



A Browser-based Tool for Collaborative Distributed Annotation for the Semantic Web

Onni Valkeapää and Olli Alm
Semantic Computing Research Group (SeCo)
Helsinki University of Technology (TKK),
Laboratory of Media Technology
and
University of Helsinki, Department of Computer Science

<http://www.seco.hut.fi/>

Introduction



- Semantically rich, ontology-based metadata is needed in developing the semantic web
- Annotation of web resources should be made possible for the people unfamiliar with the technical side of the Semantic Web
- Annotation systems currently available lack the support of distributed creation of semantic metadata based on rich metadata annotation schemas (ontologies)
- Browser-based annotation tool Saha supports collaborative annotation of web-documents and provides a simple user interface that hides the complexity of ontologies from the annotators
- Saha accepts any annotation schema defined in RDF/OWL and supports the use of external ontology services

Requirements for the annotation tool



- To be easy to use and to support the creation of semantically rich metadata, the system should...
 - hide technical concepts, e.g. URIs and complex class hierarchies, related to ontologies
 - support the use of different kinds of annotation schemas that can be used to define the structure of annotations and thus guide the annotators in their task
 - support collaborative annotation, where annotations can be produced by different annotators at different locations
 - be implemented as a web application, which will make it platform independent from the annotator's point of view

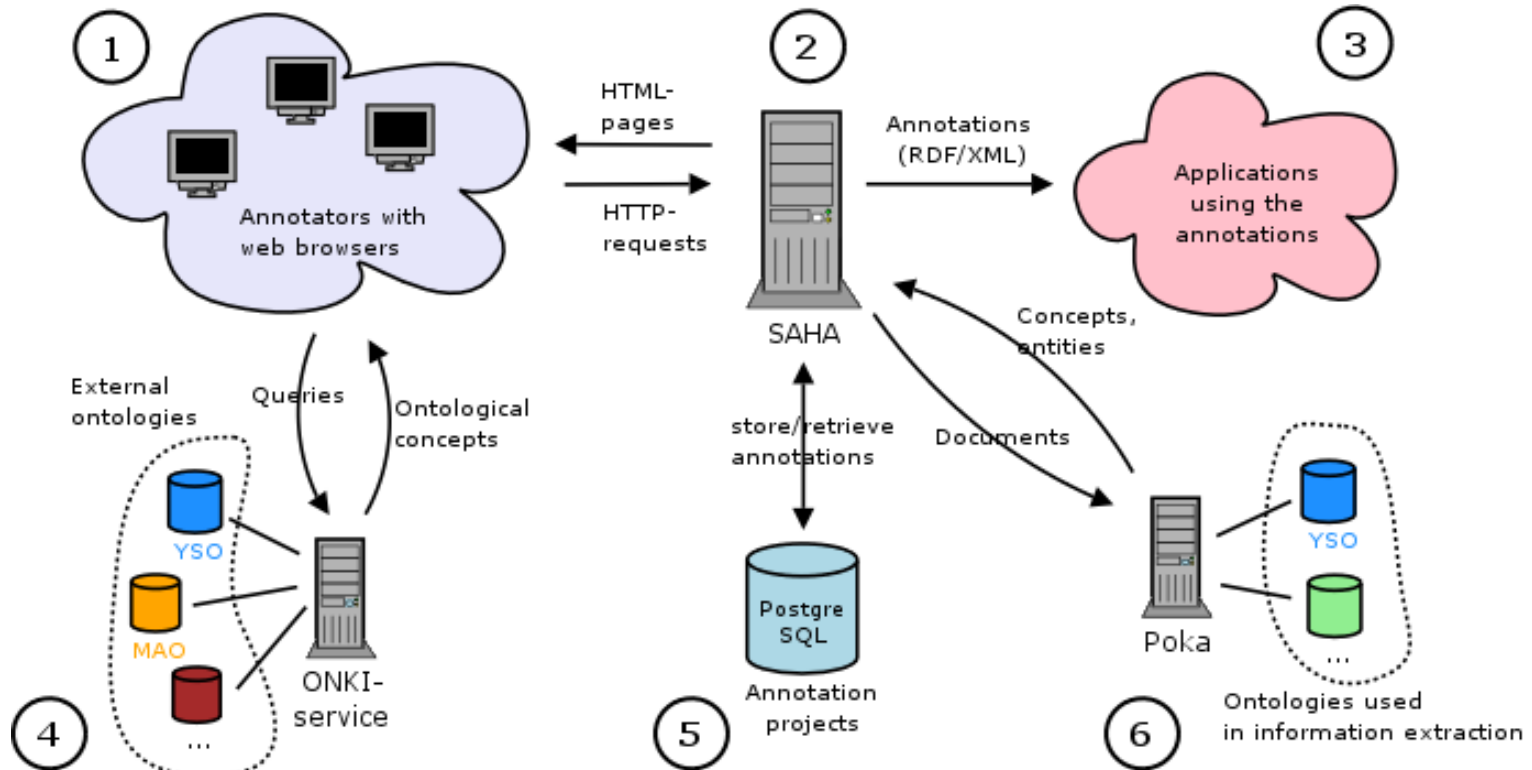
Architecture of Saha



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Saha is a web application built on Apache Cocoon and Jena frameworks
- Used with an ordinary web browser with Internet connection
- Uses extensively techniques such as Javascript and Ajax to provide an annotator with versatile and simple user interface

Architecture of Saha





Saha - Semantic Computing Research Group - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://demo.seco.tkk.fi/saha/

Saha Annotation Editor
demo

Search: | View RDF | About Saha | Language: English

Version 3.0 (30.3.2007)

Annotation classes

- Document
- Image
- Sound
- Text document
- Video
- Web page
- Semanttinen Web (Wikipedia)

Reference classes

- Document format
- Organization
- Association
- Person

Web page

Identifier: http://www.seco.tkk.fi/applications/saha#Instance_ID1177052924195
Created: 20.04.2007 10:08
Modified: 20.04.2007 10:08

Title

fi [Semanttinen Web \(Wikipedia\)](#)

en

dc:subject (Subject defined in Finnish Upper Ontology (YSO))

- teknologia (<http://yso.fi/YSO#teknologia>)
-
- yliopistot (<http://yso.fi/YSO#yliopistot>)
- suoja (<http://yso.fi/YSO#suoja>)
- teknologia (<http://yso.fi/YSO#teknologia>)
- asiakirjat (<http://yso.fi/YSO#asiakirjat>)

Document being annotated

http://fi.wikipedia.org/wiki/Semanttinen_web

Kirjautu sisään tai luo tunnus

artikkeli keskustelu muokkaa historia

Semanttinen Web
Uudelleenohjattu sivulta [Semanttinen web](#)

Semanttinen Web (engl. [Semantic Web](#)) on Internetin WWW-palvelun laajennos, jonka dokumentit on suunniteltu myös koneita silmällä pitäen. Tällöin ohjelmistot pystyvät ymmärtämään dokumenttia ja arvioimaan sen hyödyllisyyttä käyttäjän näkökulmasta. Näin käyttäjä vapautuu tiedon etsimisestä. Semanttinen verkko voisi osata esimerkiksi,

Done

Annotations in Saha (1)



- Annotations are based on user-defined metadata annotation schema, defined in RDF/OWL
- Any kind of schema can be used
- Schemas are used in Saha for defining...
 - the structure of annotations
 - types of values that can be used in annotations
 - the layout of the user interface
- Annotations are instantiated classes of an annotation schema, with properties and values defined for instances
- Annotations are linked to the documents either by stating that the document is an instance of the annotation class or using a named property (cf. property “annotates” in Annotea)

Annotations in Saha (2)



- Annotations describe documents as a whole, i.e. they are not associated with any particular section inside the documents
- Annotations can be used...
 - to find out how the document is related to other documents that have been annotated
 - to serve as a semantically described index that can be used to categorize documents or to search a document
- Classes of an annotation schema used in Saha can be distinguished by their intended use:
 - Classes that describe the documents
 - Classes that describe other resources (e.g. places, people etc.) that are somehow *related* to the documents
- Classes used to describe other resources form a local *knowledge base* that can be used and extended when new annotations are created

Annotation projects



- Collaborative annotation is supported in Saha by *annotation projects*
- Each annotation schema loaded to Saha forms an annotation project that can have multiple members as annotators
- In practice, a project is Jena's ontology model stored in a database
- Members of an annotation project can edit all the annotations and KB instances stored in the project as well as create new ones
- KB instances can be shared between different project using Onki ontology service framework



- In addition to describing annotation's structure in annotation schema, we need a way to describe *how* the schema is used during the annotation process
- With Saha, a separate *meta-schema* can be used to define e.g.
 - classes which are shown in the class hierarchy of the user interface and which user can choose when new annotation is being created
 - properties of a class whose values are used as `rdfs:label` of the instance of the class
 - the order of a class's properties, in which they are shown to the annotator
- Meta-schema improves the usability of the system and enables the use of same annotation schema in different ways in different annotation projects

Using Saha: KB instances



- Object properties of an annotation schema may have two types of values: KB instances or concepts of an external ontology
- Annotator can choose existing KB instance or create a new:

Search for existing instances:

Creator (Person responsible for creating the resource)

tom

Property

Values matching the search

Fetch all values

Define a new value (Person)

Close

Creating a new instance:

Pop-up window

Fields to define values for the instance

Specify a new value: Person

Name

Tom Whitehead en X

en +

Email address

twhite@cc.hut.fi en +

Close

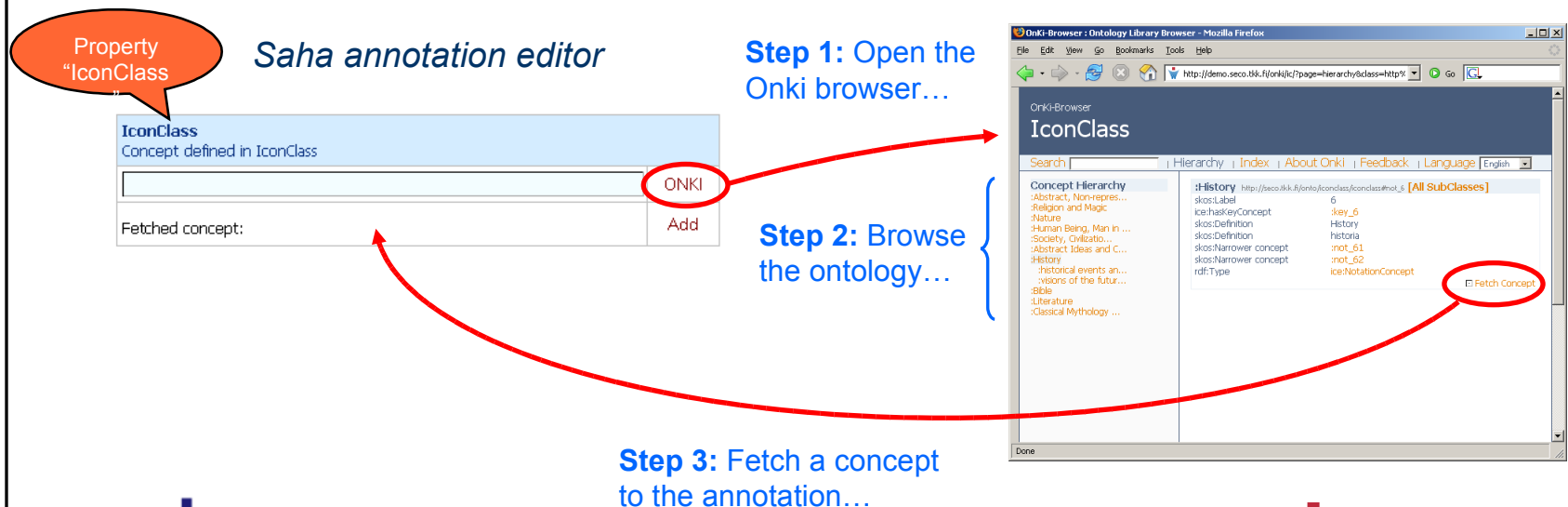
Done

Search

Using Saha: ontology services



- Saha can be connected to the Onki ontology service framework developed in the National Semantic Web Ontology Project in Finland
- Onki provides a web-based browser to search concepts defined in various ontologies and to import them to external applications, such as annotation editors



Saha: use-cases and future plans



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Saha is a working prototype
- Currently in trial use in the metadata creation for the Finnish health promotion portal
- Initial feedback from the end-users has been promising, but further experimenting is still needed
- Future plans include...
 - using Saha to provide metadata for additional semantic portal projects
 - implementing automated features to ease the annotation
 - making improvements to the user interface
- A working demo will be available at: <http://www.seco.tkk.fi/applications/saha/>



Questions?