

VocBench

Caterina Caracciolo, FAO

Armando Stellato, Uni Tor Vergata



Food and Agriculture Organization
of the United Nations



VOCBENCH 2.0

A Collaborative Environment for Management of
SKOS/SKOS-XL Concept Schemes

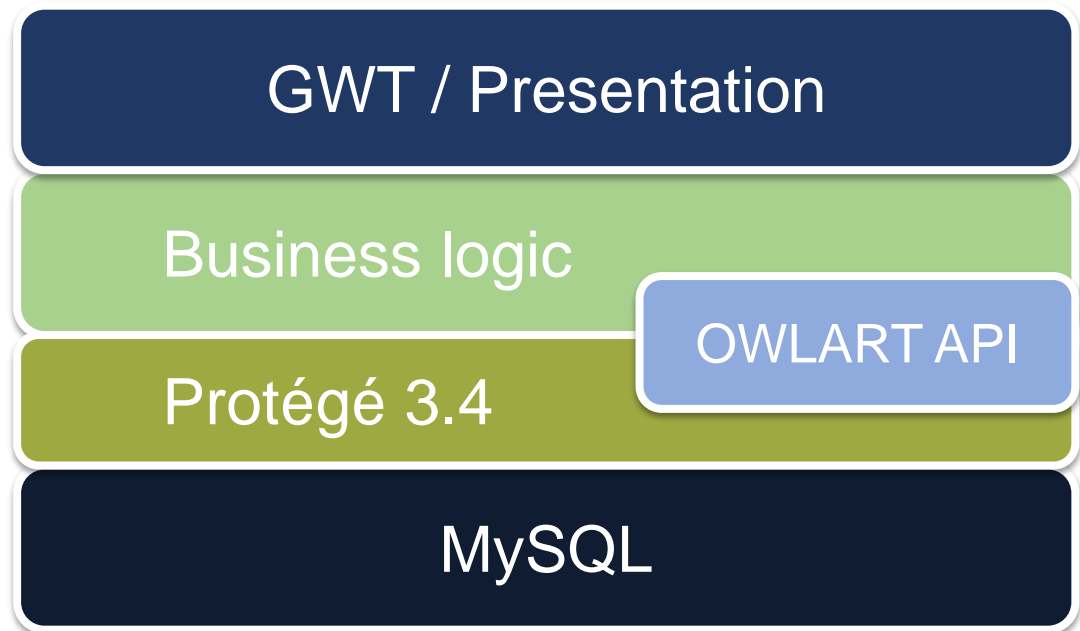


Why was it built?

AGROVOC (big agriculture vocabulary developed by FAO)

- In 2004: >32 000 concepts in up to 22 languages
- A global group of terminologists.
- No existing standard for thesauri
- No existing tool that met FAO's needs

- Google Web Toolkit
- Lucene
- Protégé API
- OWLART API
- MySQL
- Custom OWL model





V1 Problems

- Couldn't support other triple stores (Glued to Protégé API)
- Custom OWL model
- No support for emerging standards, e.g. SKOS
- No import
- No support for alignments
 - AGROVOC aligned to a dozen other vocabularies
- Complicated export
- No SPARQL endpoint



Towards V2

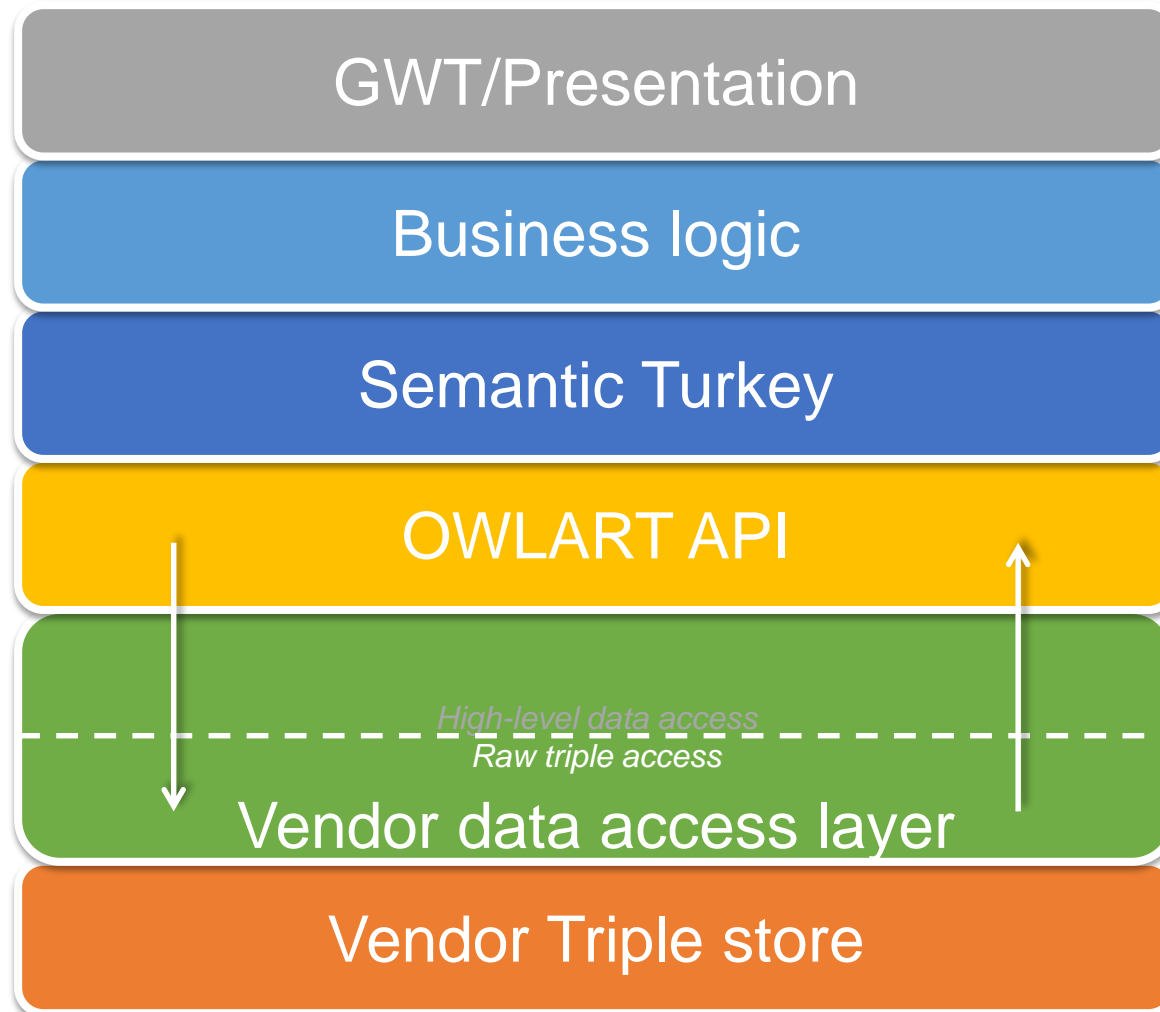
- Market Analysis
 - Choose a commercial product or
 - Build a new version
- Collaboration
 - U. of Rome Tor Vergata
 - MIMOS Berhad (FAO Center of Excellence)
 - EC FP7 projects agInfra and SemaGrow
- The Decision
 - Build a new version leveraging new frameworks and an OSGi pluggable SOA design
 - Support SKOS (& SKOS-XL) and importing

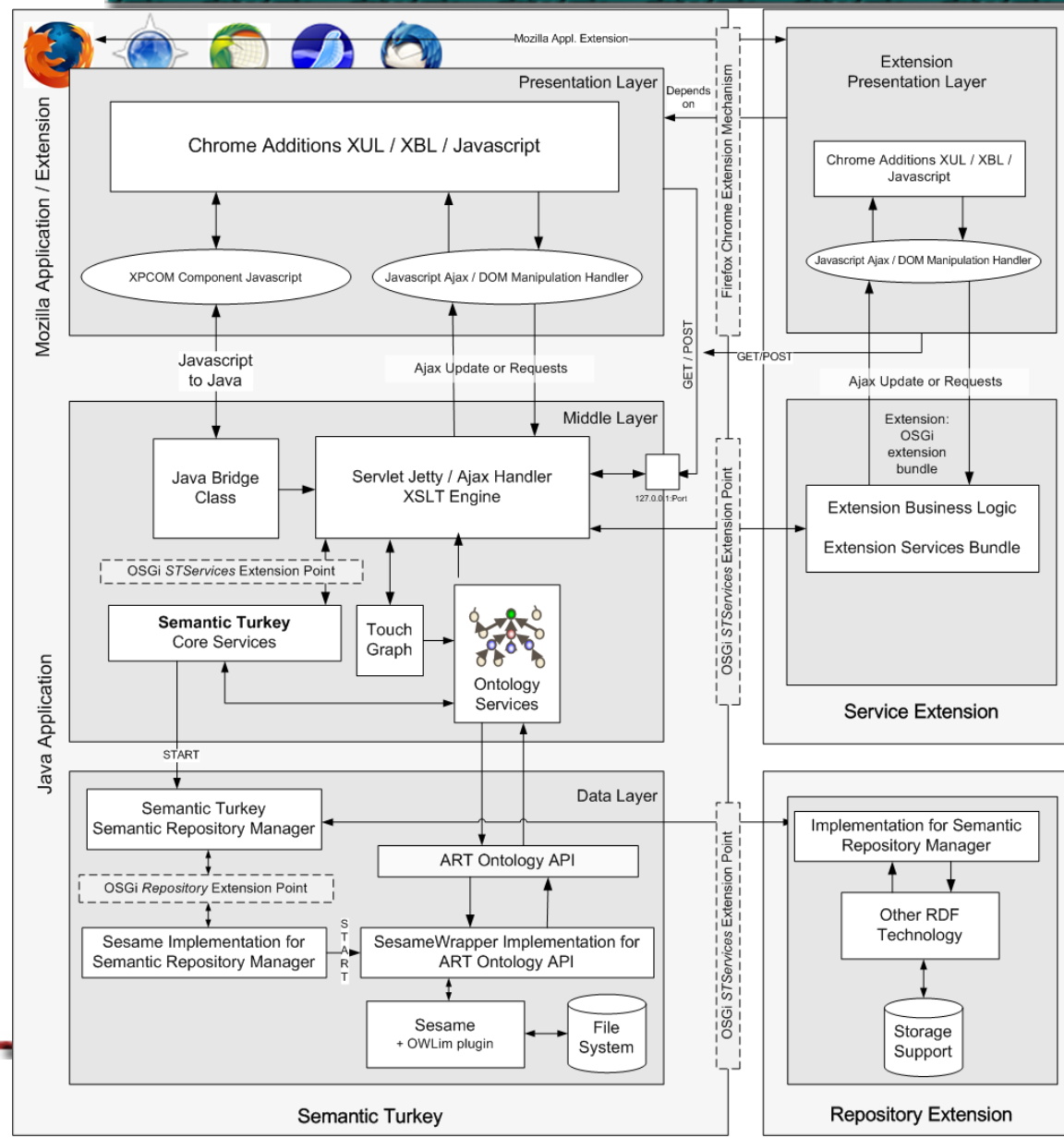


Objectives for VB2.0...

- A completely rebuilt backing framework for the service and data layers, based on an already existing open source project: Semantic Turkey¹
 - Based on OSGi Open Services Gateway
 - Open Connectivity to most notable RDF middleware and triple storing technologies (Sesame2, Jena, Allegrograph...)
 - Native support for SKOS and SKOSXL over RDF (no more conversions from internal legacy models), other than OWL
- VB1.0 User Interface remains mostly unchanged in the first release of VB2.0

1. <http://semanticturkey.uniroma2.it/>





Three layered extensible architecture

Presentation Layer

- GWT (Google Web Toolkit)
- Vocbench User Interface (*Mozilla apps in the original framework*)

Services Layer

- Enables communication between the client (Vocbench UI) and the ontology persistence layer.
- HTTP based Services accessed through the Ajax paradigm
- OSGi Extensible Servicing System

Persistence Layer

- Access to ontological knowledge.
- Based on dedicated ontology API, which can be implemented through use of different technologies.

Collaborative Management

- Validation&Publication Workflow (propose, validate, publish, revise, deprecate...)
- Fine grained user management
 - both users and functionalities may be associated in groups
 - Functionalities (or groups of) may be assigned to different users (or groups of)
- Full editing history (not only concepts, but most of the actions can be subject to validation too)
- RSS Feeds
- Fine-grained metadata and editorial notes: SKOS-XL and reified definitions allow for timestamped status and rich editorial notes

Multilinguality

- Strong support for multi-lingual thesauri management
- Application itself is also multilingual (currently support for english, dutch, spanish, more languages coming)

Native RDF support

- Support for different triple stores
- Possibility to SPARQL query/update through a dedicated interface with syntax completion/highlight
- SKOS-XL management
 - If preferred, SKOS-core export through available conversion tools

Large scale thesauri management

- Scalability limited only by the underlying triple store

Extensibility

- OSGi connectable services

Advanced skos:ConceptScheme Management

- SKOS allows for non-trivial management of multiple conceptual schemes, which is fully supported by VB

And, last but not the least: **Free and Open Source!** (<http://vocbench.uniroma2.it>)

- **Client:** GWT¹ / Mozilla² technologies

- GWT Java / Mozilla JavaScript client library
- Java/JavaScript Client API for services
- Ontology Widgets (*under development*)

- **Services:** OSGi

- Service Extensions
- Plugins

- **Data and Models:** OSGi

- Ontology Manager implementations according to different triple stores

```

import("resource://stservices/SERVICE_Cls.jsm", art_semanticturkey);
try {
    var responseXML;
    //this method
    art_semanticturkey
} catch (e) {
    //simple handling
}

```

Overview Package Class Use Tree Deprecated Index Help

OWL Art API 2.0.1 API

Package	Class	Use	Tree	Deprecated	Index	Help
it.uniroma2.art.owlart.exceptions						
it.uniroma2.art.owlart.filter						
it.uniroma2.art.owlart.io						
...						
it.uniroma2.art.owlart.query.io						
...						
it.uniroma2.art.owlart.util						

OntologyManager <<interface>>

Sesame2 OntManager <<Implementation>>

Jena OntManager <<Implementation>>

... OntManager <<Implementation>>

Sesame Repository

Jena Graph

...

1. Google Web Toolkit Java Client on the VOCBENCH Collaborative Suite
2. Mozilla Javascript Client on Semantic Turkey based desktop tools

Current service implementation

```
if (request.equals(Req.isTopConceptRequest)) {  
    String skosConceptName = setHttpPar(Par.concept);  
    String schemeName = setHttpPar(Par.scheme);  
    checkRequestParametersAllNotNull(Par.concept, Par.scheme);  
    response = isTopConcept(skosConceptName, schemeName);  
}
```

```
public Response isTopConcept(String skosConceptName, String schemeName) {  
    SKOSModel skosModel = getSKOSModel();
```

```
    try {  
        ARTResource[] graphs = getUserNamedGraphs();  
        ARTURIResource skosConcept = retrieveExistingResource(skosModel, skosConceptName, graphs);  
        ARTURIResource skosScheme = retrieveExistingResource(skosModel, schemeName, graphs);  
        return createBooleanResponse(skosModel.isTopConcept(skosConcept, skosScheme, graphs));  
    } catch (NonExistingRDFResourceException e) {  
        return logAndSendException(e);  
    } catch (ModelAccessException e) {  
        return logAndSendException(e);  
    }  
}
```

Annotations may contain explicit validation checks with respect to the application's semantics

Arguments and returned values are now explicitly managed with their native types. Marshalling/Unmarshalling to the serialization formats adopted by the service is demanded to dedicated components

Separation of service method-controller / automatic generation of controller

```
public boolean isTopConcept(@present ARTURIResource skosConcept, @Optional(defaultValue = ":mainScheme") ARTURIResource skosScheme)  
    throws NonExistingRDFResourceException, ModelAccessException {  
    SKOSModel skosModel = getSKOSModel();  
    ARTResource[] graphs = getUserNamedGraphs();  
    return skosModel.isTopConcept(skosConcept, skosScheme, graphs);  
}
```

The method signature then drives the generation of the controller, which is the direct frontend for the service. Exceptions are serialized in the response (the content of which codes both data and application-level error codes) and data validation annotations are managed by Spring data validation methods



V2: Triple Store Support

- What triple stores do we currently support?
 - Sesame2 (standard internal triple stores, both in-memory and native) and
 - OWLIM (through sesame2 remote connection), plus
 - ANY repository which can be accessed through a plain sesame2 remote connection.
- For OWLIM, we make use of the specific triple indexing it offers
- Non-updated interfaces for Allegrograph and Jena
- Support for Virtuoso coming from a VocBench partner

Signed in as Administrator (Administrator) to: AGROVOC_SKOS_VB_2013-07-12_SIMPLIFIED

Administration | About VocBench | RSS feed | Preferences | Help | Sign out

VocBench

VERSION 2.0.1-SNAPSHOT [Build 20140213] (DEVELOPMENT)

Exact word Go [Advanced search](#)

Recent changes **Concepts** Properties Schemes Validation Load data Export Statistics SPARQL [Concept navigation history](#) [Content language](#)

Concepts Show URI Show non-preferred

URI: http://aims.fao.org/aos/agrovoc/c_330892

- activities (en); activité (fr); Attività (it)**
- entities (en); entité (fr)**
 - Accounts (en); Compte (fr); Conti (it)
 - Acts (en); Decrees (en); Laws (en); **legislation (en)**; Décret (fr); Loi (fr); **Législation (fr)**; Atti legislativi (it); Decreti (it); Leggi (it); **Legislazione (it)**
 - agencies (en); Agenzie (it)
 - Agreements (en); accord (fr); Accordi (it)
 - Assets (en); Actif (fr); Attivo patrimoniale (it)
 - balance of payments (en); Balance des paiements (fr); Bilancia dei pagamenti (it)
 - Bank deposits (en); Dépôt bancaire (fr); Depositi bancari (it)
 - Biographies (en); Biographie (fr); Biografie (it)
 - boards (organizations) (en); conseil (fr)
 - Boundaries (en); Territorial boundaries (en); **Délimitation (fr)**; Confini (it); Confini territoriali (it)
 - branches (enterprises) (en)
 - Brand names (en); Brands (en); Proprietary names (en); **Trade marks (en)**; Trade names (en); **Marque commerciale (fr)**; Marque déposée (fr); Nom commercial (fr); Denominazioni commerciali (it); Marche (it); Marche commerciali (it); Marchi depositati (it); **Marchi di fabbrica (it)**
 - brigades (en); **work teams (en)**; Brigade (fr); **Équipe de travail (fr)**; Gruppi di lavoro (it); Gruppi organizzati (it)

entities (en); entité (fr) Show inferred and explicit

Terms (11) Definition (1) Note (0) Attribute (0) Notation (0) Relationship (0) History (0) Image (0) Hierarchy

+ Add new term

Language	Term
English (en)	<input checked="" type="checkbox"/> entities (Preferred) W
Español (es)	<input checked="" type="checkbox"/> entidades (Preferred) W
Français (fr)	<input checked="" type="checkbox"/> entité (Preferred) W
中文 (zh)	<input checked="" type="checkbox"/> 实体 (Preferred) W
Cesky (cs)	<input checked="" type="checkbox"/> celky (Preferred) W
Deutsch (de)	<input checked="" type="checkbox"/> Einheit (Preferred) W <input checked="" type="checkbox"/> Entität W <input checked="" type="checkbox"/> Entitaet W
Türkçe (tr)	<input checked="" type="checkbox"/> Varlık (Preferred) W
українська мова (uk)	<input checked="" type="checkbox"/> ради (організацій) (Preferred) W <input checked="" type="checkbox"/> господарсько-економічні одиниці W

Legend Proposed Validated Published Revised Proposed deprecated Depreated [Show more](#)

© FAO, 2014



VB "desktop version": Semantic Turkey

Firefox

Strange Cases from the Files of Astrono... +

people.physics.tamu.edu/krisciunas/strange/strange.html

Google

SKOS Panel (UAT)

Concepts Properties Schemes

- Interdisciplinary astronomy
 - Astrophysical fluid dynamics
 - Hydrodynamical simulations
 - Archaeoastronomy
 - Megalithic astronomy
 - Ethnoastronomy
 - Molecular astrophysics
 - Sociology of astronomy
 - Astrobiology
 - Xenobiology
 - Exobiology
 - Astroparticle physics
 - Gamma rays
 - Gamma-ray sources
 - Gamma-ray lines
 - Cosmological neutrinos
 - Urca process
 - Neutrino oscillation
 - Supernova neutrinos
 - Cosmic rays
 - Solar neutrinos
 - Gravitational waves
 - Astrochemistry
 - Philosophy of astronomy
 - Asteroseismology
 - Observational astronomy

[Next: About this document](#)

Strange Cases from the Files of Astronomical Sociology

Kevin Krisciunas

Texas A&M University
Department of Physics and Astronomy
College Station
Texas 77843

Abstract:

What astronomer could not use his own observatory worth \$5 billion in today's money? Who had one of his anti-correlation between scientific output and honeymoons? Who wrote the most egotistical opening of King Tutankhamen's tomb and observatory? What is the strangest abstract well-known bad tempered astronomer wrote in Bulgaria, which all vampirologists recognize? Who embarked on a sea voyage to London? In the

One of the rapidly growing social sciences keep wondering why we are the way we

That astronomers are considered unusual

concept Editor

Sociology of astronomy	Remove Value
skosxl:prefLabel	
xl_en_7b83288e	Remove Value
skos:semanticRelation	
T828	Remove Value
T93	Remove Value
skos:topConceptOf	
skos:inScheme	
http://purl.org/astronomy/uat	Remove Value

WebLinks:

Web documents in topic:

[Strange Cases from the Files of Astronomical Sociology](#)

Close

Armando Stellato Home Page - Mozilla Firefox

File Modifica Visualizza Cronologia Segnalibri Strumenti ?

http://ai-nlp.info.uniroma2.it/stellato/

Come iniziare Ultime notizie

Enter query...

Google Cerca

getRDFa

Ontology Panel

Class Panel Property Panel

Classes

- rtv:Person(1)
- rtv:Role
- rtv:Organization
- rtv:Project(4)
- rtv:Entity
- rtv:Standard
- rtv:Activity

Instances of rtv:Person


- Armando Stellato

Enter query...

Completato

Armando Stellato

I am a member, since 2002, of the [Artificial Intelligence Research Group](#) at the Department of Computer Science, Systems, and Production in the [University of Rome "Tor Vergata"](#). I have graduated in Computer Science Engineering in 2002 with a thesis on "Ontological Mediation in a community of Intelligent Linguistic Agents" and took my PhD in 2006 with a thesis on Alignment and Mediation of Distributed Information Sources in the Semantic Web.



Research Interests

- Knowledge Representation
- KS: Integration of
- Information Retrieval
- Information Extraction

Contact Info

e-mail: stellato@info.uniroma2.it

Publications

- Savino Sguera, Armando Stellato, Teresa Paziienza, Stefano Scandura, *Reusing software S...*
- 4th Italian Semantic Web Workshop (SWAP2007) Bari, Italy, December 18-20, 2007

Instance Annotation

"University of Rome "Tor Vergata"" is a further annotation of "Armando Stellato"

"University of Rome "Tor Vergata"" is a range instance of a property of "Armando Stellato"

Show all properties

Property

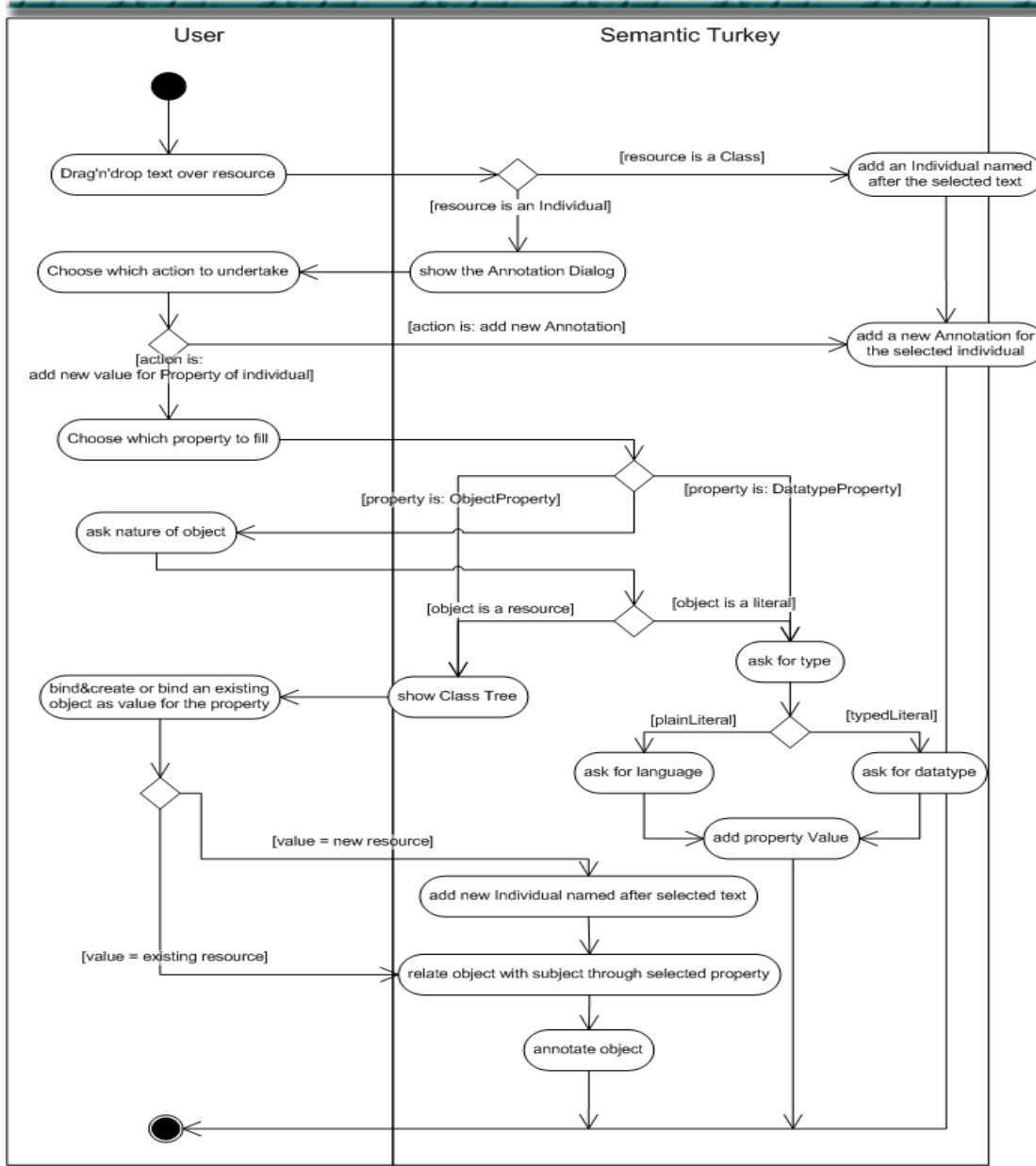
- rtv:worksIn
- rtv:occupation
- rtv:phoneNumber

ok cancel



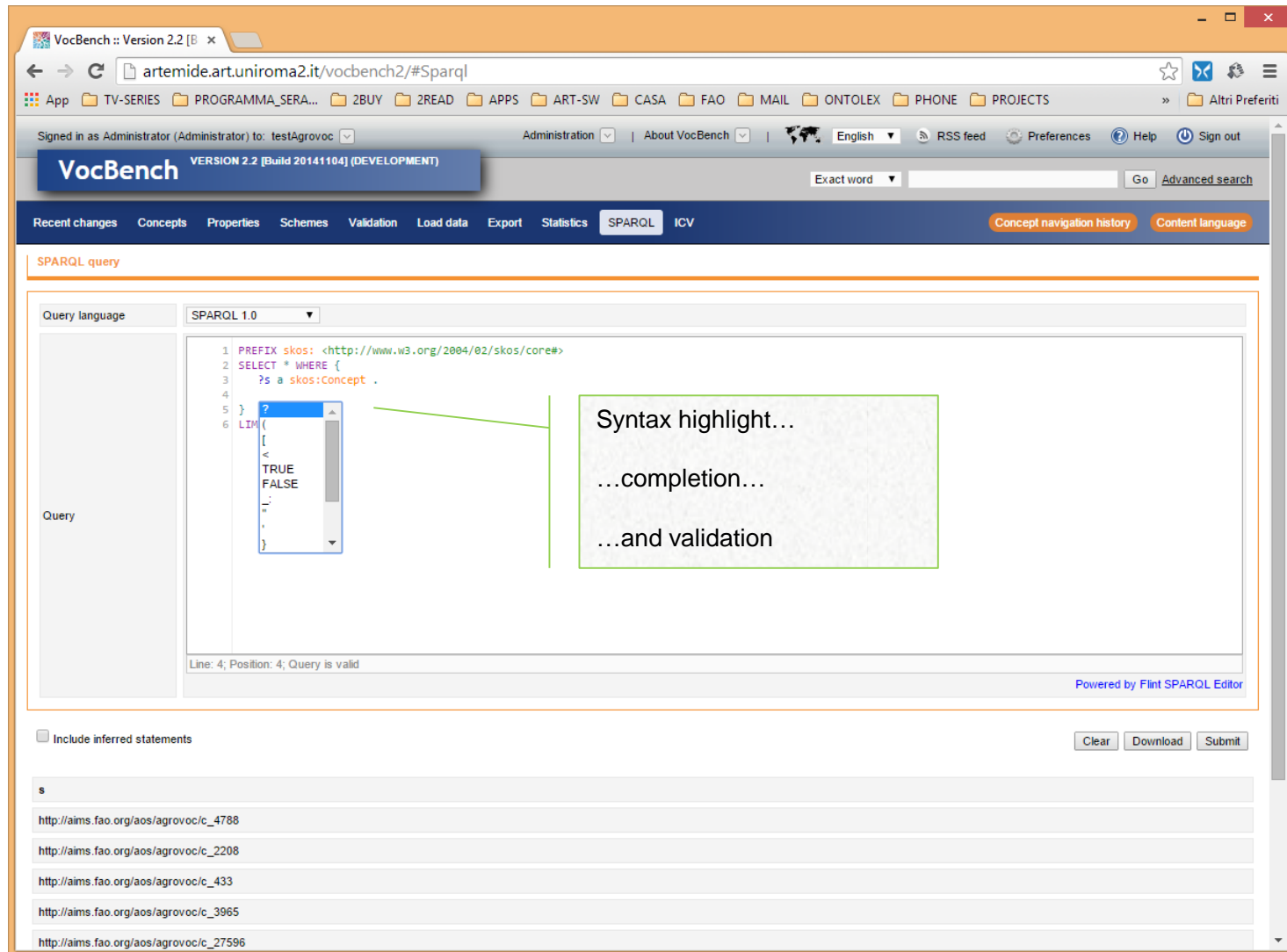
Semantic Turkey in Action: Semantic Annotation

- No automatic ontology building from text but...
- with just one intuitive drag'n'drop operation (and few HC interactions), the system:
 - Creates a new *Domain Object* instance
 - (and/or builds a new lexicalization for the already existing instance on the annotate page)
 - Creates a new *SemanticAnnotation* instance
 - Creates a new *WebPage* instance
 - Relate all of them through dedicated properties
 - ...(depending on the specific operation)



Specialized operations based on context

- Full OWL Editing
 - Individual creation
 - Web reference annotation
 - Multilingual annotations
 - Property valuing, depending on property type and range restrictions
- but...
- Knowledge Acquisition must be kept simple for the user!
- Result: context-sensitive drag'n'drop based operations



The screenshot shows the VocBench web application interface. The browser address bar displays `artemide.art.uniroma2.it/vocbench2/#Sparql`. The application header includes the VocBench logo, version information (VERSION 2.2 [Build 20141104] (DEVELOPMENT)), and navigation links. The main content area is titled "SPARQL query" and features a query editor with the following SPARQL query:

```
1 PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
2 SELECT * WHERE {
3   ?s a skos:Concept .
4 }
5 }
6 LIMIT {
```

A dropdown menu is open over the `LIMIT` keyword, showing options: `<`, `TRUE`, `FALSE`, `..`, `+`, and `}`. A green box highlights the dropdown menu with the text "Syntax highlight... ..completion... ..and validation". Below the query editor, the status bar indicates "Line: 4; Position: 4; Query is valid". The interface also includes a "Query" label on the left, a "Powered by Flint SPARQL Editor" note, and a section for "Include inferred statements" with "Clear", "Download", and "Submit" buttons. The results section shows a list of URIs:

- s
- `http://aims.fao.org/aos/agrovoc/c_4788`
- `http://aims.fao.org/aos/agrovoc/c_2208`
- `http://aims.fao.org/aos/agrovoc/c_433`
- `http://aims.fao.org/aos/agrovoc/c_3965`
- `http://aims.fao.org/aos/agrovoc/c_27596`

This is a non-exhaustive list of features added along the various versions. Only the major news are reported here

VB2.2

- important change in ST: moved to the karaf OSGi container. Just run it through the bin/st_server_run batch/bash file. The new ST is totally compatible with the same data folder (SemanticTurkeyData) of the previous Semantic Turkey , so no need to update/change the existing data.
- ST Instance Manager: no more need to pass an ST connection URL (which was rather annoying to be copied and pasted from the installation instructions). The static part of the connection URL is hardcoded in the system and is always paired (and updated) with the ST version being used by the current VB, while the dynamic part can be defined only in terms of port and host. Also, through the ST instance manger, users can define multiple ST installations in terms of the sole baseURL and port and then just invoke them when defining a new project. A default ST configuration is already stored in the system, and corresponds to the default settings of ST (localhost, 1979).
- Running more projects on the same ST instance: from ST 0.10 (paired with VB2.2), it is now possible to manage more projects in parallel with the same ST instance (technical details on the ST site: <http://semanticturkey.uniroma2.it>). Through the ST instance manager, it is possible to associate and host different projects on ST instances, though this is no more necessary.
- User experience: New features/improvements:
 - A much requested feature: user-customizable entity id-generator (e.g. the way the name of concepts, xlabels etc.. is generated can be customized)
 - Validation for move/link/unlink actions
 - downloadable SPARQL results
 - scheme management: now possible to edit lexicalizations for schemes as well
 - Interface for checking and repairing dangling concepts
- bug fix:
 - Fixed locking of visitor only mode. In visitor mode, user can login as Administrator to get out of visitor mode.
 - In ST: no more projects deleted due to a lost connection to a remote repository
 - RDF Export not working
 - Removed the user/group management capability from the Publisher. Note this is not removed from existing installation
- New experimental resource view. To test it: login as administrator, you will see new button on top of concept tree. Select any concept and click "<s-p-o>" button near the visualization button. It will popup the resource view.
- *Developers only*: Context injectable services: now services address objects injected through an extensible "context": the context currently includes the: read and write graph(s) and the project to be considered (following the multi-project management enabled in this version)

VB2.1

- A completely rebuilt installation mechanism for an headache-free installation experience!
 - Self-installing DB, with auto-updating scripts
 - Wizard-driven system configuration, with import/export of configuration profiles
- SPARQL module: query/update content directly through the SPARQL query language for RDF; syntax completion & highlight
- Multi scheme management: now concepts can be shared among different schemes
- RSS feeds for all editing actions

VB2.0

- A Completely re-engineered RDF backend, based on RDF Management platform Semantic Turkey
 - Support for different triple stores
 - Extension mechanism based on OSGi
- Multi scheme management. Several skos:ConceptSchemes can be developed for the same dataset, providing different views on the data
- Statistics module: a module providing resumung information about the loaded data.
- Export module: for exporting all or part of the content of a project according to several existing RDF serialization standards
- Load data module: for loading bulk data serialized in some RDF serialization standard
- Ontology Import Management (Administration-->Ontologies): to owl:import ontologies to be used as property vocabularies for the modeled thesauri
- New tabs under the concept view for covering extensively the SKOSXL standard (note, notations)

Firefox

W Maize weevil - Wikipedia, the...

file:///D:/ingegneria/dottorato/presentazioni/DemoST-AgroIE/Demo AgroIE-st/Mai

SKOS Panel (DemoAgroVoc)

Concepts Properties Schemes

- Product
 - Insecticide
 - Diazinon
 - Animal
 - Insect
 - Curculionidae
 - Sitophilus Zeamais
 - Sitophilus Oryzae
 - Pseudococcidae
 - Hemiptera
 - Mealybug
 - Plant
 - Crop
 - Cassava
 - Cereal
 - Rise
 - Grain
 - Oryza
 - Rye
 - Sorghum
 - Oat
 - Wheat
 - Barley
 - Buckwheat
 - Orchidacea
 - Cotton
 - Fruit
 - Cactus
 - Carnivorous Plant
 - Helianthus Annuus
 - Pea

Maize weevil

From Wikipedia, the free encyclopedia

The **maize weevil** (*Sitophilus zeamais*), known in the United States as the **greater rice weevil**,^{[1][2]} is a species of beetle in the family Curculionidae. The **maize weevil can be found in numerous tropical areas around the world, and in the United States, and is a major pest of maize** ^[3] This species attacks both standing crops and stored cereal products, including wheat, rice, sorghum,^{[4][5][6]} oats, barley, rye, buckwheat,^[6] peas, and cottonseed. The maize weevil also infests other types of stored, processed cereal products such as pasta, cassava,^[5] and various coarse, milled grains. It has even been known to attack fruit while in storage, such as apples.^[7]

Maize weevil

Scientific classification

Kingdom: [Animalia](#)
 Phylum: [Arthropoda](#)
 Class: [Insecta](#)
 Order: [Coleoptera](#)
 Family: [Curculionidae](#)
 Genus: [Sitophilus](#)
 Species: ***S. zeamais***

Binomial name

Sitophilus zeamais
 (Motschulsky), 1855

(elytra). It has a long, thin snout, and **oryzae**), but has more clearly marked
 ne another, there are several

Select triples to Add

Select	Subject	Predicate	Object
<input type="checkbox"/>	sitophilus zeamai	isPestOf	barley
<input type="checkbox"/>	sitophilus zeamai	isPestOf	oats
<input type="checkbox"/>	sitophilus zeamai	isPestOf	sorghum
<input type="checkbox"/>	sitophilus zeamai	isPestOf	oryza
<input type="checkbox"/>	sitophilus zeamai	isPestOf	wheat
<input type="checkbox"/>	sitophilus zeamai	isPestOf	cereal
<input type="checkbox"/>	sitophilus zeamai	isPestOf	crop
<input type="checkbox"/>	pea	hasPest	sitophilus zeamai
<input type="checkbox"/>	buckwheat	hasPest	sitophilus zeamai
<input type="checkbox"/>	rye	hasPest	sitophilus zeamai
<input type="checkbox"/>	barley	hasPest	sitophilus zeamai

Add Triple(s) Toggle human readable Cancel

Maize weevil (*S. zeamais*)

Rice weevil (*S. oryzae*)



Contacts

Caterina Caracciolo caterina.caracciolo@fao.org

Armando Stellato stellato@info.uniroma2.it

VocBench site: <http://vocbench.uniroma2.it/>

VocBench pages@FAO: <http://aims.fao.org/tools/vocbench-2/>

VocBench SandBox installation: <http://202.73.13.50:55481/vocbench/>

You can also follow VB by registering to:

- AIMS Community Site: <http://aims.fao.org/> (you can selected the topics you are interested into)
- VocBench Mailing Lists:
 - User: <http://groups.google.com/group/vocbench-user>
 - Developer: <http://groups.google.com/group/vocbench-developer>
- Semantic Turkey Mailing Lists:
 - User: <http://groups.google.com/group/semanticturkey-user>
 - Developer: <http://groups.google.com/group/semanticturkey-developer>

VocBench – features overview

- Full support to multilinguality
- Supports collaborative editing
- Formalized workflow
- Scheme and data management

- Open source

VocBench - Pointers

- VocBench Sandbox v2.2 (latest stable):
<http://202.73.13.50:55481/vocbench/>
- VB project
- VB users mailing list
- AIMS community, for news:
<http://aims.fao.org/user/register>
- ..