Ontology-related data and reengineering

Design by Re-engineering

Aldo Gangemi

Laboratory for Applied Ontology (ISTC-CNR), Roma

{aldo.gangemi, valentina.presutti}@istc.cnr.it

Course Outline

- Ontologies and the Semantic Web
- Ontology Design and Ontology Design Patterns
- Content Ontology Design Patterns
- Design by Re-Engineering

Reengineering OPs

Definition

- Reengineering OPs are transformation rules applied in order to create a new ontology (target model) starting from elements of a source model

- The target model is an ontology, while the source model can be either an ontology, or a non-ontological resource
  - e.g., a thesaurus concept, a data model pattern, a UML model, a linguistic structure, etc.

- Two types:
  - Schema reengineering OPs are rules for transforming a non-OWL DL metamodel into an OWL DL ontology

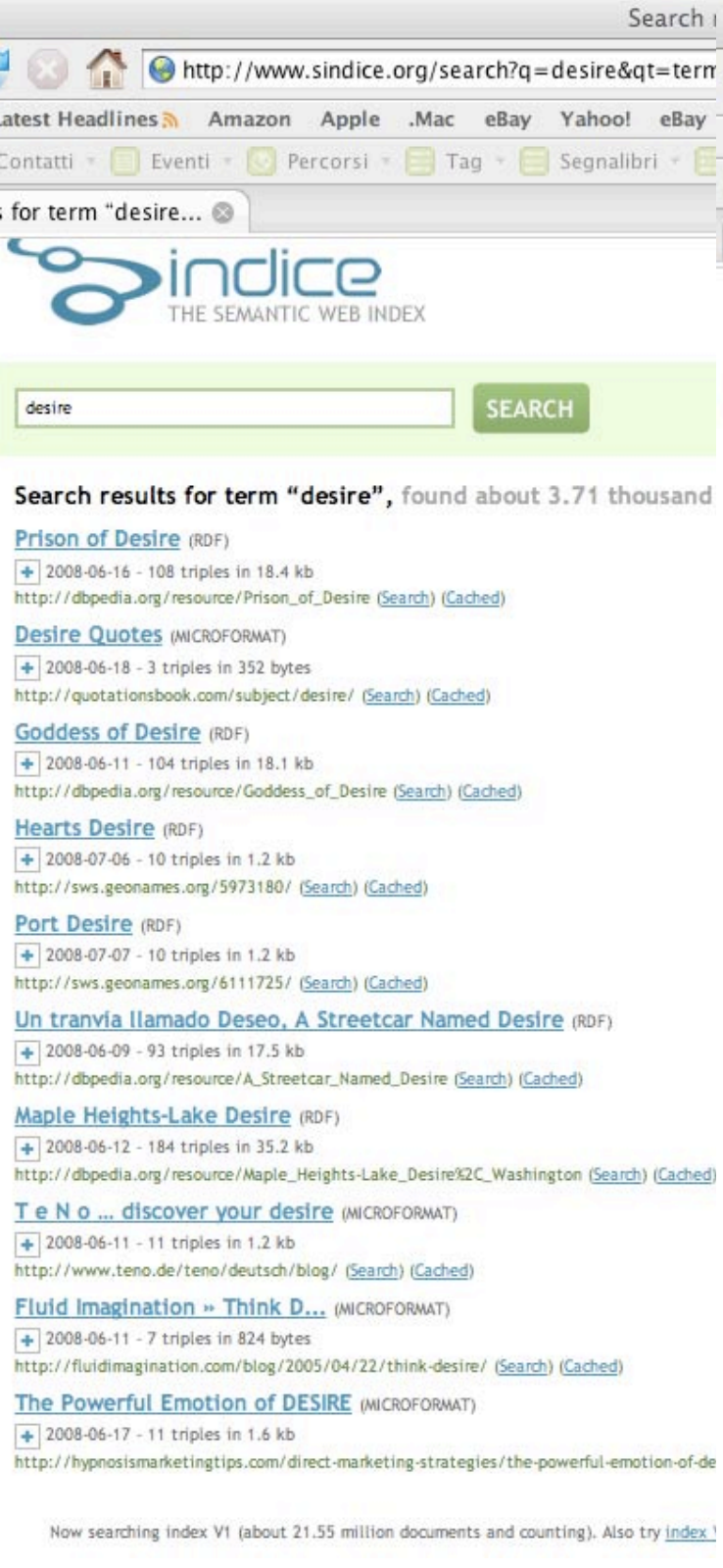
  - Refactoring OPs provide designers with rules for transforming, i.e. “refactoring”, an existing OWL DL “source” ontology into a new OWL DL “target” ontology
    - E.g. a guideline to reengineer a piece of an OWL ontology in presence of a requirement change, as when moving from individuals to classes, or from object properties to classes. See also N-ary relation tranformation pattern

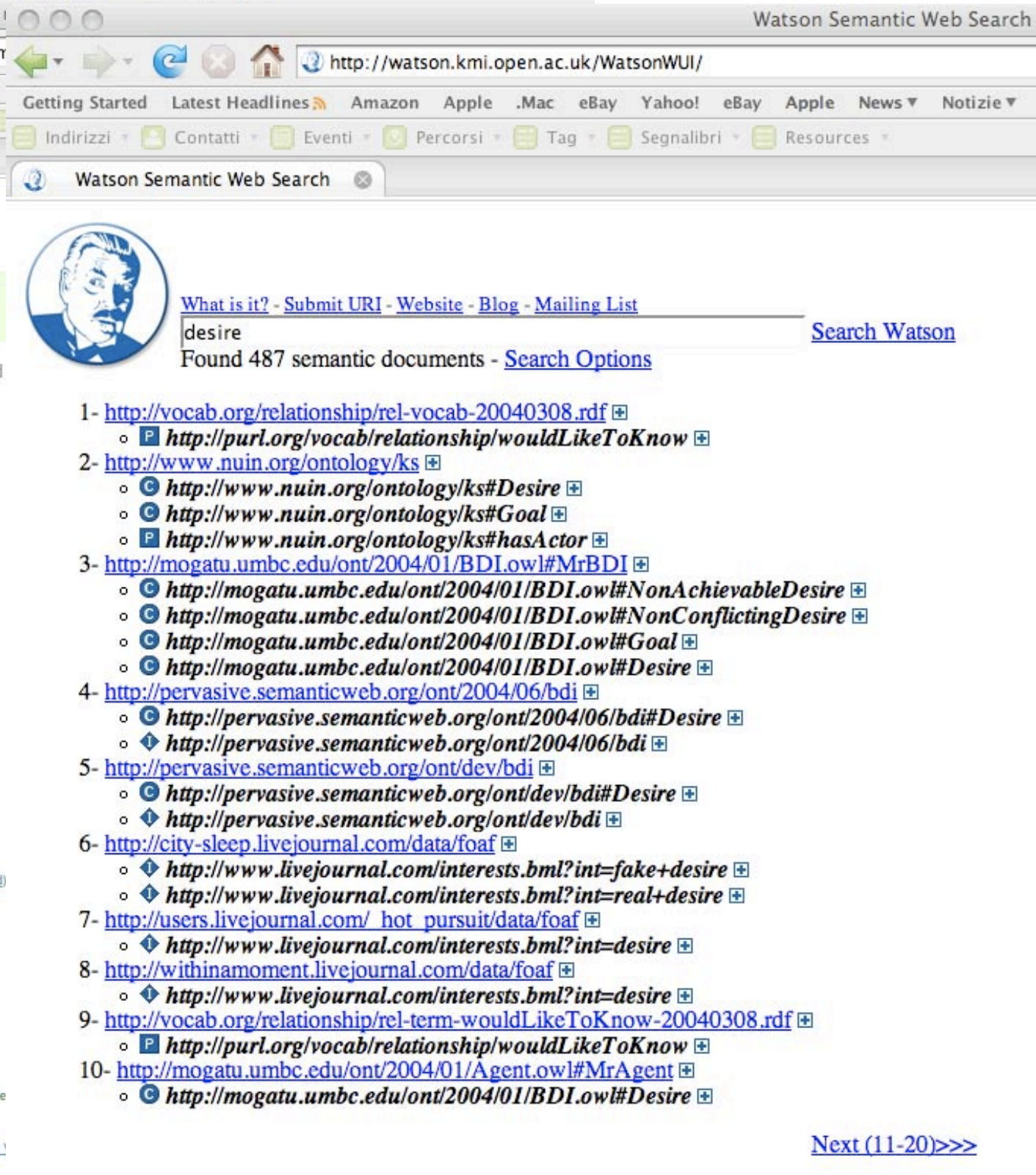
Ontology-related data: knowledge resource types

- Modeling Languages
  - E/R, UML, XSD, Petri Nets, ebXML, BPEL4WS
- Conceptual models
  - Database schemas, UML diagrams, XSD schemas, etc.
- Informal Data Structures
  - Spreadsheets, tables, etc.
- Lexical resources
  - WordNet, FrameNet, Oxford Dictionary, etc.
- Concept Schemes
  - Thesauri, classifications, nomenclatures, etc.
- Open tag systems
  - Flickr, Wikipedia, MySpace, ...
- Linked Open Data
  - DBpedia, Microformats, RDFa, etc.
- Text extractors
  - Text2Onto, TermExtractor, SST, Frame Detector, ...

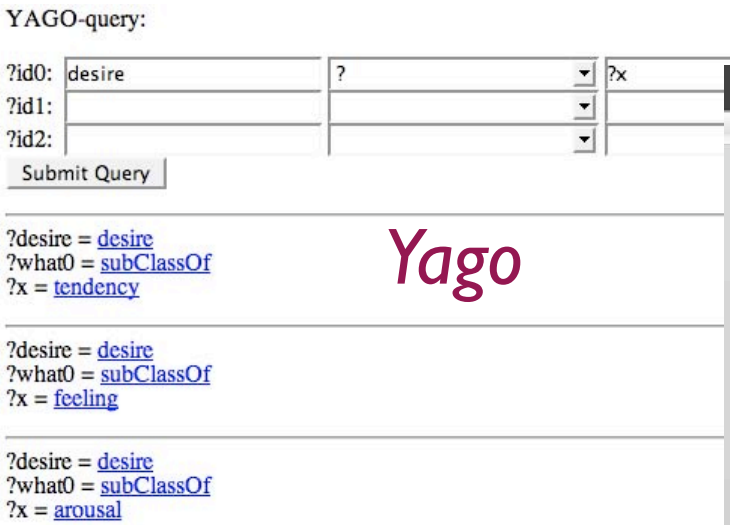
Searching and using ontologies, on-the-fly data reengineering

- Watson and the NeOn Toolkit
- Sindice
- Yago
- Umbel
- Freebase
- OpenLink Data Explorer over Linked Open Data
- GRDDL, RDFa and Microformats

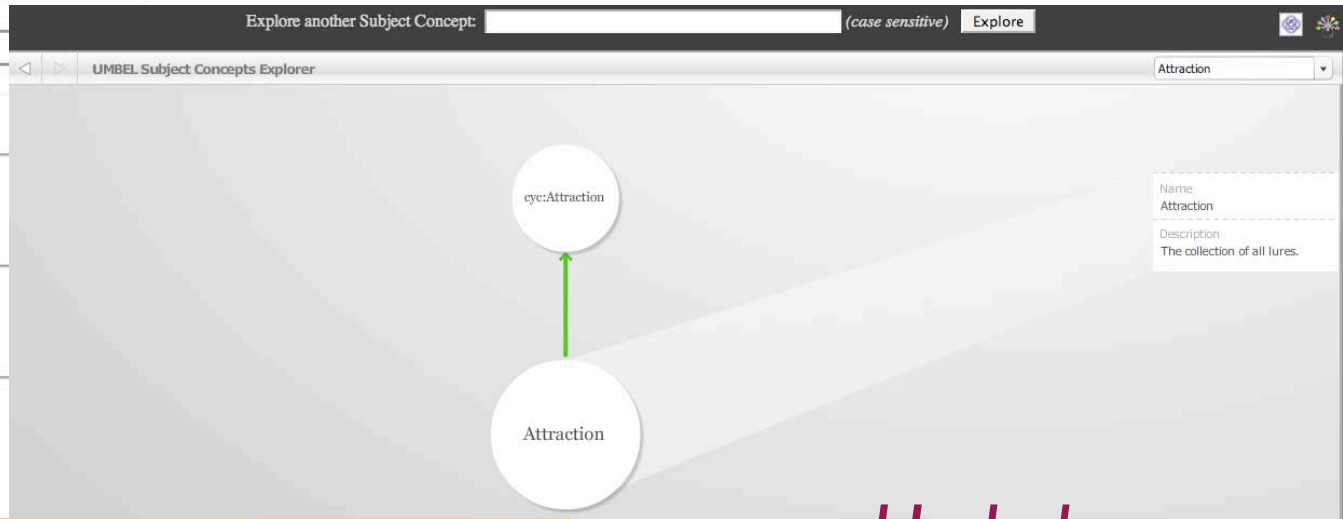

Search
http://www.sindice.org/search?q=desire&qt=term
Latest Headlines
Amazon
Apple
.Mac
eBay
Yahoo!
eBay
Contatti
Eventi
Percorsi
Tag
Segnalibri
for term "desire...
indice
THE SEMANTIC WEB INDEX
desire
SEARCH
Search results for term "desire", found about 3.71 thousand
Prison of Desire (RDF)
2008-06-16 - 108 triples in 18.4 kb
http://dbpedia.org/resource/Prison_of_Desire (Search) (Cached)
Desire Quotes (MICROFORMAT)
2008-06-18 - 3 triples in 352 bytes
http://quotationsbook.com/subject/desire/ (Search) (Cached)
Goddess of Desire (RDF)
2008-06-11 - 104 triples in 18.1 kb
http://dbpedia.org/resource/Goddess_of_Desire (Search) (Cached)
Hearts Desire (RDF)
2008-07-06 - 10 triples in 1.2 kb
http://sws.geonames.org/5973180/ (Search) (Cached)
Port Desire (RDF)
2008-07-07 - 10 triples in 1.2 kb
http://sws.geonames.org/6111725/ (Search) (Cached)
Un tranvia llamado Deseo, A Streetcar Named Desire (RDF)
2008-06-09 - 93 triples in 17.5 kb
http://dbpedia.org/resource/A_Streetcar_Named_Desire (Search) (Cached)
Maple Heights-Lake Desire (RDF)
2008-06-12 - 184 triples in 35.2 kb
http://dbpedia.org/resource/Maple_Heights-Lake_Desire%2C_Washington (Search) (Cached)
T e N o ... discover your desire (MICROFORMAT)
2008-06-11 - 11 triples in 1.2 kb
http://www.teno.de/teno/deutsch/blog/ (Search) (Cached)
Fluid Imagination » Think D... (MICROFORMAT)
2008-06-11 - 7 triples in 824 bytes
http://fluidimagination.com/blog/2005/04/22/think-desire/ (Search) (Cached)
The Powerful Emotion of DESIRE (MICROFORMAT)
2008-06-17 - 11 triples in 1.6 kb
http://hypnosismarketingtips.com/direct-marketing-strategies/the-powerful-emotion-of-de
Now searching index V1 (about 21.55 million documents and counting). Also try index


Watson Semantic Web Search
http://watson.kmi.open.ac.uk/WatsonWUI/
Getting Started
Latest Headlines
Amazon
Apple
.Mac
eBay
Yahoo!
eBay
Apple
News
Notizie
Indirizzi
Contatti
Eventi
Percorsi
Tag
Segnalibri
Resources
Watson Semantic Web Search
What is it? - Submit URI - Website - Blog - Mailing List
desire
Search Watson
Found 487 semantic documents - Search Options
1- http://vocab.org/relationship/rel-vocab-20040308.rdf
P http://purl.org/vocab/relationship/wouldLikeToKnow
2- http://www.nuin.org/ontology/ks
C http://www.nuin.org/ontology/ks#Desire
C http://www.nuin.org/ontology/ks#Goal
P http://www.nuin.org/ontology/ks#hasActor
3- http://mogatu.umbc.edu/ont/2004/01/BDI.owl#MrBDI
C http://mogatu.umbc.edu/ont/2004/01/BDI.owl#NonAchievableDesire
C http://mogatu.umbc.edu/ont/2004/01/BDI.owl#NonConflictingDesire
C http://mogatu.umbc.edu/ont/2004/01/BDI.owl#Goal
C http://mogatu.umbc.edu/ont/2004/01/BDI.owl#Desire
4- http://pervasive.semanticweb.org/ont/2004/06/bdi
C http://pervasive.semanticweb.org/ont/2004/06/bdi#Desire
I http://pervasive.semanticweb.org/ont/2004/06/bdi
5- http://pervasive.semanticweb.org/ont/dev/bdi
C http://pervasive.semanticweb.org/ont/dev/bdi#Desire
I http://pervasive.semanticweb.org/ont/dev/bdi
6- http://city-sleep.livejournal.com/data/foaf
I http://www.livejournal.com/interests.bml?int=fake+desire
I http://www.livejournal.com/interests.bml?int=real+desire
7- http://users.livejournal.com/_hot_pursuit/data/foaf
I http://www.livejournal.com/interests.bml?int=desire
8- http://withinamoment.livejournal.com/data/foaf
I http://www.livejournal.com/interests.bml?int=desire
9- http://vocab.org/relationship/rel-term-wouldLikeToKnow-20040308.rdf
P http://purl.org/vocab/relationship/wouldLikeToKnow
10- http://mogatu.umbc.edu/ont/2004/01/Agent.owl#MrAgent
C http://mogatu.umbc.edu/ont/2004/01/BDI.owl#Desire
Next (11-20)>>>

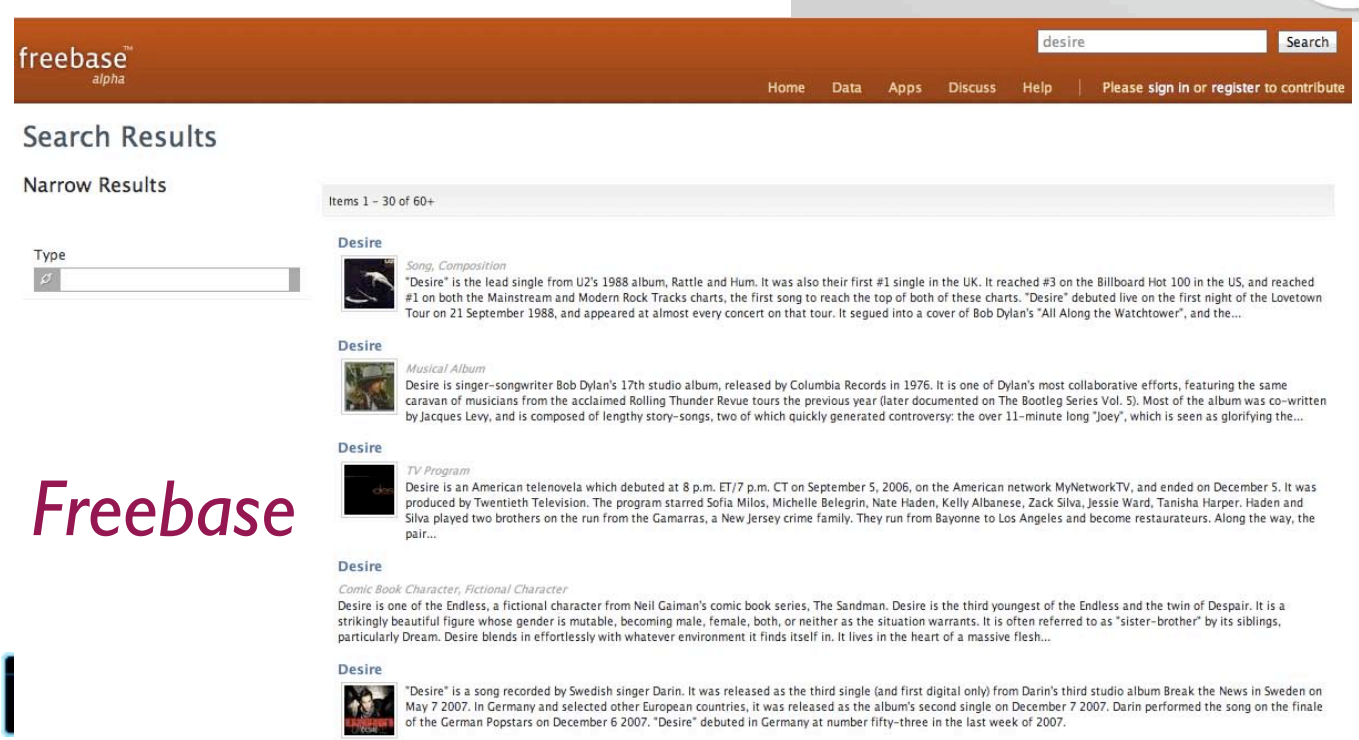
Integrated knowledge search



Yago

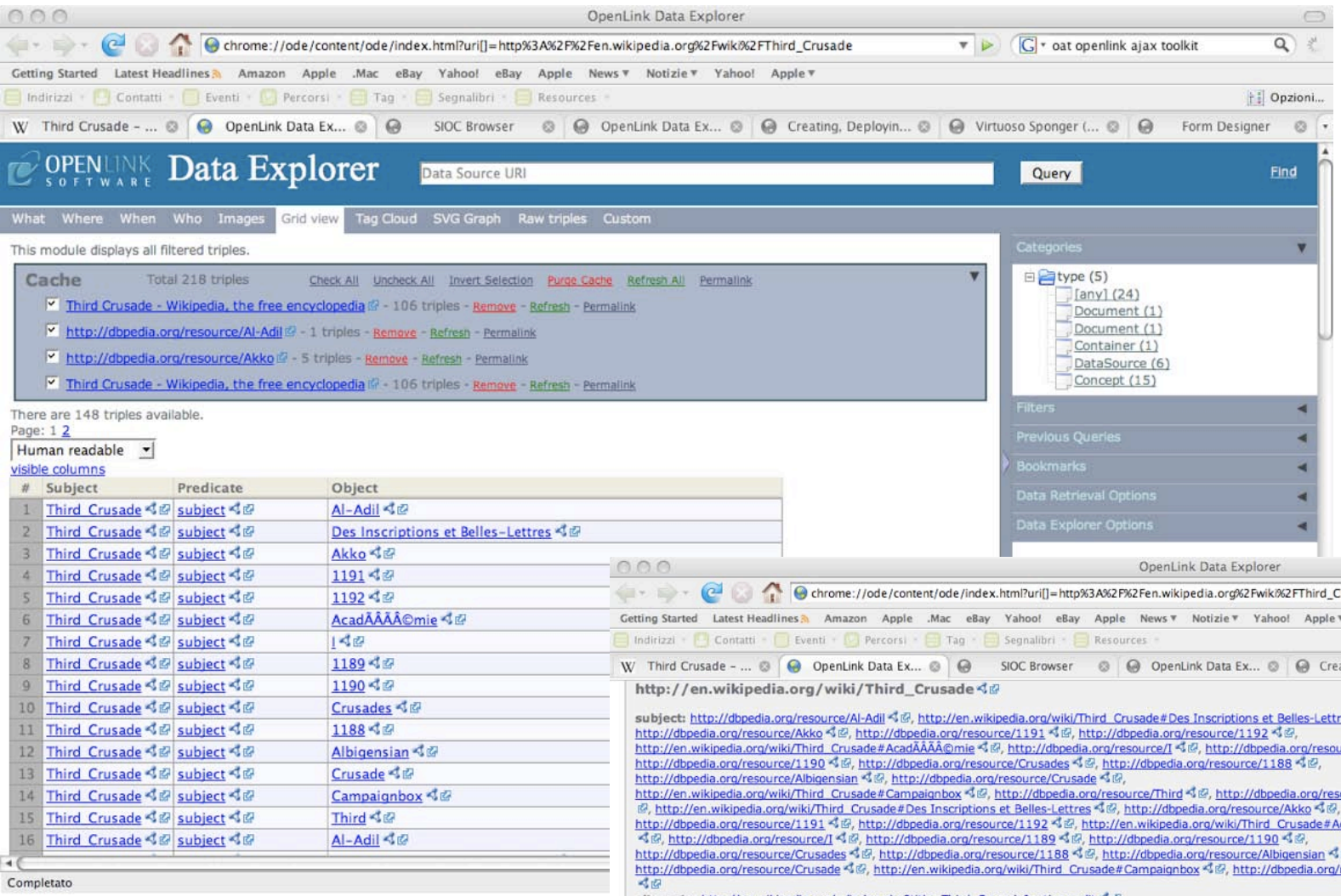


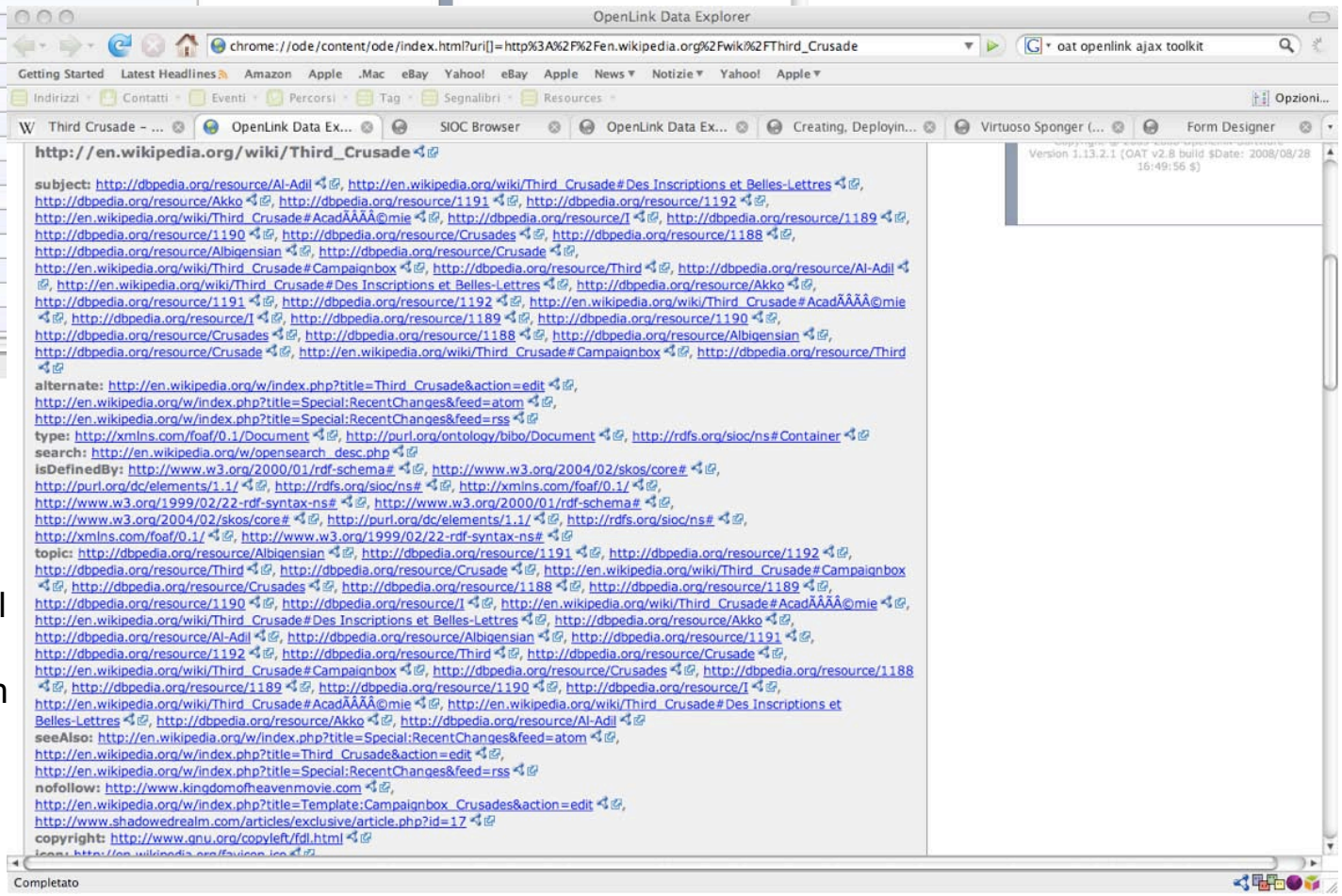
Umbel



Freebase

everything based on a centralized ontology ... of mixed quality





RDF triples can contain all sorts of relations, as shown with reference to DBpedia triples related to the Third Crusade: *subjects, alternate pages, copyright forms, types, icons*, etc. are all related to the Third Crusade page in a way that makes the results of the search still confusing (although better than googling, or browsing WikiPedia for a

How Linked Data Materialize on the Web

Generated "on the fly" via RDF middleware from:

- Existing Web Pages (POSH, Microformats, eRDF, RDFa, GRDDL)
- Web 2.0 Data Spaces (Social Networks, Blogs, Wikis, Bookmarks, Online Discussions / Conversations etc)
- Web Services (SOAP and REST)

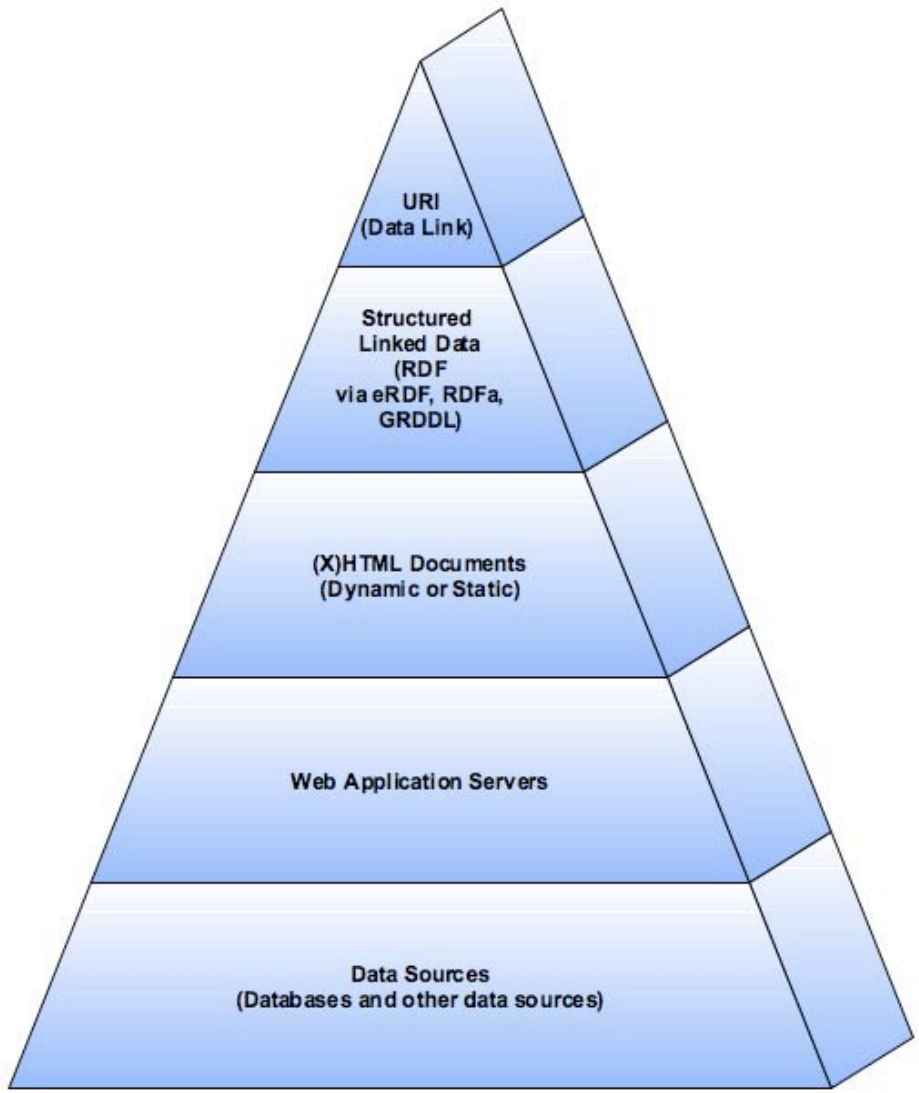
Enterprise Data Sources

- SOA oriented Web Services
- XML based Data Warehouses and Views
- Enterprise Databases (ODBC, JDBC, OLE-DB, ADO.NET, XMLA, Native CLIs)

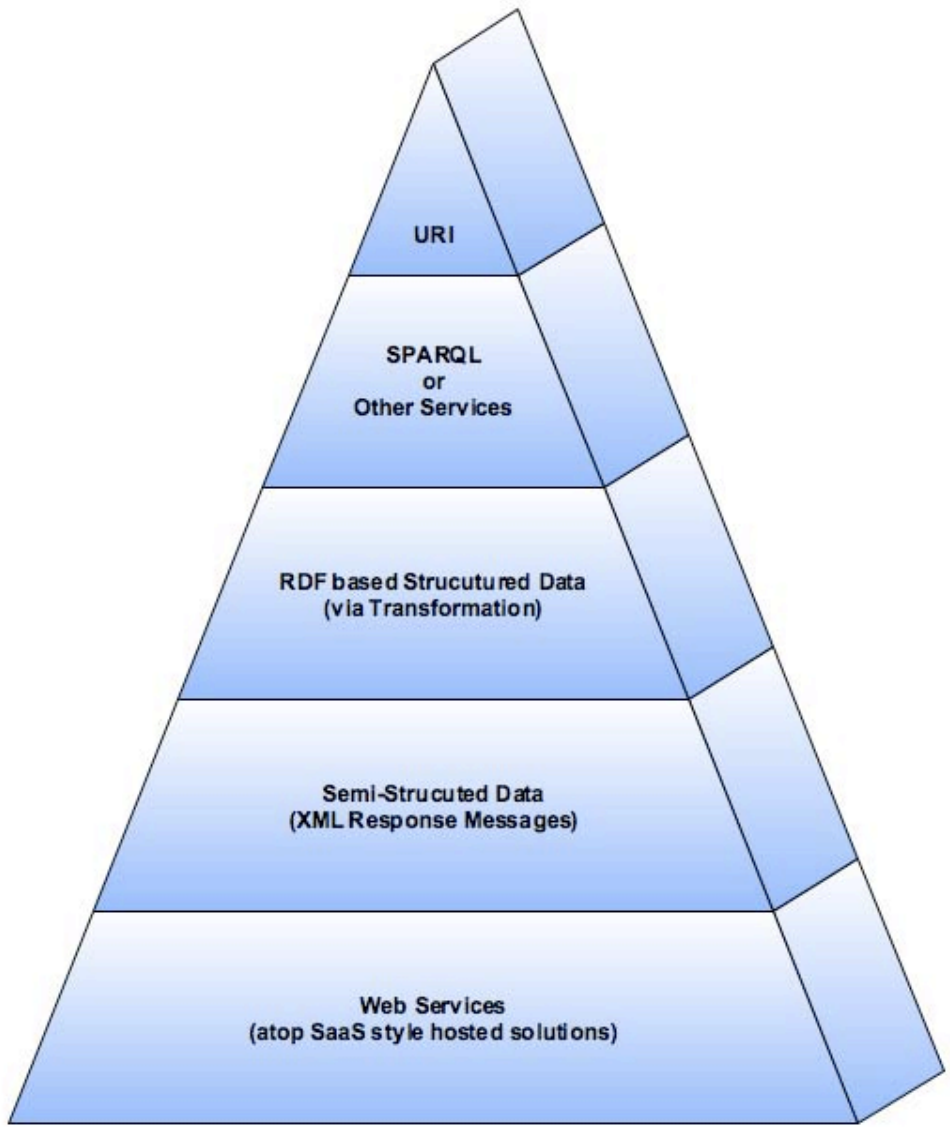
Community driven extraction efforts

- DBpedia, Bio2RDF, and many other Linking Open Data projects

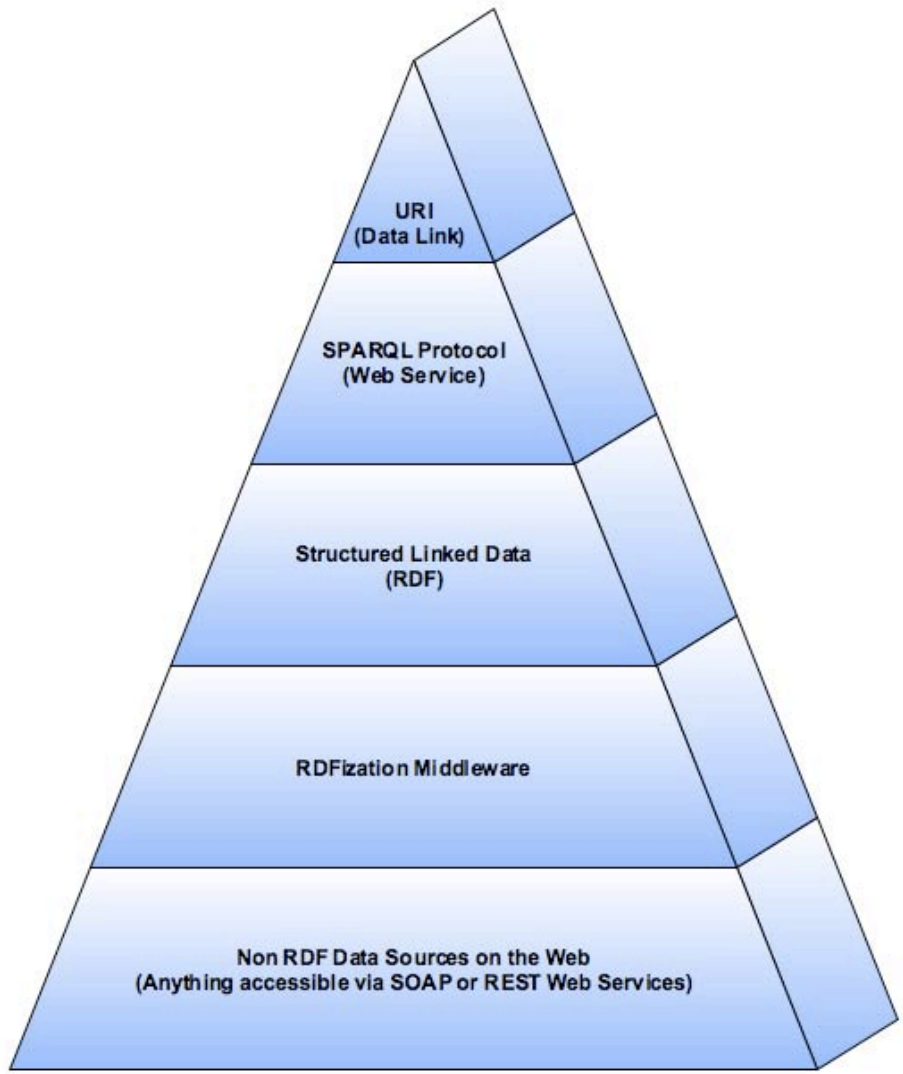
Reengineering dynamic web content



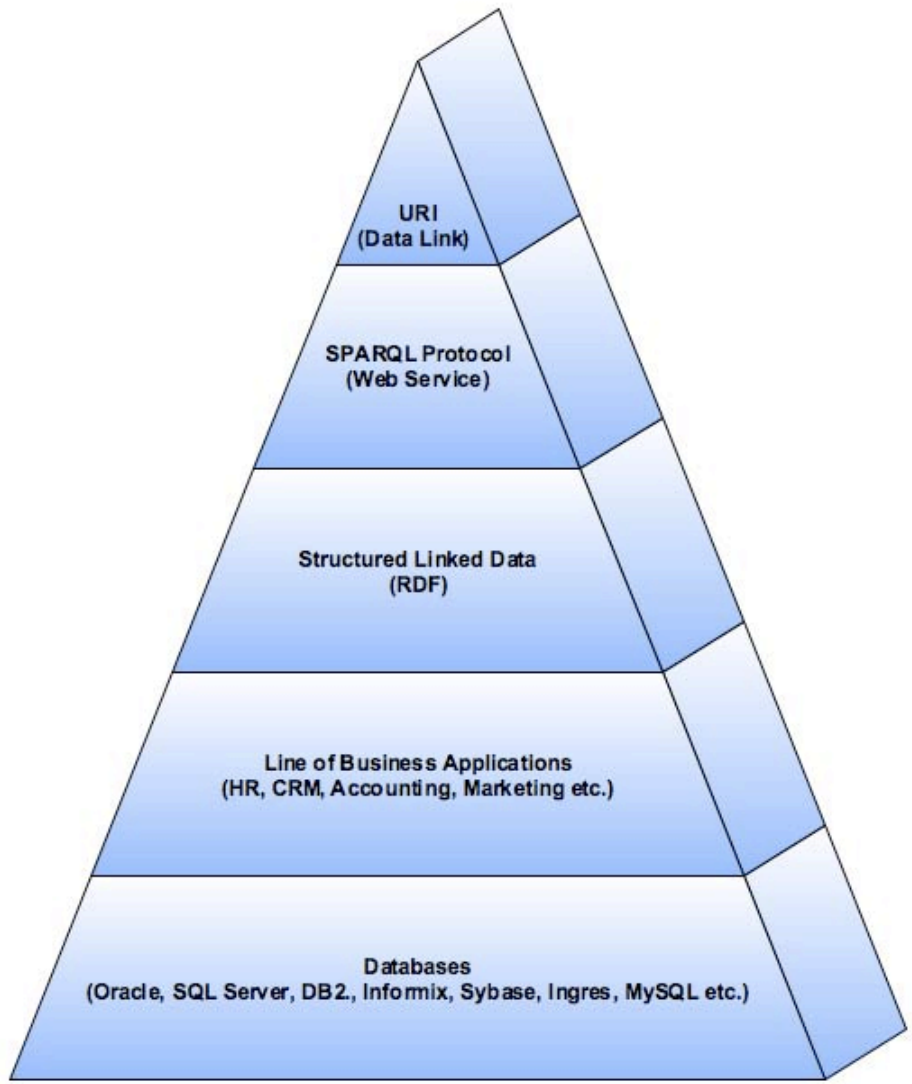
Reengineering web service content



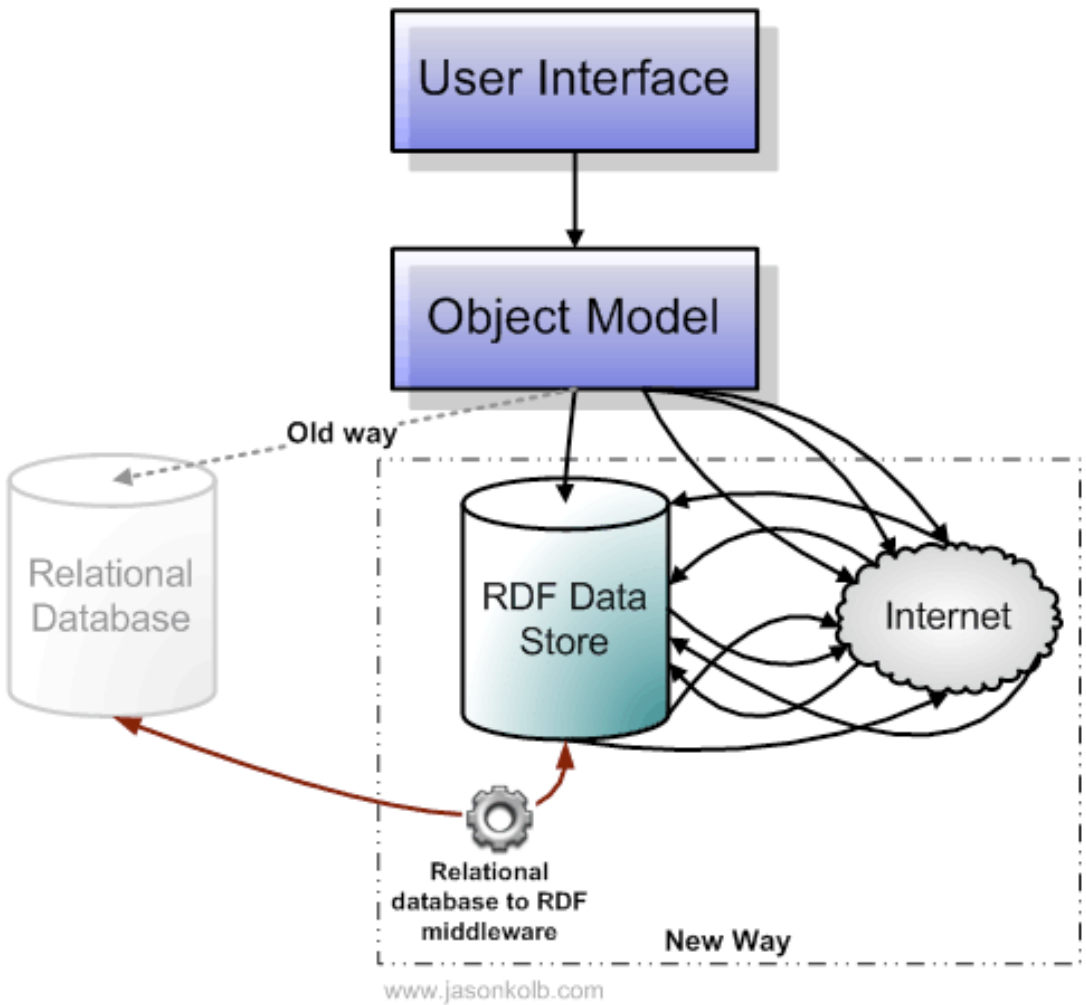
Reengineering web content



Reengineering enterprise data



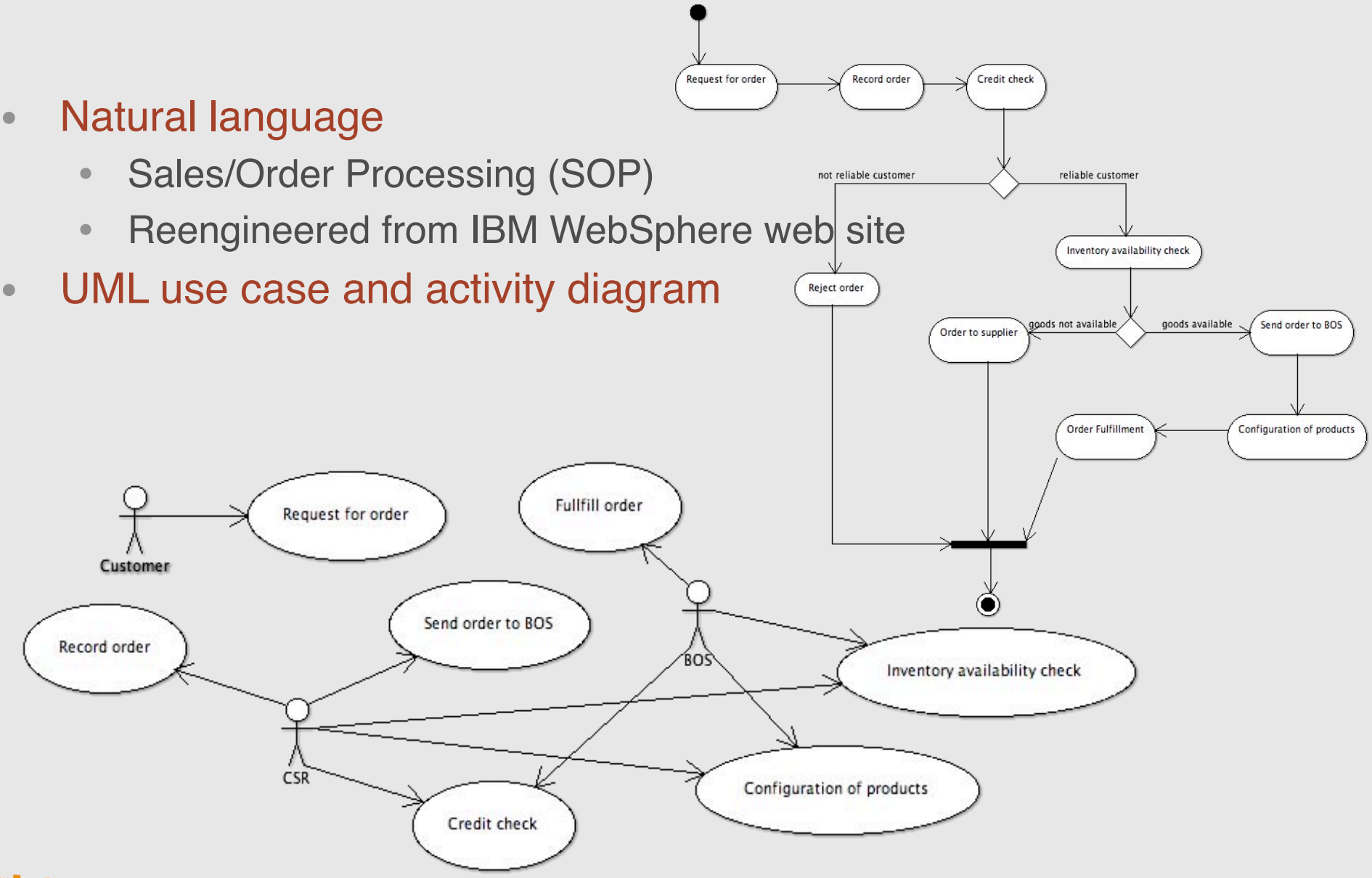
Reengineering relational databases



Modeling Languages, conceptual schemas, and informal data structures

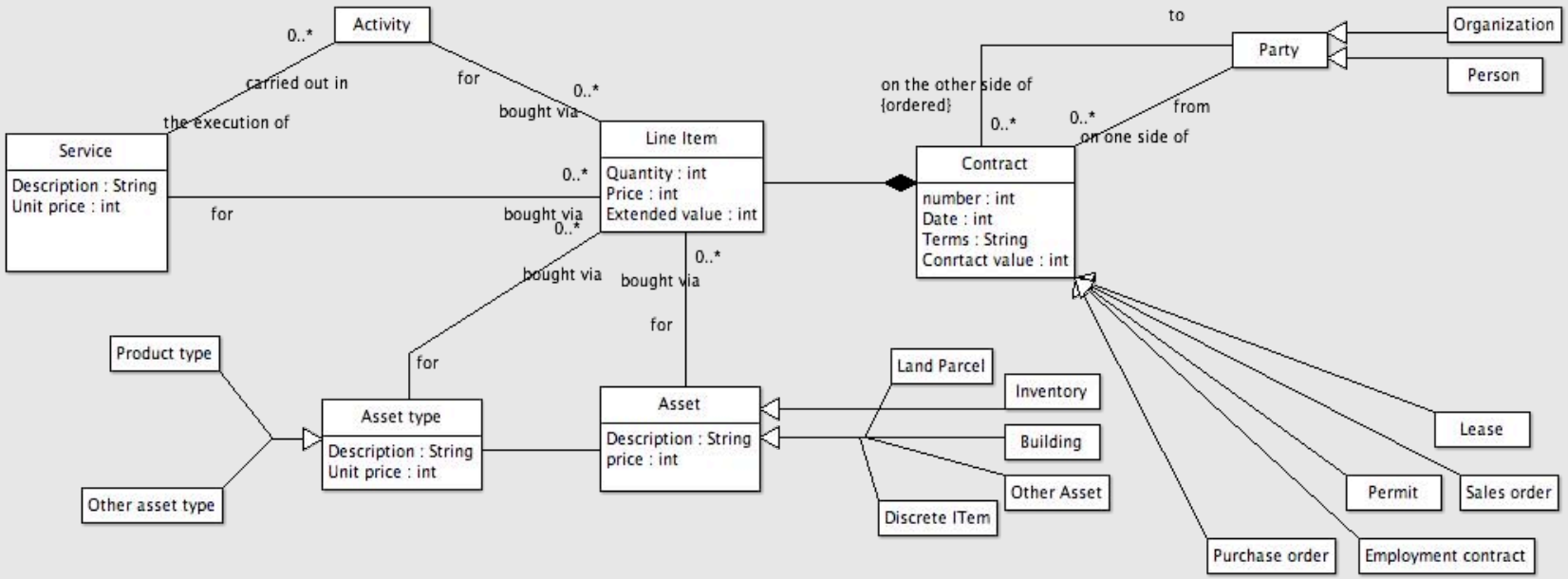
Example: Sales/Order Process 1/2

- Natural language
  - Sales/Order Processing (SOP)
  - Reengineered from IBM WebSphere web site
- UML use case and activity diagram

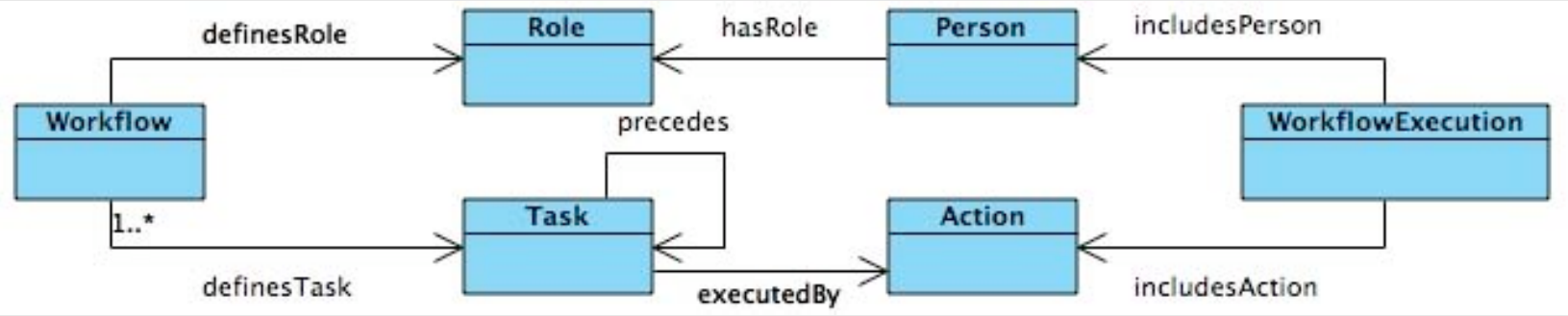


Example: Sales/Order Process 2/2

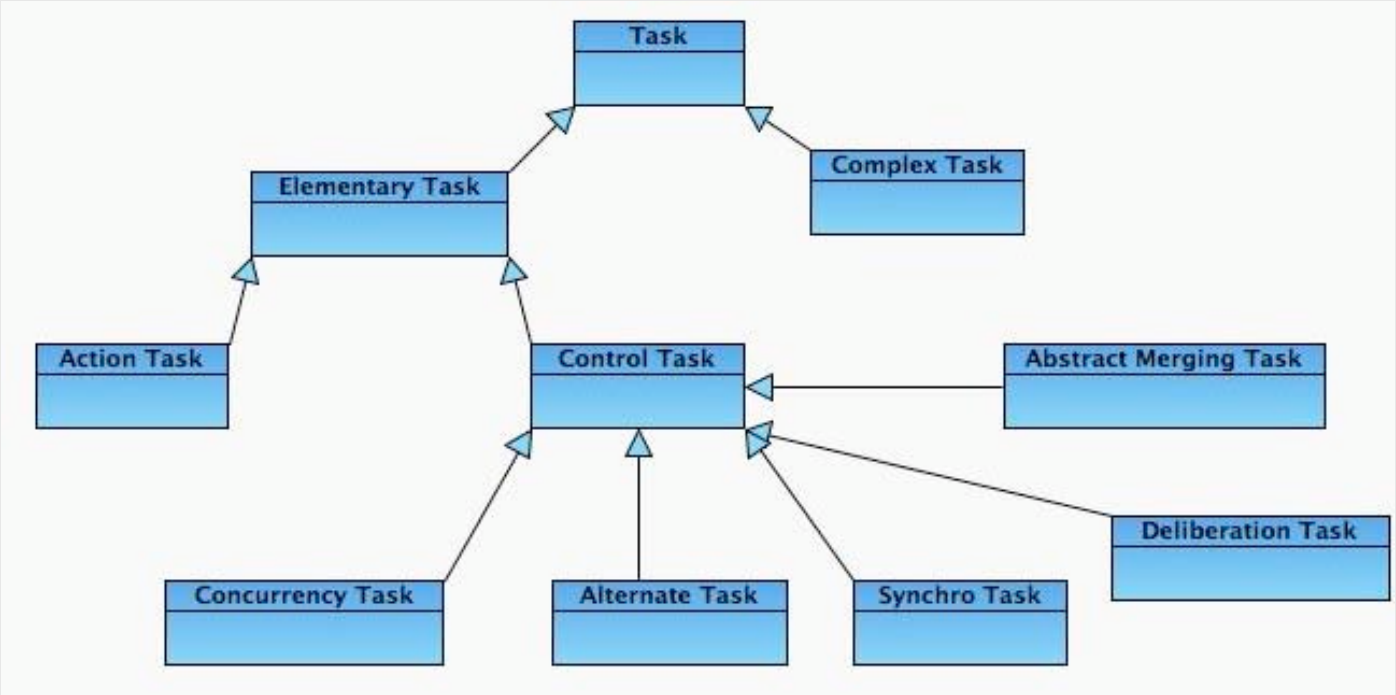
- Data model patterns
  - *Kinds of Contracts*
  - re-engineered from 'Data Model Patterns' (D.C. Hay. 96) Workflow patterns



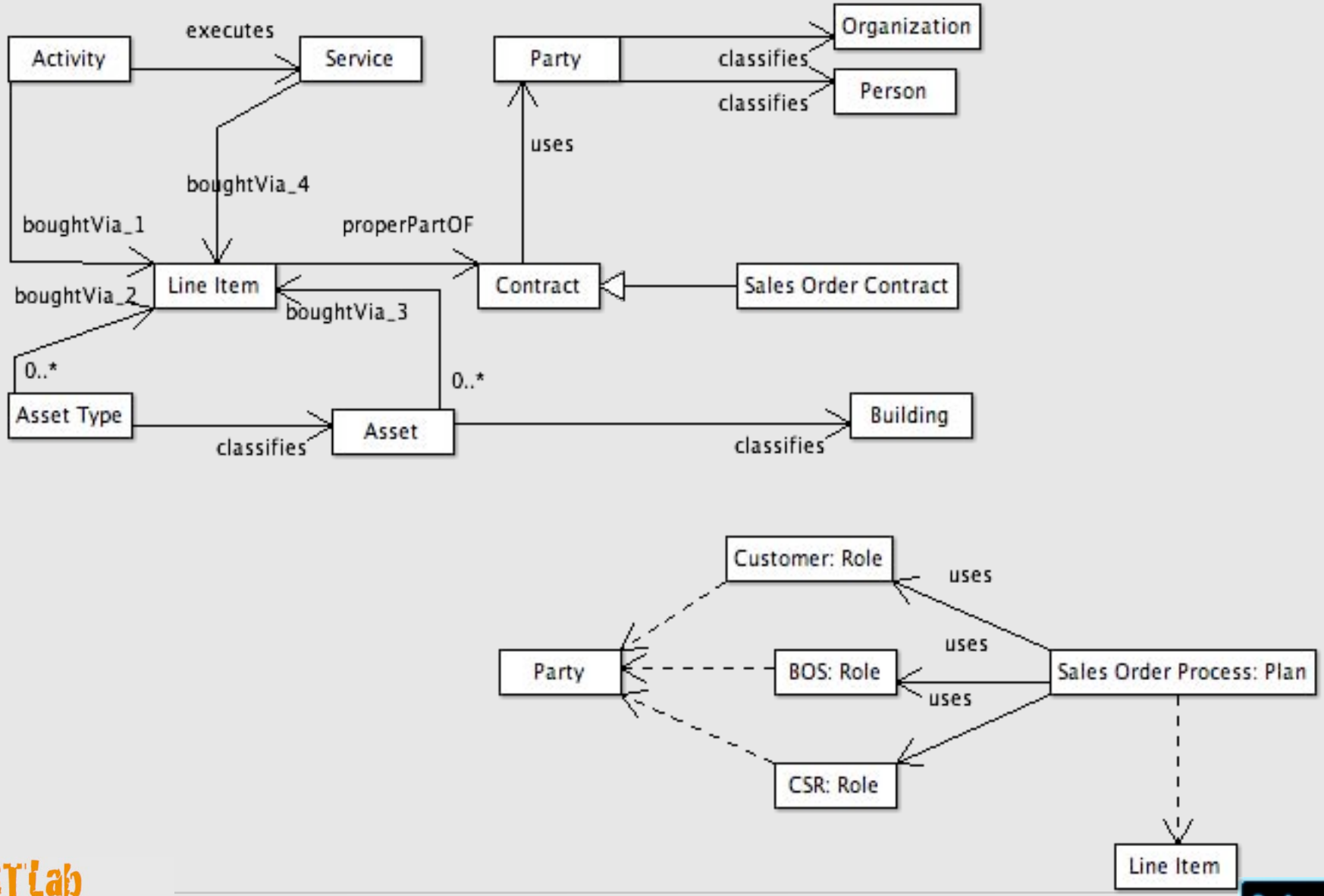
Workflow: CP specialization



Some workflow patterns (re-engineered from van der Aalst)

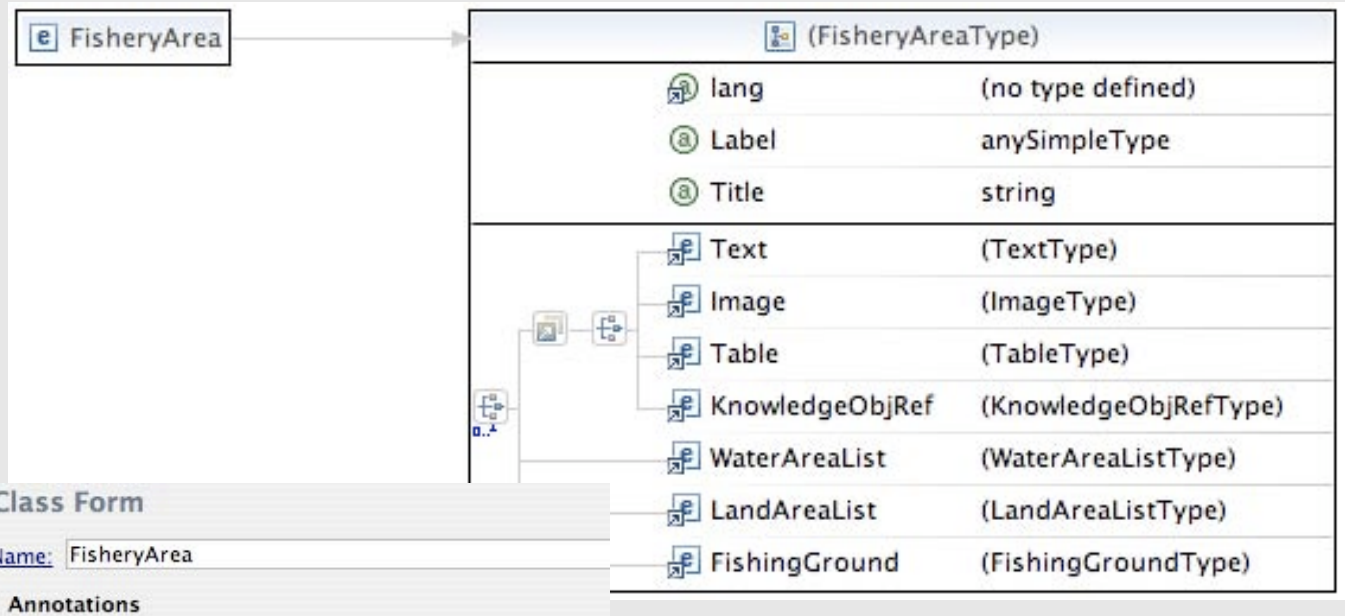


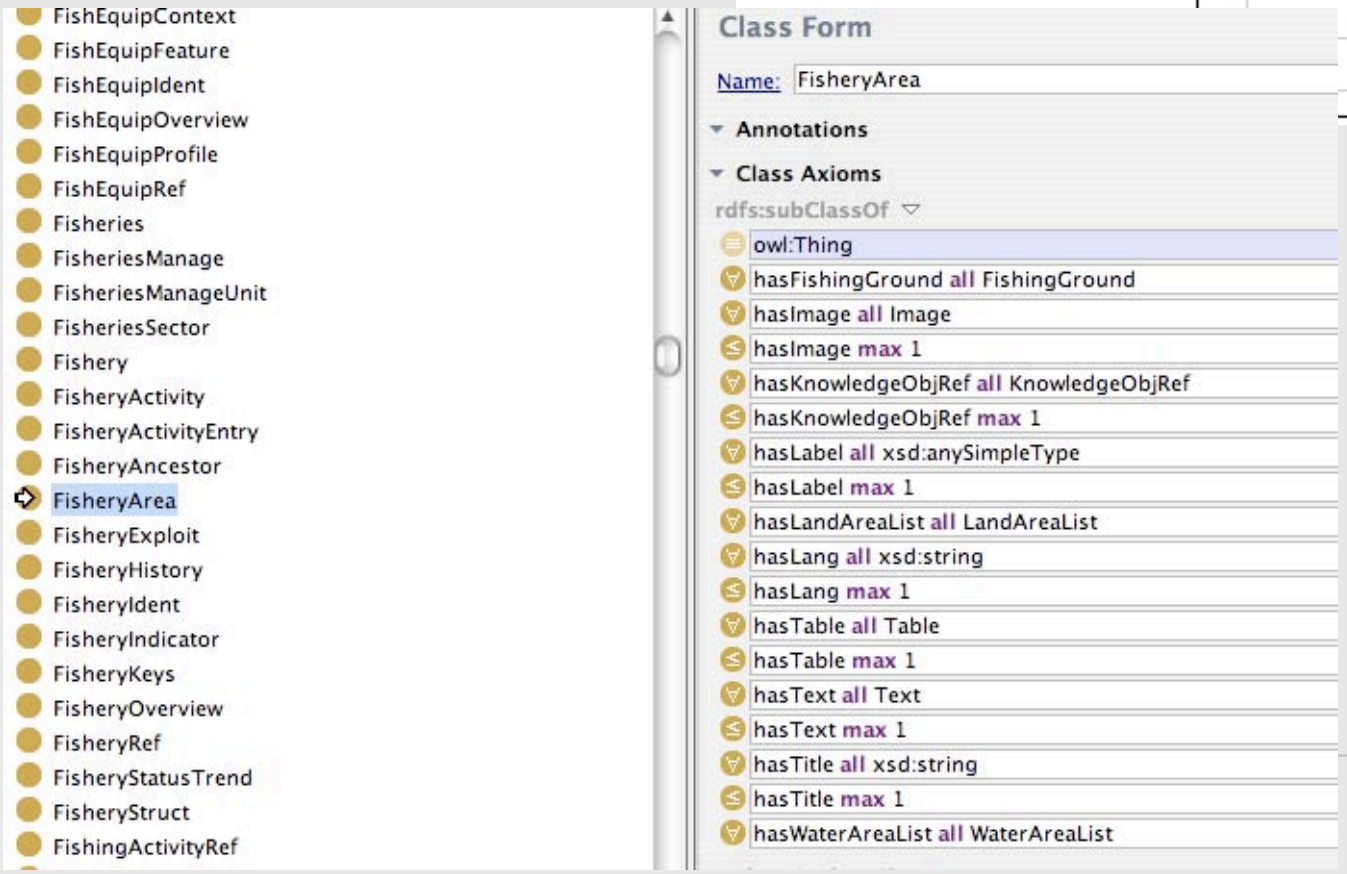
Merging data models and workflow patterns in OWL CPs



Handbook of Ontologies for Business Interaction. The IDEA Group 2007

An example from a DTD-based XSD





Spreadsheet2RDF (e.g. rdf123)

```
rdf123:metadata,,,,,,,,,
title,Computing lab members,,,,,,,,
comment,8 June 2007,,,,,,,,
row head,true,,,,,,,,
start row,8,,,,,,,,
type,rdf123:ConvertedSpreadsheetInRDF,,

NAME,EMAIL,OFFICE,,,,,,,
Al Turing,amt@umbc.edu,ITE332,,,,,,,
Don Knuth,dek@umbc.edu,ITE332,,,,,,,
Marvin Minsky,mlm@umbc.edu,ITE442,,,,,,,
```

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix rdf123: <http://rdf123.org/> .


< > a rdf123:ConvertedSpreadsheetInRDF;
dc:title "Computing lab members;
rdfs:comment "8 June 2007".

[]  a foaf:Person;
foaf:name "Al Turing";
foaf:mbox "amt@umbc.edu";
foaf:offceNumber "ITE332".

[]  a foaf:Person;
foaf:name "Don Knuth"; ...
...
```

Lexical resources, concept schemas, and web 2.0

Linguistic dictionaries and thesauri

Oxford American Dictionary

desire |dəˈzī(ə)r| |dəˌzaɪ(ə)r| |diˌzaɪ(ə)r| |dɪˌzʌɪə|

noun

a strong feeling of wanting to have something or wishing for something to happen : [with infinitive] *a desire to work in the dirt with your bare hands.*

• strong sexual feeling or appetite : *they were clinging together in fierce mutual desire.*

verb [trans.]

strongly wish for or want (something) : *he never achieved the status he so desired* | [as adj.] (**desired**) *it failed to create the desired effect.*

• want (someone) sexually : *there had been a time, years ago, when he had desired her.*

• archaic express a wish to (someone); request or entreat.

ORIGIN Middle English : from Old French ***desir*** (noun), ***desirer*** (verb), from Latin ***desiderare*** (see **desiderate**).

Thesaurus

desire

noun

1 *a desire to see the world* wish, want, aspiration, fancy, inclination, impulse; yearning, longing, craving, hankering, hunger; eagerness, enthusiasm, determination; informal yen, itch, jones.

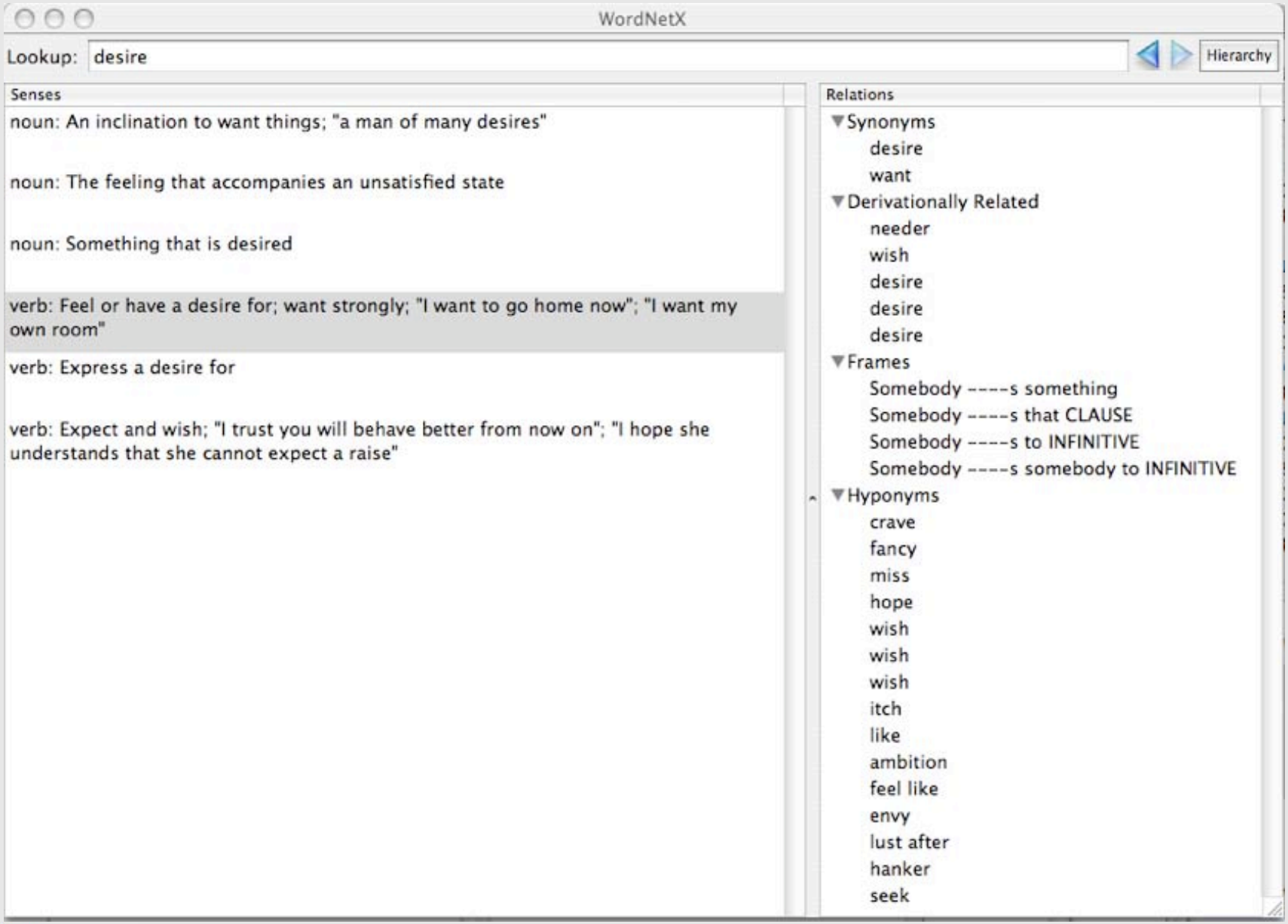
2 *his eyes glittered with desire* lust, sexual attraction, passion, sensuality, sexuality; lasciviousness, lechery, salaciousness, libidinousness; informal the hots, raunchiness, horniness.

verb

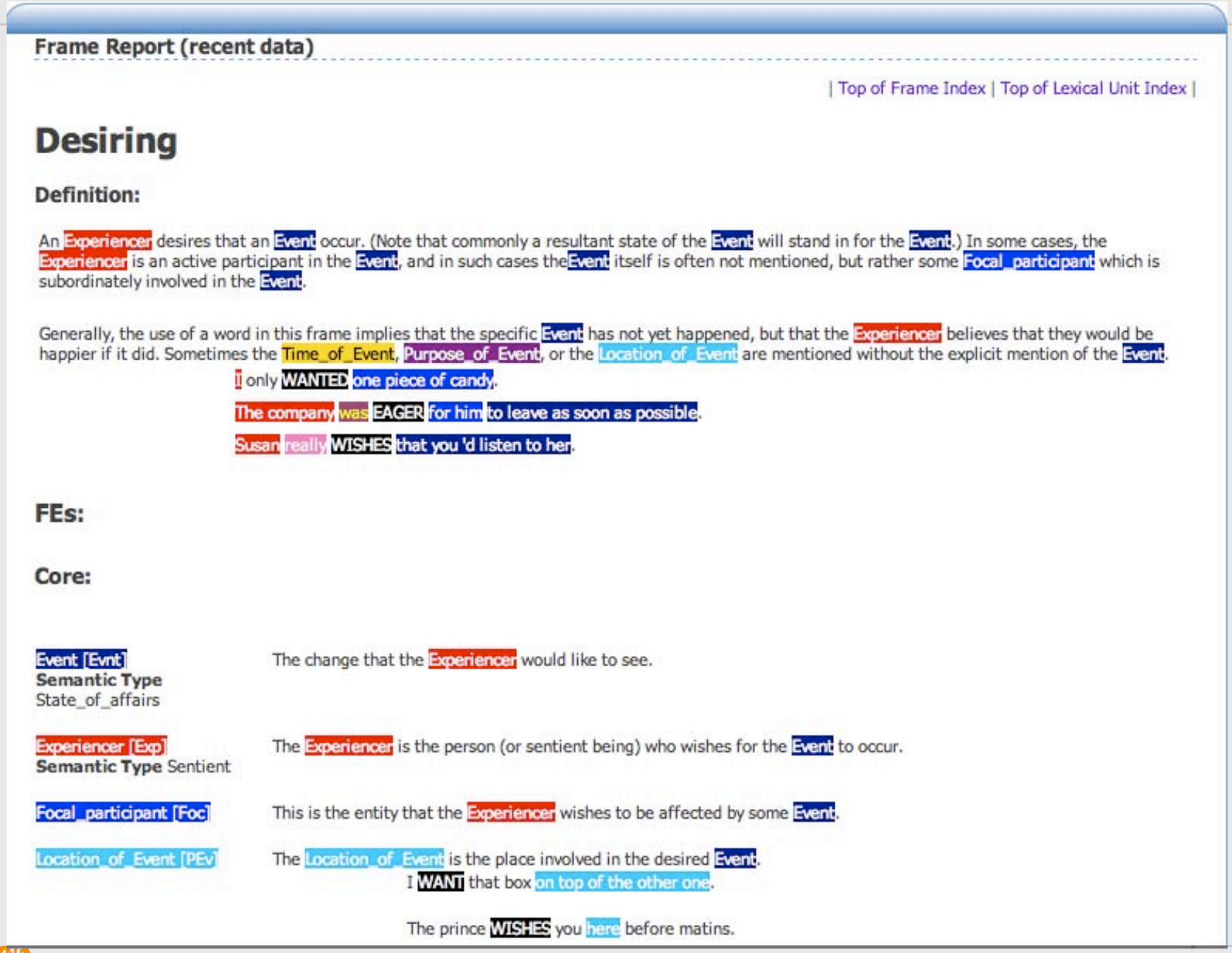
1 *they desired peace* want, wish for, long for, yearn for, crave, hanker after, be desperate for, be bent on, covet, aspire to; fancy; informal have a yen for, have a jones for, yen for, hanker after/for.

2 *she desired him* be attracted to, lust after, burn for, be infatuated by; informal fancy, have the hots for, have a crush on, be mad about, be crazy about.

WordNets


WordNetX
Lookup: desire
Hierarchy
Senses
noun: An inclination to want things; "a man of many desires"
noun: The feeling that accompanies an unsatisfied state
noun: Something that is desired
verb: Feel or have a desire for; want strongly; "I want to go home now"; "I want my own room"
verb: Express a desire for
verb: Expect and wish; "I trust you will behave better from now on"; "I hope she understands that she cannot expect a raise"
Relations
Synonyms
desire
want
Derivationally Related
needer
wish
desire
desire
desire
Frames
Somebody ----s something
Somebody ----s that CLAUSE
Somebody ----s to INFINITIVE
Somebody ----s somebody to INFINITIVE
Hyponyms
crave
fancy
miss
hope
wish
wish
wish
itch
like
ambition
feel like
envy
lust after
hanker
seek

FrameNets



Frame Report (recent data)

| Top of Frame Index | Top of Lexical Unit Index |

Desiring

Definition:

An Experiencer desires that an Event occur. (Note that commonly a resultant state of the Event will stand in for the Event.) In some cases, the Experiencer is an active participant in the Event, and in such cases the Event itself is often not mentioned, but rather some Focal_participant which is subordinately involved in the Event.

Generally, the use of a word in this frame implies that the specific Event has not yet happened, but that the Experiencer believes that they would be happier if it did. Sometimes the Time_of_Event, Purpose_of_Event, or the Location_of_Event are mentioned without the explicit mention of the Event.

I only WANTED one piece of candy.

The company was EAGER for him to leave as soon as possible.

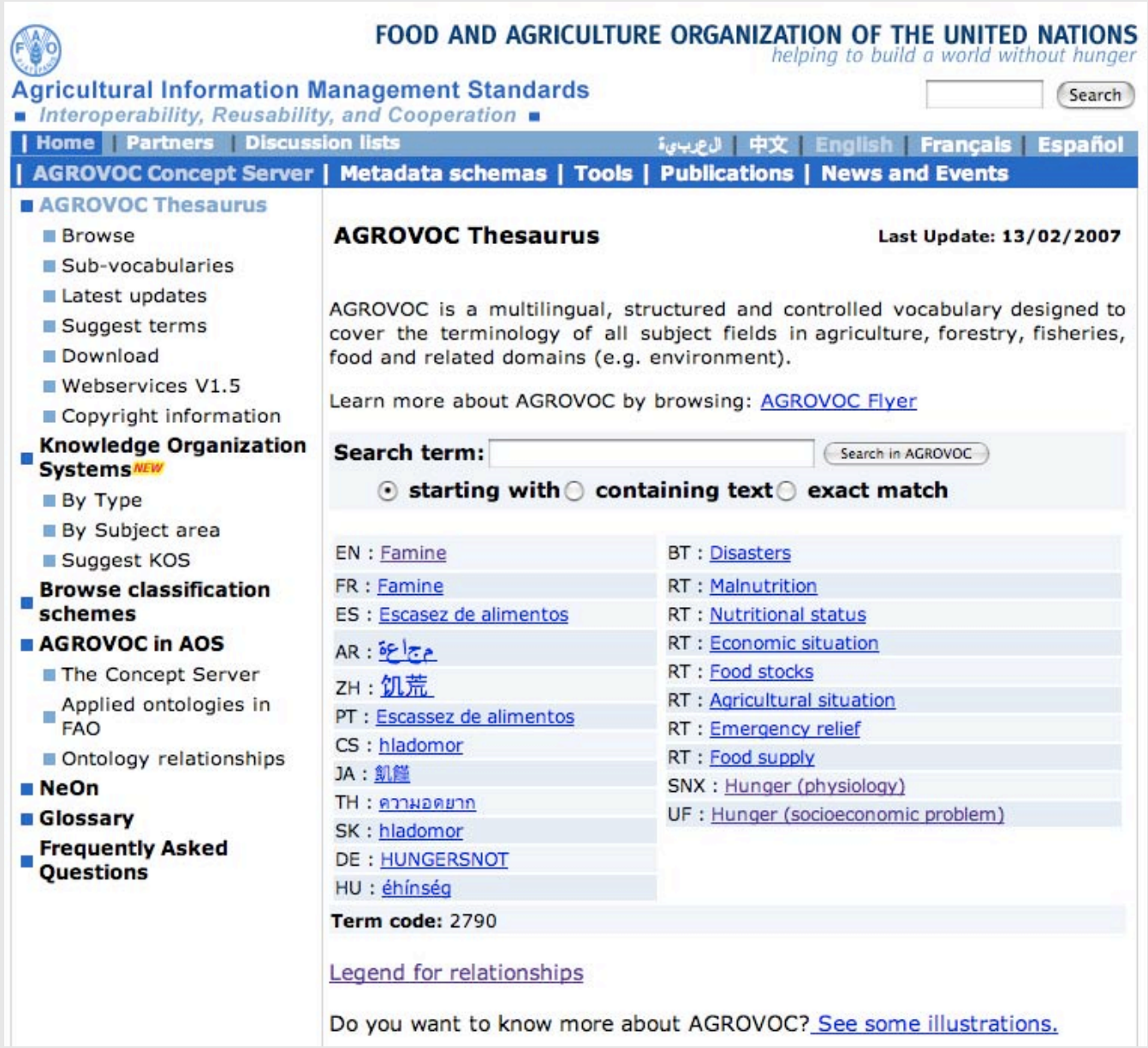
Susan really WISHES that you 'd listen to her.

FEs:

Core:

Event [Evnt] **Semantic Type** State_of_affairs	The change that the Experiencer would like to see.
Experiencer [Exp] **Semantic Type** Sentient	The Experiencer is the person (or sentient being) who wishes for the Event to occur.
Focal_participant [Foc]	This is the entity that the Experiencer wishes to be affected by some Event.
Location_of_Event [PEv]	The Location_of_Event is the place involved in the desired Event. I WANT that box on top of the other one. The prince WISHES you here before matins.

Thesauri: Agrovoc



FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS
helping to build a world without hunger

Agricultural Information Management Standards
Interoperability, Reusability, and Cooperation

Search

Home | Partners | Discussion lists | العربية | 中文 | English | Français | Español

AGROVOC Concept Server | Metadata schemas | Tools | Publications | News and Events

- AGROVOC Thesaurus
  - Browse
  - Sub-vocabularies
  - Latest updates
  - Suggest terms
  - Download
  - Webservices V1.5
  - Copyright information
- Knowledge Organization Systems NEW
  - By Type
  - By Subject area
  - Suggest KOS
- Browse classification schemes
- AGROVOC in AOS
  - The Concept Server
  - Applied ontologies in FAO
  - Ontology relationships
- NeOn
- Glossary
- Frequently Asked Questions

AGROVOC Thesaurus — **Last Update: 13/02/2007**

AGROVOC is a multilingual, structured and controlled vocabulary designed to cover the terminology of all subject fields in agriculture, forestry, fisheries, food and related domains (e.g. environment).

Learn more about AGROVOC by browsing: AGROVOC Flyer

Search term: Search in AGROVOC

starting with **containing text** **exact match**

EN : Famine	BT : Disasters
FR : Famine	RT : Malnutrition
ES : Escasez de alimentos	RT : Nutritional status
AR : مجاعة	RT : Economic situation
ZH : 饥荒	RT : Food stocks
PT : Escassez de alimentos	RT : Agricultural situation
CS : hladomor	RT : Emergency relief
JA : 飢饉	RT : Food supply
TH : ความอดยาก	SNX : Hunger (physiology)
SK : hladomor	UF : Hunger (socioeconomic problem)
DE : HUNGERSNOT	
HU : éhínség	

Term code: 2790

Legend for relationships

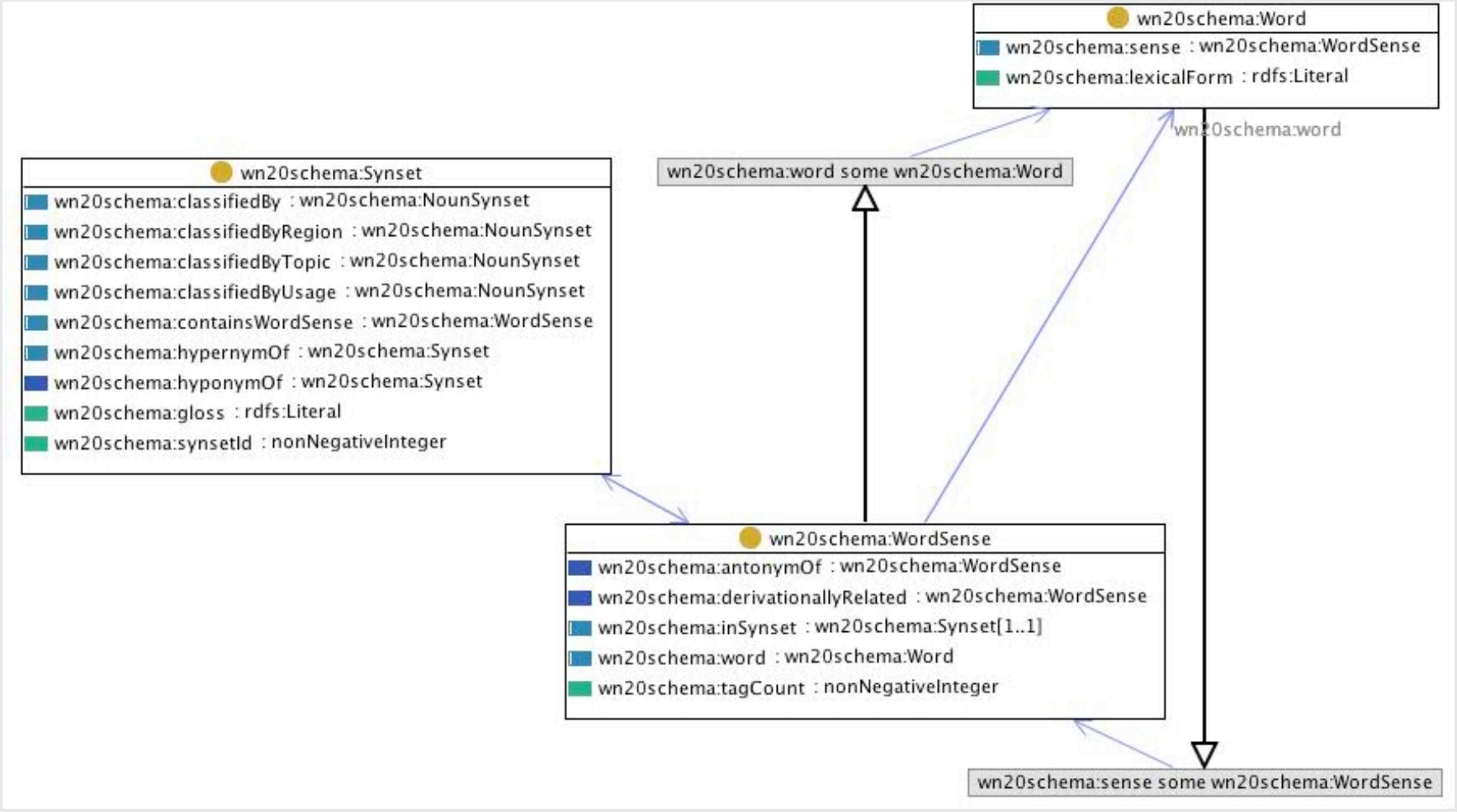
Do you want to know more about AGROVOC? See some illustrations.

How to compare them?

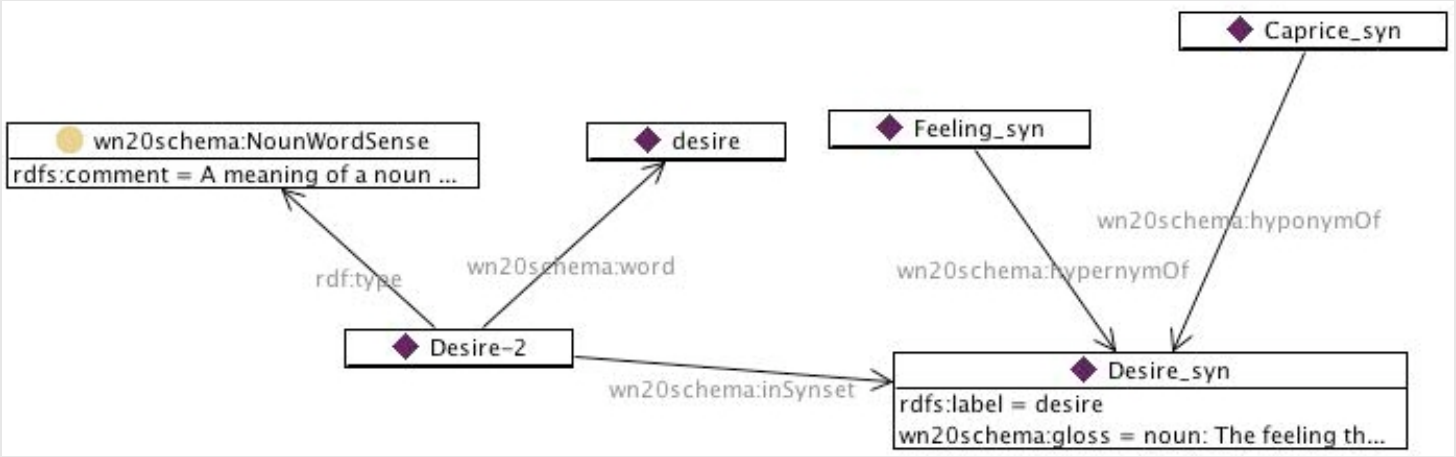
- How to compare descriptions that use different representations and are mostly missing formal semantics?
  - e.g. logic, informal, linguistic, topic-based, ...
- Current trend: meta-modeling
  - e.g. LMM for OWL, WordNet, FrameNet, KOS, LMF

W3C WNET Schema (Lexicon2ABox approach)

- http://www.w3.org/2006/03/wn/wn20/
- http://www.w3.org/2001/sw/BestPractices/WNET/wn-conversion.html#primer

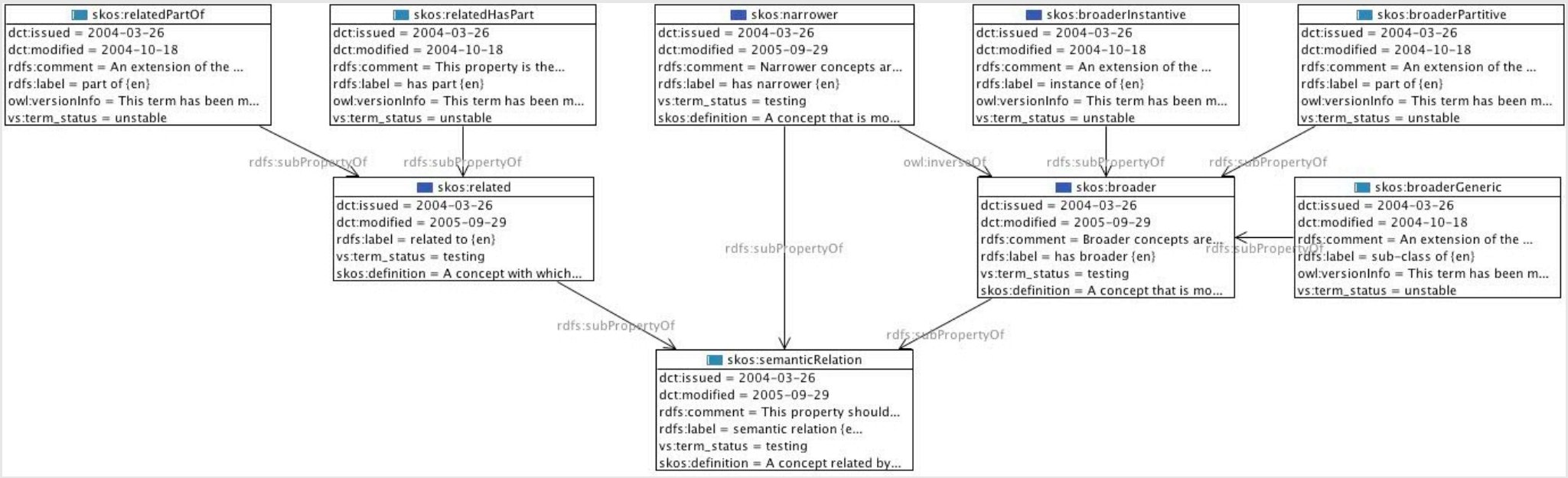


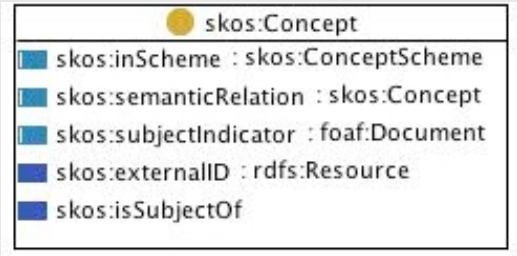
An example of wordnet in owl (Lexicon2ABox approach)

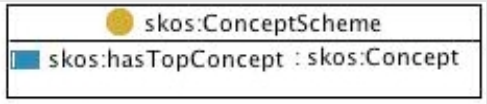


SKOS Vocabulary (KOS2ABox approach)

- http://www.w3.org/2004/02/skos/vocabs

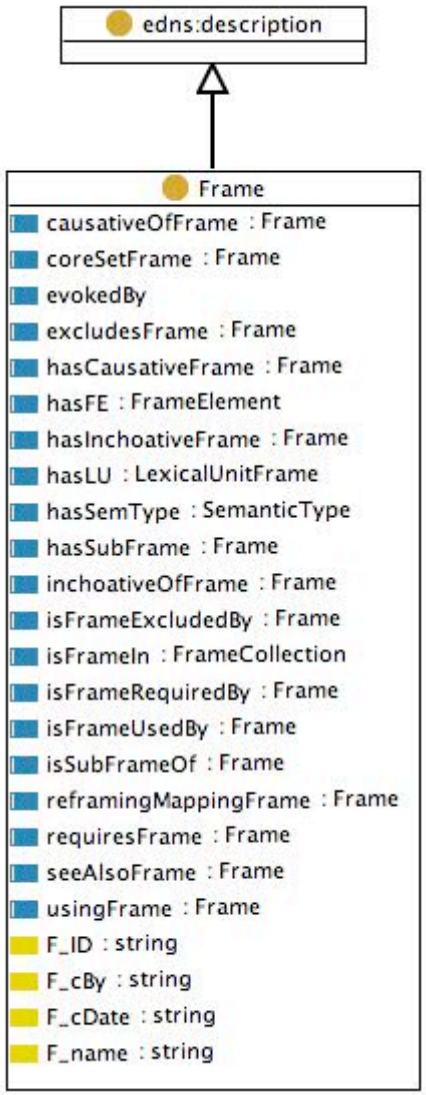






OntoFrameNet (a different Lexicon2ABox approach)

- http://www.loa-cnr.it/ontologies/FrameNet/ofn.owl
- based on cDnS ontology: http://www.loa-cnr.it/ontologies/OFN.owl



From “raw” data to patterns

- Moving from “raw” knowledge resources to networked ontologies require:
  - Ontology requirement analysis (domain(s), task(s), and sustainability constraints for ontologies to be built/managed)
  - Tool/resource requirement analysis (functionalities to be covered by tools, and competences needed)
  - Project planning (deciding on knowledge resources, economic resources, team composition and responsibilities, data copyright management, tools)
  - Workflow decision (specially for **reengineering** and **argumentation**)
  - Rationale elicitation (“critiquing” the reengineered data)
  - Providing solutions (e.g. based on design patterns, or conveying new ones)
- Not one, “best” waterfall methodology
  - A project can start spontaneously to solve a rationale elicitation problem, can be planned in order to reengineer knowledge resources, or to reuse existing ontologies or patterns, etc.
  - A project can be started either with or without requirement analyses.
  - Even the solutions can consist only of a “bulk” reengineering process, without explicit patterns

Legacy *aquaculture* hierarchies from fishery terminology systems

AQUACULTURE (AGROVOC)
- NT1 fish culture
  - NT2 fish feeding
- NT1 frog culture
- …
- rt agripisciculture
- rt aquaculture equipment
- …
- Fr aquaculture
- Es acuicultura

AQUACULTURE (ASFA)
- NT Brackishwater aquaculture
- NT Freshwater aquaculture
- NT Marine aquaculture
- rt Aquaculture development
- rt Aquaculture economics
- rt Aquaculture engineering
- rt Aquaculture facilities

Biological entity (FIGIS)
- Taxonomic entity
- Major group
- Order
- Family
- Genus
- Species
- Capture species (filter)
- Aquaculture species (filter)
- Production species (filter)
- Tuna atlas spec

SUBJECT (OneFish)
- Aquaculture
- Aquaculture development
- Aquaculture economics @
- Aquaculture planning

Sample data model analysis/conversion (KOS2TBox approach)

Term ≠ Concept

Term = String (or Lexical Item)

Concept = Class

BT ≈ subsumption between classes

RT ≈ top-level conceptual relation

{Descriptors} = $\bigcup${Classes},{Individuals}

Individual $\in$ Class

Concept ≠ Subject/Topic/Domain

Conversion: effects on translation (1)

- agrovoc_schema:Descriptor
  - agrovoc:River
  - agrovoc:Amazon

- owl:Class(agrovoc:River)
- owl:Individual(agrovoc:Amazon(rdf:type agrovoc:River))

Conversion: effects on translation (2)

- agrovoc:RT
- agrovoc_schema:Descriptor
  - agrovoc:Fishing_vessel
  - agrovoc:Fishing_gear
  - agrovoc:Fishing_vessel,RT,Fishing_gear

- Class(agrovoc:Fishing_vessel partial
  (restriction(agrovoc:RT
    someValuesFrom(agrovoc:Fishing_gear))))

Ontology evaluation

- Domain: entity types, expertise patterns
  - is the ontology appropriate to context?
- Task: competency questions
  - is the ontology appropriate to support relevant queries?
- Resources: tools and personnel
  - is the ontology (structure, function, annotations) manageable and cost-effective?
- Direct measuring of graphs and annotations
- Black-box/glass-box measuring of admissibility wrt conceptualization
- Indirect measuring via user feedback, and correlation
- Principles, diagnosis and trade-offs

Possible theses and research issues at our Lab

- Implementation of reengineering patterns over information resources and linking open data
- Implementation of tools supporting pattern-based design
- Collaborative tools for ontology design
- ...

Events: conferences, tutorials, schools

- EKAW2008: 29/9 - 3/10
- SWAP2008: 15/12 - 17/12
- SSSW2009: Early July 2009