

Knoodl.com

Semantic Wiki

Creating and using OWL vocabularies in a wiki

Michael Lang

Revelytix

June 18, 2007

Agenda

- ▶ What is a Semantic Wiki
- ▶ Building the Semantic Models
- ▶ Bootstrapping COI based vocabularies
 - With WordNet context and description
- ▶ COI vocabularies in a semantic Wiki
 - OWL models
- ▶ Semantic Wiki

Wiki

- ▶ A website where anyone can edit the content of the site easily
- ▶ Wiki's are now established as mainstream technology for collaboration
 - On the world wide web
 - Within the enterprise
- ▶ Also managing a lot of content
 - Many kinds of files can be linked to or embedded into the wiki

Wiki Drawbacks

- ▶ Information is organized in a manner similar to a file system
 - It can be very difficult to find documents on a wiki after the wiki reaches a certain size
 - Just like the file system on your personal computer
 - ▶ Except: you organized everything on your laptop
 - ▶ Everyone else organized content on the wiki

Wiki Drawbacks

- ▶ Even though wikis are collaboration and content management systems
 - There is no information model that can be used to manage the content
 - Wikis contain structured, unstructured and others sorts of content

Semantic Wiki 1.0

- ▶ A wiki based tool for building formal semantics
 - Community based, collaborative
 - Both structured and unstructured content is managed in the same collaborative framework
 - Imports and exports OWL
 - Accessible by non domain experts

Semantic Wiki 2.0

- ▶ A wiki that enables any content, structured and unstructured, to be “tagged” so the content can be queried and reasoned over
 - Tagged means adding content to an OWL based ontology
- ▶ An integrated and queryable knowledgebase
 - Query is very different from search
 - Queries can be embedded into the wikitext

Wiki Vocabularies



Wiki Vocabularies

- ▶ The semantics for any domain are created within the wiki as an OWL vocabulary
 - Project management, event management, social networks, logistics, acquisition, bioinformatics, CRM
- ▶ Multiple domain models might be available concurrently
- ▶ Once the vocabulary is created and published the semantics can be leveraged to achieve
 - Interoperability
 - Integration
 - Discovery
 - Semantic matching
 - Semantic Wikis

Bootstrapping COI Vocabularies in a Semantic Wiki

Bootstrapping Ontologies

▶ Step 1: Start at the bottom

- Build vocabularies from existing physical systems

▶ Step 2: Collaborate

- The community can document, review, discuss and change
- Human-readable documentation *and* formal ontology definition

▶ Step 3: Share and Use

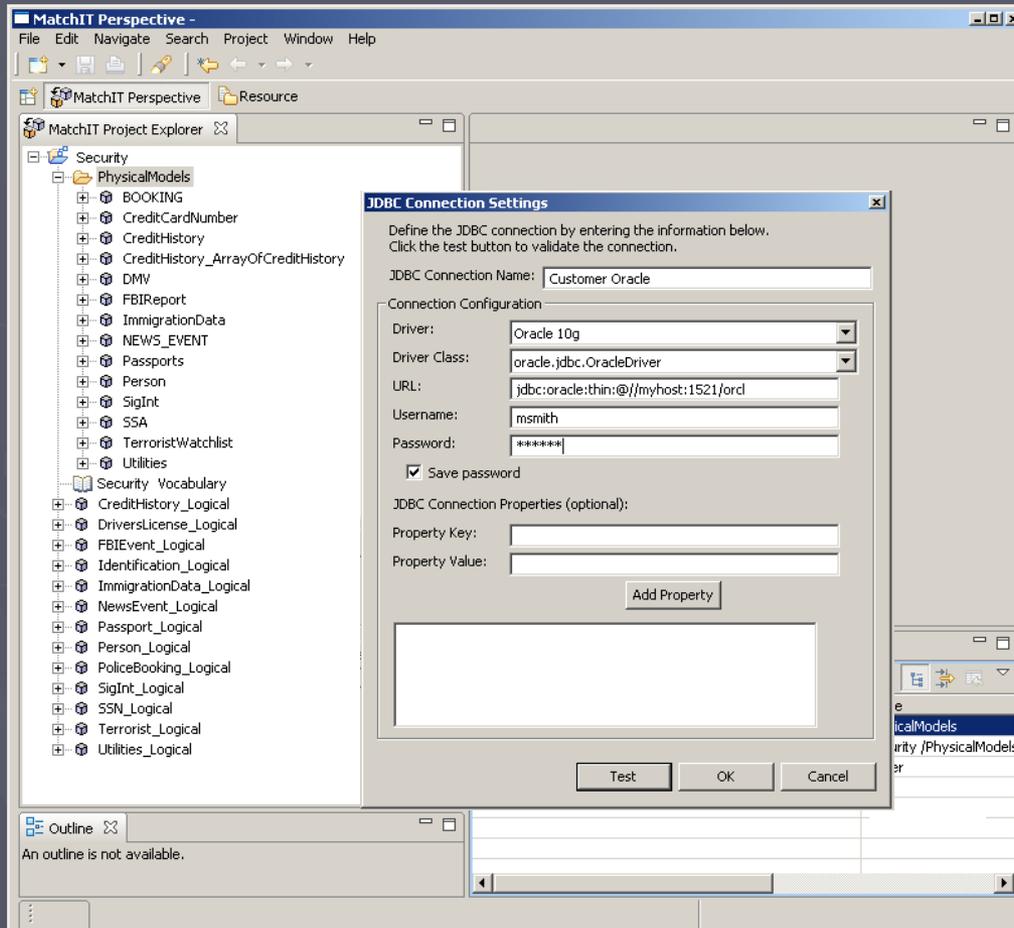
- People access the vocabularies through web browsers to view the natural language documentation and navigate formal relationships
- Machines can download OWL ontologies and use for automated reasoning

Step 1: Start at the Bottom

- ▶ Bootstrap from existing systems and models
 - Import the schemas from databases to start building the terms in the vocabulary
 - Messages, Excel, metadata repositories
- ▶ Use a semantically enabled matching tool to associate semantics with the bootstrapped terms
 - Combine the terms used with knowledge bases to discover and assign semantics to information
 - Store the terms, definitions and semantics in vocabularies
 - Built-in knowledge base is WordNet, but can also use custom domain-specific

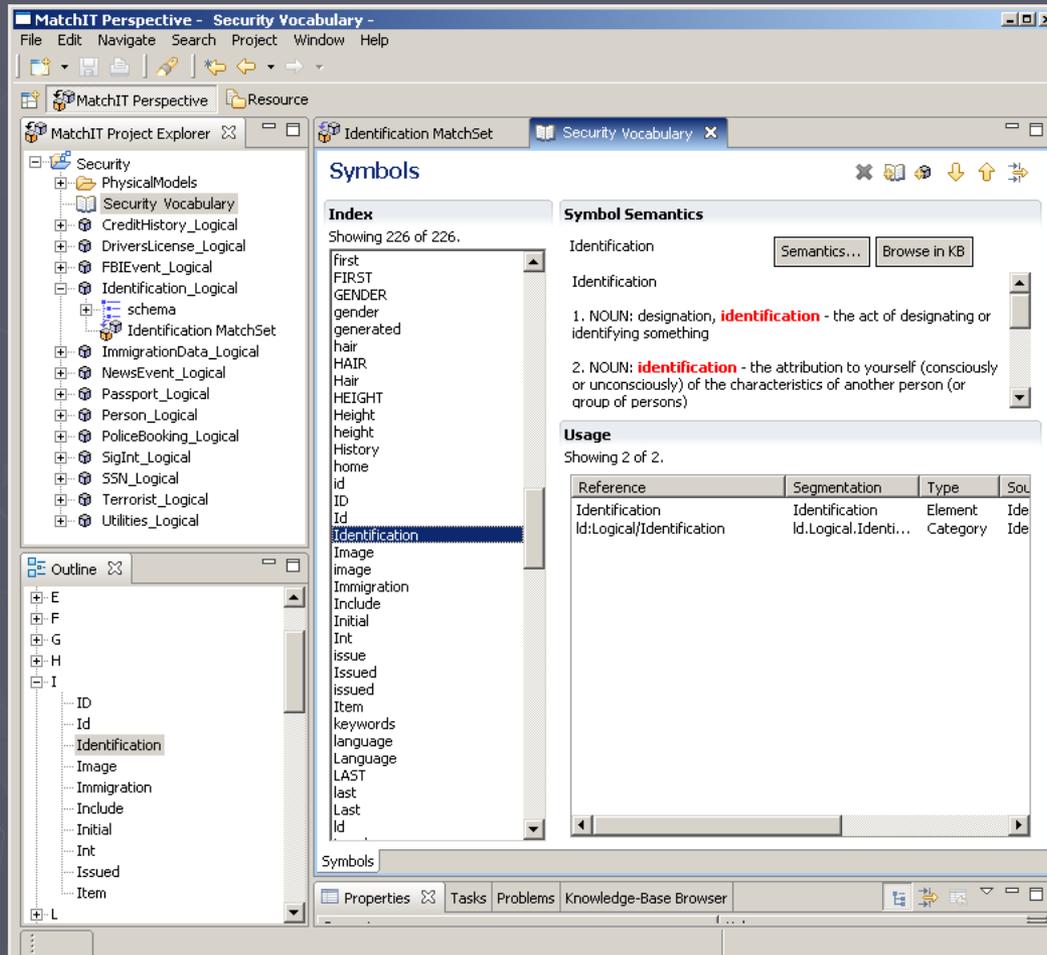
Vocabulary Management

Step 1: Extract semantics from existing data



Vocabulary Management

Step 2: Create bootstrapped vocabulary



The screenshot shows the MatchIT Perspective software interface. The main window is titled "MatchIT Perspective - Security Vocabulary". The left pane shows a project explorer with a tree view of the "Security" project, including folders like "PhysicalModels", "Security Vocabulary", and various logical models. The right pane is titled "Symbols" and displays a list of terms in the "Index" section, with "Identification" selected. The "Symbol Semantics" section shows two definitions for "Identification". The "Usage" section shows a table with columns "Reference", "Segmentation", "Type", and "Source".

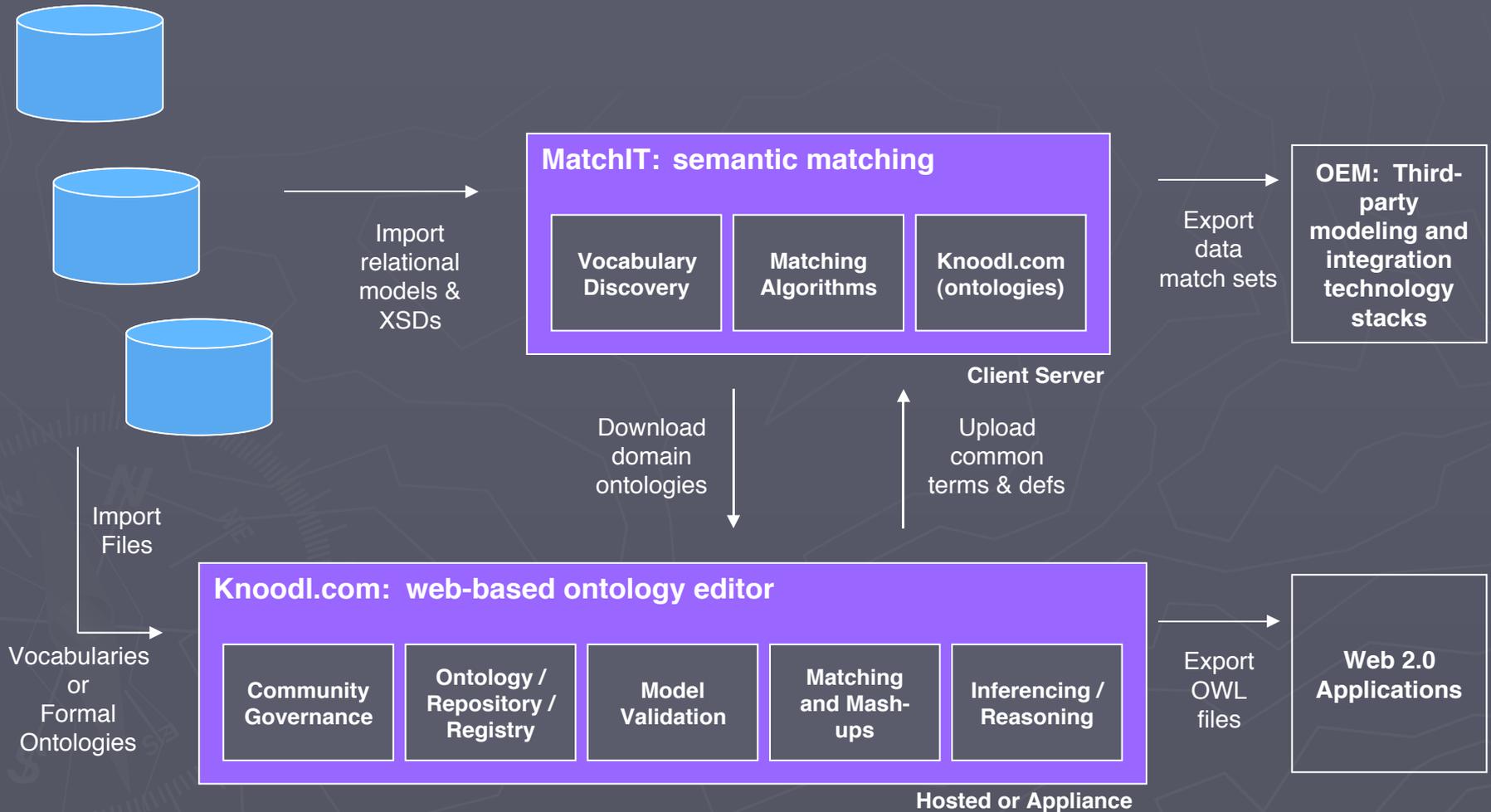
Reference	Segmentation	Type	Source
Identification	Identification	Element	Id
Id:Logical/Identification	Id.Logical.Identi...	Category	Id



Step 2: Collaborate

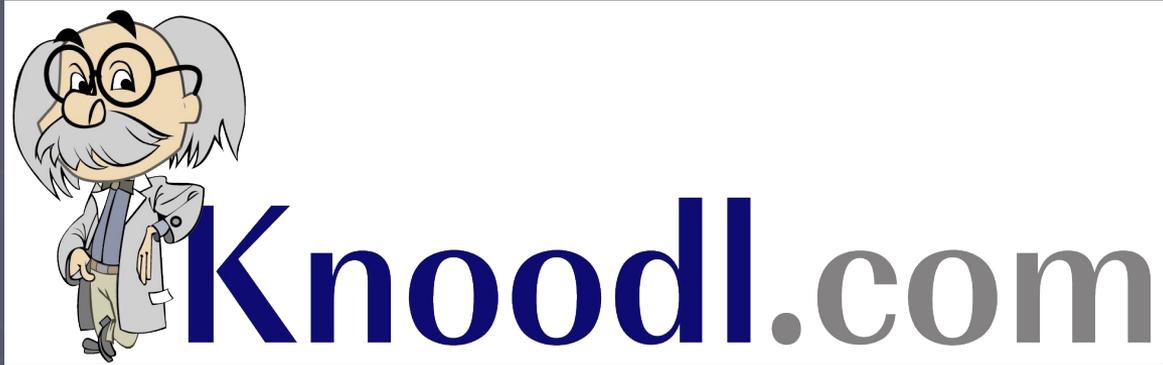
- ▶ Creating vocabularies is naturally collaborative
 - identify, define, document, standardize, edit, review, audit
 - Involve the right people
 - Reuse other vocabularies: benefit from the experts
- ▶ Community-oriented
 - A community consists of members that share experience, expertise and interest in a particular domain
 - Communities manage memberships, content, and access privileges
- ▶ Semantic Wiki
 - Captures the efforts of many over time
 - Adds semantic richness to wiki markup language

Vocabulary Development



Step 3: Share and Use

- ▶ Machines use ontologies
 - The vocabularies are represented with formalism that are rich and precise enough for software
 - Vocabularies can be downloaded as OWL ontologies
- ▶ People use natural language
 - (Most) People don't understand XML, OWL, RDF, or even HTML
 - People understand text, images, tables, charts, links
 - Follow existing web paradigms that people are comfortable with (browsers, links, pages, addresses, search, discussions, etc.)
- ▶ Keep the two parts together
 - People have to understand the vocabulary to maintain and use it
 - If parts are kept separate, more difficult to diverge
 - It's simply easier this way! (Manually aligning documentation with models is too much work)



Semantic Wiki

Knoodl.com

- ▶ Uses the Wiki paradigm to enable the development and use of OWL vocabularies by Communities of Interest (COIs)
 - W3C-based OWL editor, registry/repository
 - Facilitate sharing

Knoodl.com is ...

- ▶ An internet application where people can collaborate with others in their communities of interest to
 - Create, edit, share and find
 - Vocabularies / ontologies
- ▶ OWL Repository
 - Free, but licensing controlled by COI's
- ▶ Institutional Knowledge Management
 - Users contribute content and benefit from the content
 - Vocabularies capture much of the institutional knowledge of an enterprise or community
 - Gain value over time

Knoodl.com

- ▶ Knoodl is a collaborative framework
- ▶ We need three groups of stakeholders contributing to the description and context of the domain
 - ▶ Businesspeople
 - ▶ Technical people
 - ▶ Data people
- Knoodl provides the features for the business people to participate

Vocabulary Management

Use vocabulary to understand

Knoodl.com

Home
Communities
Directory List
Create New
Library
Directory List
Search
My Account
My Communities
My Bookshelf
My Preferences
Help
FAQ
News and Announcements

FOAF
Community examples vocabulary FOAF version 2.0
View Discussion Edit History

Note: The content was copied from a portion of that found at <http://xmllns.com/foaf/0.1/>

Contents

- 1 FOAF Basics
 - 1.1 Example
- 2 The Basic Idea
- 3 Technical Specifications
 - 3.1 Status
 - 3.2 Overview
 - 3.3 Classes
 - 3.4 Properties

FOAF Basics

The FOAF project is based around the use of machine readable Web homepages for people, grr collection of basic terms that can be used in these Web pages. At the heart of the FOAF project is about the world. The initial focus of FOAF has been on the description of people, since people are documents, attend meetings, are depicted in photos, and so on.

The FOAF Vocabulary definitions presented here are written using a computer language (RDF/OWL vocabulary, and consequently about the things described in FOAF documents. A FOAF document, database of information.

Example

Here is a very basic document describing a person:

```
<foaf:Person>  
<foaf:name>Dan Brickley</foaf:name>  
<foaf:mbox_sha1sum>241021f0e6289f92815fc210f9e9137262c252e</foaf:mbox>  
<foaf:homepage rdf:resource="http://rdfweb.org/people/danbr1/" />  
<foaf:img rdf:resource="http://rdfweb.org/people/danbr1/mugshot/danbr1</foaf:Person>
```

This brief example introduces the basics of FOAF. It basically says, 'there is a foaf:Person with a f241021f0e6289f92815fc210f9e9137262c252e; this person stands in a foaf:homepage relati http://rdfweb.org/people/danbr1mugshot/danbr1-small.jpeg.

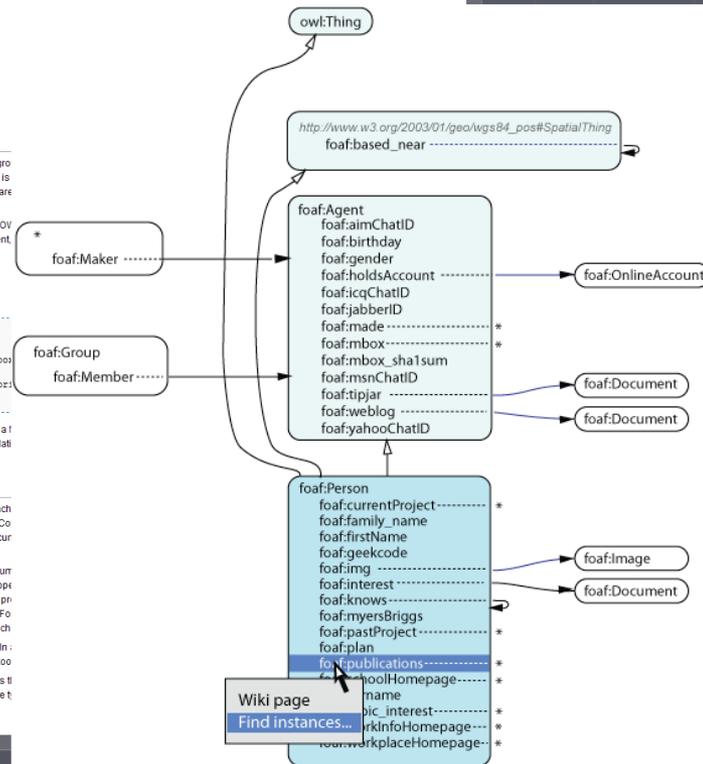
The Basic Idea

The basic idea is pretty simple. If people publish information in the FOAF document format, much such documents in the Web, we will have a machine-friendly version of today's hypertext Web. Co than humans, storing the information they find, keeping a list of "see also" pointers to other docur question-answering services based on the harvested documents.

So, what is the FOAF document format? FOAF files are just text documents (well, Unicode docum Framework (RDF). In addition, the FOAF vocabulary defines some useful constructs that can appl categories ("classes") such as foaf:Person, foaf:Document, foaf:Image, alongside some handy pri as well as some useful kinds of relationship that hold between members of these categories. Fo foaf:Person) to a foaf:Image. The FOAF demos that feature photos and listings of who is in which

The specific contents of the FOAF vocabulary are detailed in this FOAF namespace document. In contain "see also" pointers to other FOAF files. This provides a basis for automatic harvesting too

The remainder of this specification describes how to publish and interpret descriptions such as th number of categories (RDF classes such as "Person") and properties (relationship and attribute t machine-readable form, hyperlinked for quick reference.



Semantic Wiki

- ▶ Incorporate formal semantic technology into the preeminent collaboration technology
 - Features that facilitate the construction of formal semantic models
 - Features that make it simple and even automated to some extent to “link” facts to one or more models
 - ▶ Structured
 - ▶ Unstructured

Semantic Wiki

- ▶ A collaborative tool to build integrated knowledge bases
 - Formal queryable information stores
 - Context of the queries is determined by the structure of the ontology
 - ▶ Axioms
 - ▶ Relationships
 - ▶ Assertions
 - Facts are captured as RDF
 - Integrates structured and unstructured content

Semantic Tagging

- ▶ Browser plug in facilitates tagging any content while browsing any website to any collection of semantic models
- ▶ Form based fact creation
 - Form is generated from the ontology
- ▶ Markup text or data
 - Highlight text or data and dynamically add the markup to the knowledgebase as facts
- ▶ Semantic models are selected from the registry at knoodl.com

Thank You

