Navigation: Categories, Properties and Instances

- Example 
- **alphabetical lists of instances** of
  - individual **categories** (e.g. web services, web shops)
  - individual **properties**
- accessible on the wiki page of the individual categories or properties

Pages using the category „Web Shop“

- **Fraunhofer IAO-Shop**
- **Fraunhofer-Vision Web-Shop**
- ...

Pages using the property “Service-URL“

[Baufachinformation.de](#) + ⓘ <http://www.baufachinformation.de> + 🔍

[CiTyServer3D](#) + ⓘ <http://www.igd.fhg.de/igd-a5/projects/index.html> + 🔍

# Search Facility: Semantic Navigation: Category Tree

- Example →
- Theoretical Status
  - Structure: selected trees from the structured inheritance network
  - Function
    - extension of the vocabulary of the wiki pages
    - providing one or more predefined views on wiki
- Practical Use
  - search assistance
  - „guided tour to services“
  - product presentation for selected target groups

- Service Types
  - Web Shops
    - Fraunhofer IAO-Shop
    - Fraunhofer-Vision Web-Shop
    - ...
  - ...
- Service Domains
  - Adaptronics
    - ..
  - Water



# Search Facility: Semantic Navigation: Tags

- Example
  - Theoretical Status
    - Structure:
      - set of tags with weights
      - assigned by Wiki-authors and -users
    - Function:
      - tags: extension of the vocabulary of the text corpus
      - weights: (changing) distributions of writers' and of users' views and interests on the wiki pages
  - Practical Use
    - search assistance
    - product presentation (influence on search behaviour)
-

# Search Facility: Semantic Search

- Query

- ① query description
- ② printout statements
- ③ appearance parameters

Page Discussion Edit History Delete

## Editing AskServiceAttributes



**B** / **Ab**  **A**      

```
① {{ #ask: [[Categorie:Web Service]]  
② |? Service Provider  
}}
```

Page Discussion Edit History Delete Move Pro

- Result: sortable table

## AskServiceAttributes

	 Service Provider
Product X	Fraunhofer Gesellschaft
...	...

# Search Facility: Semantic Search

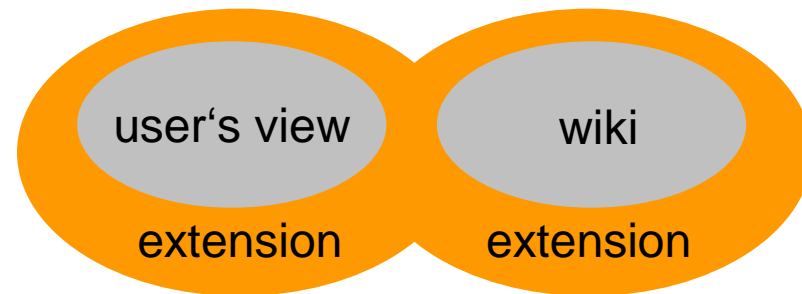
- Example
- Theoretical Status
  - Views on database (tables with selected attributes)
  - queries may be
    - predefined by wiki authors and
    - specified by (experienced) users
- Practical Use
  - presentation of selected data for different user groups:
    - end users
    - maintenance staff / WikiAdmins
    - developers

## 4 Discussion of the Modelling

- Application
- Wiki
  - Description
  - Hypertext Functionality
  - Knowledge Representation Functionality
  - Database Functionality
- Modelling of the Catalogue of Web Services
- [Discussion of the Modelling](#)
- Desired Extensions
- Conclusion and Future Research

# General Task: Extension of Matching of users' worlds and wiki world

- users' worlds and wiki world do not always coincide
- wiki-approach: extension of
  - wiki-descriptions
  - users' views
- theoretical status of extensions
  - extension and change of knowledge
  - extension and change of perspectives on knowledge



# Extension Methods

- manually, individually
  - quality professional
  - view individual, static / inflexible
  - vocabulary controlled or free
  - effort for creation & maintenance enormous; work of individuals
- manually, collaboratively
  - quality mass of contributors as corrective
  - views manifold, dynamic
  - vocabulary free
  - effort low; many individuals share the work load
- automatically
  - quality depending on state of the art
  - views manifold, dynamic
  - vocabulary determined by the involved sources

categories

categories  
tags

## Extension Method: Category Tree

- possible developers
  - administrators
  - users
- network of categories is handcoded
- categories for the tree are selected manually
- categories in the wiki pages are handcoded
- list of services with particular categories are assembled automatically
  
- category tree does not necessarily reflect the actual state of the wiki contents

## Extension Method: Categories as Tags

- Tagging mechanism of „WikiCategoryTagCloud“: categories serve as tags
- intermixture of
  - editorially maintained categories
  - folksonomy categories
- advantages
  - views: manifold, multifaceted
  - vocabulary: free, multilingual
  - costs: low

# Mixture of Tags and Categories: Problems and possible Solutions

- linguistic variations must be merged manually
  - e-Learning, E-Learning, eLearning, ...
    - dictionary
    - automatic detection (clustering)
- category structure does not reflect the latest state of tags
  - additional use of tag clustering
- update of page tags: distribution of new tags to all appropriate pages
  - by WikiAdmins?
  - by users? incrementally during usage
  - less problematic if trends of use are to be reflected
- name conflicts can corrupt the editorial category system
  - separate tagging and editorial category system

## 5 Desired Extensions

- Application
- Wiki
  - Description
  - Hypertext Functionality
  - Knowledge Representation Functionality
  - Database Functionality
- Modelling of the Catalogue of Web Services
- Discussion of the Modelling
- [Desired Extensions](#)
- Conclusion and Future Research

## Desired Extension: Granularity of Search Results

- current granularity: [wiki page](#)
- desired extension: [direct search/navigation for other functionalities](#) of the Wiki-System, such as
  - direct navigation in the inheritance network
  - direct search for all values of a property
- current method:
  - manual coding of inline query
  - automatic generation of a table of pages containing the selected attributes
  - sorting the table by clicking to a column

## Desired Extension: Separation of the Attributes „Tag“ and „Category“

- Current state: tags are treated as categories
- possible undesired side effects of intermixture

## Desired Extension: Dictionaries

- alternative page names
  - currently managed via redirect pages (cumbersomely)
- alternative tag names
  - eLearning, e-Learning, E-Learning, ...
- ...

## 6 Conclusion and Future Research

- Application
- Wiki
  - Description
  - Hypertext Functionality
  - Knowledge Representation Functionality
  - Database Functionality
- Modelling of the Catalogue of Web Services
- Discussion of the Modelling
- Desired Extensions
- Conclusion and Future Research

# Conclusion

- Semantic MediaWiki provides many advantages for modelling a product catalogue
  - integration of hypertexts, databases and knowledge base
  - collaboration
- Desired extensions can be
  - put on a request-list  
[http://www.mediawiki.org/wiki/Project:Extension\\_requests](http://www.mediawiki.org/wiki/Project:Extension_requests) ,  
or
  - contributed to the open source library of extensions (interfaces are well-defined)  
<http://www.mediawiki.org/wiki/Category:Extensions>

# General Status of a Semantic Wiki and Tasks for Future Research

- A Semantic Wiki
    - can provide more views on data than previous tools
    - cannot solve the dynamic aspect of information retrieval:
    - task of being a users agent that searches and finds for every user the best solution in any particular task
  - Tasks for Future Research
    - further development and integration of the aspects: hypertext, database, knowledge base
    - generating views on a site dynamically, based on
      - a user's actual situation
      - a knowledge context
      - a social context
      - an author's strategy
    - combination with information extraction
    - combination with semantic/pragmatic agents
-