



A Pattern Science for the Semantic Web

Aldo Gangemi

Semantic Technology Lab

Institute of Cognitive Science and Technology

CNR, Rome

aldo.gangemi@cnr.it

Thanks to: Valentina Presutti, Eva Blomqvist, Alfio Gliozzo

What kind of talk to give as a dinner keynote?



What kind of talk to give as a dinner keynote?

- Seriously funny?
 - Uhm ... better for a dinner speech

What kind of talk to give as a dinner keynote?

- Seriously funny?
 - Uhm ... better for a dinner speech
- Serious with lots of technical details?
 - Enzymatic actions in your stomach can suffer from a severe blockage

What kind of talk to give as a dinner keynote?

- Seriously funny?
 - Uhm ... better for a dinner speech
- Serious with lots of technical details?
 - Enzymatic actions in your stomach can suffer from a severe blockage
- A mix?
 - How can you mix serious things with jokes?
 - We already have Dieter who pretends to do it well

So what?

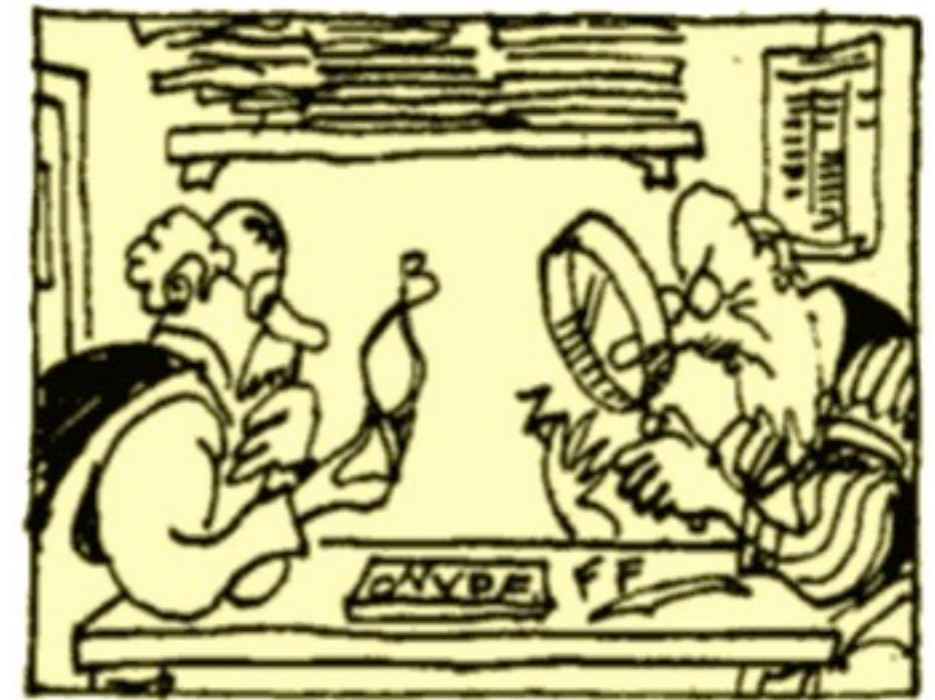
- I decided to tell you what I want to tell you
- Some visionary, a little bit crazy things, which one can tell only in a keynote
- A report of stuff done
- A final surprise

Outline

- A pattern science for the SW
- A scenario
- Frames and knowledge patterns
- Ontology design patterns
- Experimental evidence
- Tools and initiatives

The objects of an empirical science

- Should be shared by the community
- They can change
- A (partly) shared language (not necessarily formal) describes them
- Patterns typically emerge and change
 - E.g. paradigm shift from alchemy to chemistry (Lavoisier)
 - E.g. paradigm breaking in patient safety (Semmelweis)
 - Assessment and experimental practices



A science of patterns

- Ontology design for the Semantic Web *can* be practiced as an empirical science that has design patterns as its objects

Can is a claim: “yes we can!”

Patterns in general

- “Invariances across observed data or objects”
- They exist in natural, social, cognitive, or abstract worlds
- Mathematical pattern science is about symbols, i.e. non-interpreted information objects
- Objects of knowledge engineering are interpreted (either formally or cognitively)
- Mutual support/dependencies

Signs of a knowledge pattern science?

- Are its objects and language eligible to testing and manipulation?
✓
- Possible conflicting descriptions?
✓
- Evidence emerging from analysis of data?
✓
- Possibility to experiment and evaluate alternative theories and methods?
✓
- Correspondence in the real (social) world?
✓

Introducing Qfwfq



He's an Analogic Network (Von Glitschka)

A simple ordinary task

- Qfwfq wants to arrange a trip to San Francisco
- Qfwfq goes there for the first time
- Qfwfq likes hotels in the city center, close to shops and clubs
- Qfwfq travels a lot, hence there are many hotels that he liked, and many others that he disliked
- Qfwfq want to go there in a period with no close hard deadlines at work
- Qfwfq wants to book a flight at a reasonable price

Semantic Web 2010

- In order to solve his task Qfwfq
 - searches forums and social networks for good hotels, shops and clubs in San Francisco
 - looks on trip advisor for those and other hotels in San Francisco
 - checks hotel prices
 - looks photo galleries and hotel facilities descriptions
 - looks at google maps to see where the hotels are located
 - compare such information with his previous experiences i.e. select some candidates
 - checks his calendar and select a couple of possible periods
 - checks flight prices in those periods
 - checks hotel availability in those periods
 - ...
 - Qfwfq uses his personal preferences and knowledge, based on personal past traveling experiences and compares them to what he can extract from the information he gathered from the web
- how long does it take?

Semantic Web 2020

- Qfwfq asks his agent three possible travel arrangements based on those constraints
- Qfwfq goes back to his business while Qfwfq's agent performs the task
 - Qfwfq's agent uses a semantic representation of the task, exploits its knowledge about past Qfwfq's traveling experiences, and compares them to the appropriate Semantic Web knowledge
- After a while, Qfwfq's agent returns three proposals
- Qfwfq checks such proposals and selects the preferred one
- How long does it take?

In other words

- The final goal of the Semantic Web is to drastically decrease the cognitive load of humans when performing some task
 - Humans can delegate the execution of such tasks to smart software agents
- The Semantic Web can do that by enabling the interpretation and consumption of data in the form of knowledge
- It's fairly clear how to represent knowledge in general
- It is far less clear how to shape knowledge that has to be represented for a certain task, i.e. *meaningfully*

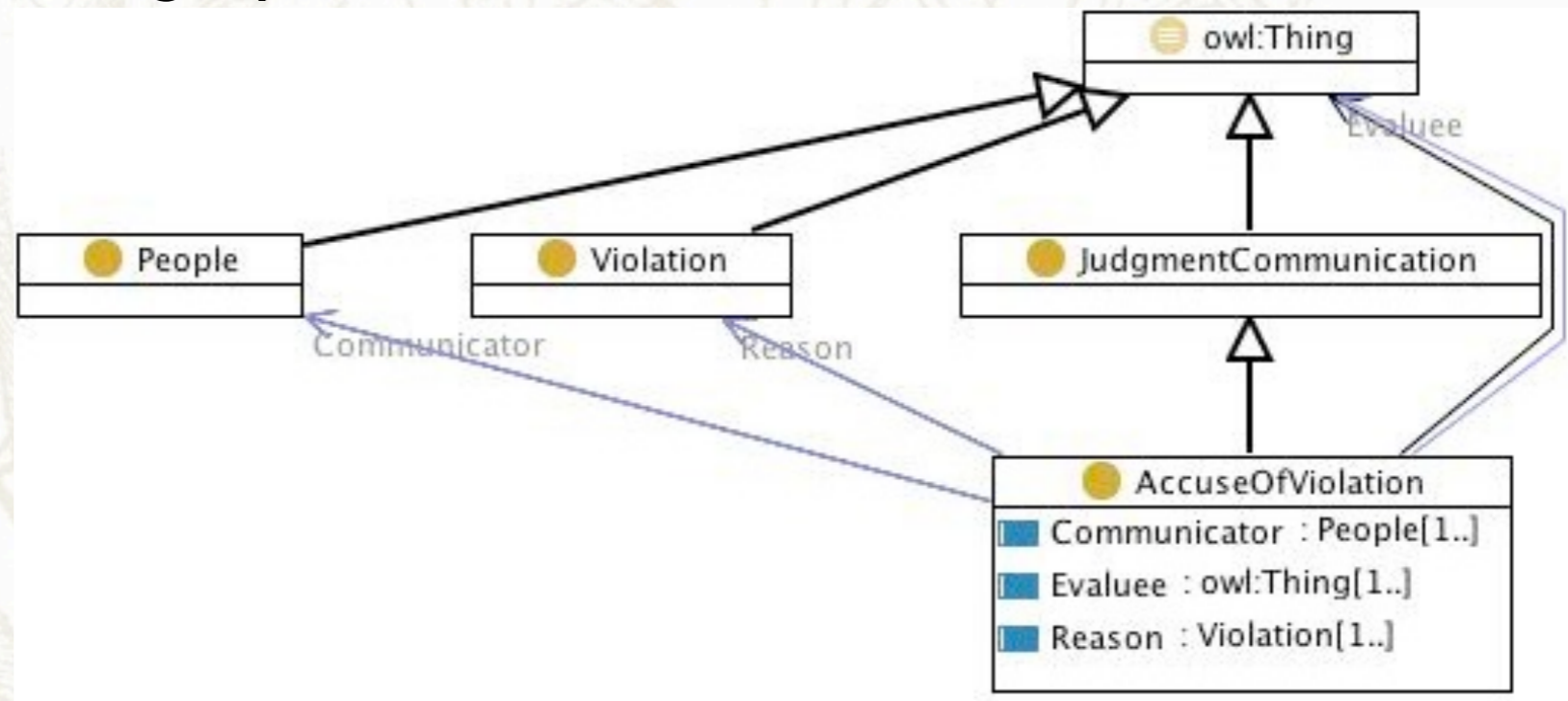
The frame hypothesis (1/2)

- Most realistic tasks require knowledge shaped in a way that allows to navigate through the data meaningfully
 - i.e. similarly to how a human interprets and aggregates them
- The typical shape of meaningful knowledge is that of a *frame* (or *situation*, *context*, *schema*)
- Minsky 1974:
 - *there would be large advantages in having mechanisms that could use these same structures both for thinking and for communicating*



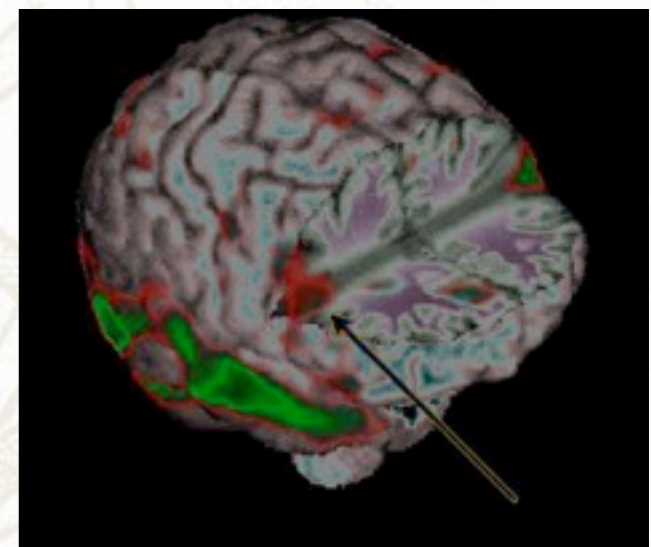
The frame hypothesis (2/2)

- The *unit of meaning* of the Semantic Web technologies should then be the *frame*, as opposed to scattered classes or properties (binary relations)
- Agents/reasoners must be able to recognize such *frames* and reason on them
- In ontology design, frames are called *knowledge patterns*, a special kind of *design patterns*



Cognitive foundations

- A search for the *relevant* units of meaning (not the primitives)
- Agents' understanding is based on [*patterns*] abstracted from previously occurred situations involving those agents, which are adapted and recombined on-the-fly in novel situations
 - Bartlett's experiments on schemata (1932), Neisser 1967
 - Piaget's experiments on schemata (1954)
 - Fillmore's frame semantics (1968)
 - Minsky's frames (1974)
 - Schank's scripts (1977)
 - Gibson's affordances (1977)
 - Biederman's experiments on scene recognition (1982)
 - Barsalou's ad-hoc goal-oriented categories (1983)
 - Lakoff's conceptual metaphors (1987), Langacker's compositional paths (1987)
 - Barsalou's simulators (1999)
 - Bar's associative-analogical network of frames for anticipation (2007)
 - Rizzolatti, Iacoboni, Gallese, etc. results and plausibility of mirror-neuron-system-based frames (2008)



A lexical frame

This frame deals with a **Healer** treating and curing an **Affliction** (the injuries, disease, or pain) of the **Patient**, sometimes also mentioning the use of a particular **Treatment** or **Medication**.

Medication [Med]
Semantic Type
 Physical_entity

The injected, applied, injected, etc. substance designed to cure the Patient.

He needs prolonged **TREATMENT** with **antibiotics**.

Note the tight relationship between Treatment and Medicine.

Healer [Hlr]
Semantic Type
 Sentient

The Healer, anyone who treats or cures the Patient, occurs as the External Argument of verbs:

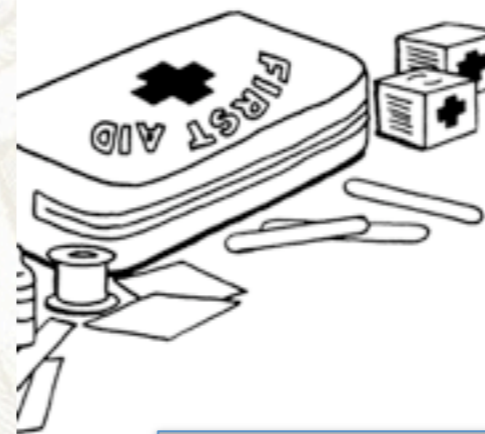
Doctors **ALLEVIATED** his suffering.

Cure

Healer



Patient



Medication

Does it really understand?

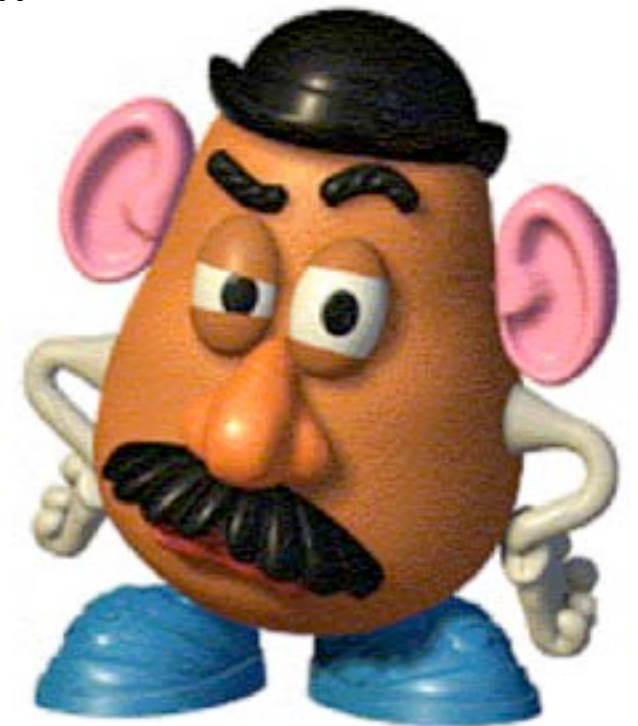
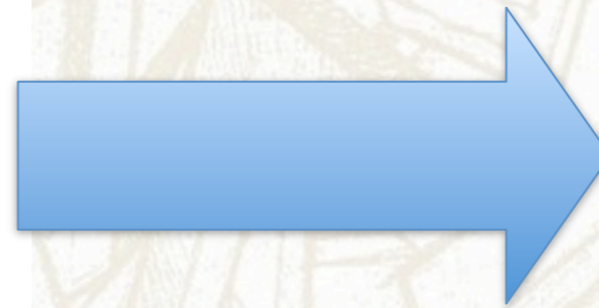


© guy & rodd, Dist. by UFS

The **driving** Frame

Driver: This is the being, typically human having a driving license, that controls the Vehicle as it moves.

Vehicle: This is the means of conveyance controlled by the Driver.



Frames we live by

- Frames and stories seem to drive our lives
 - Lakoff's *The Political Mind*, 2008
- Let's be aware of the frames and stories that we live by
- Some of them seem to be used to machine-washing our brains
 - maybe also linked data? ;)
- Some examples

Frames we live by (1/5)

- The *Godfather* frame



Frames we live by (2/5)

- The *Holy Inquisition* frame



Frames we live by (3/5)

- The *Logician* frame



Frames we live by (4/5)

- The *Pipeline* frame



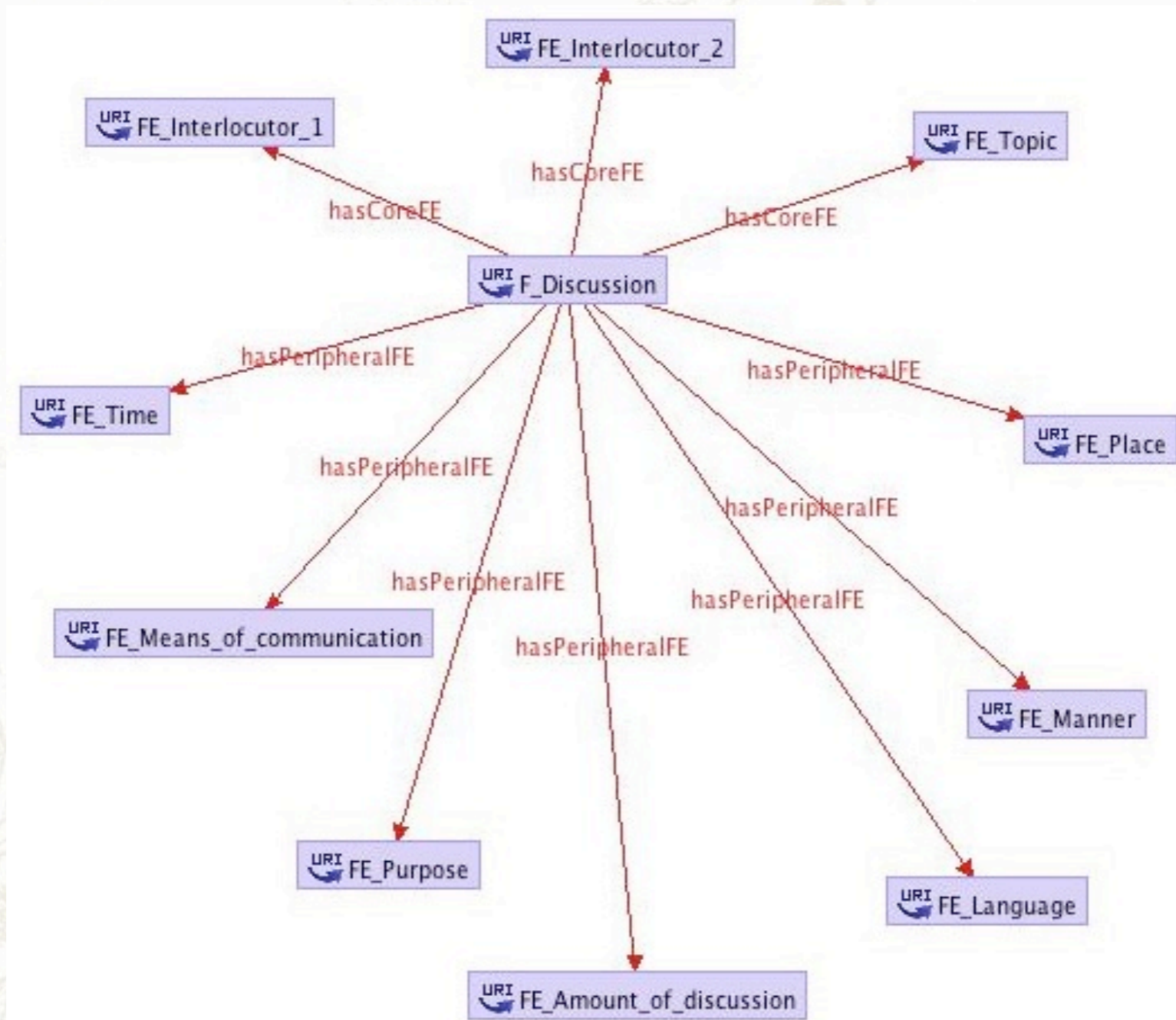
Frames we live by (5/5)

- The *Religious Libertine* frame
- ehm ...

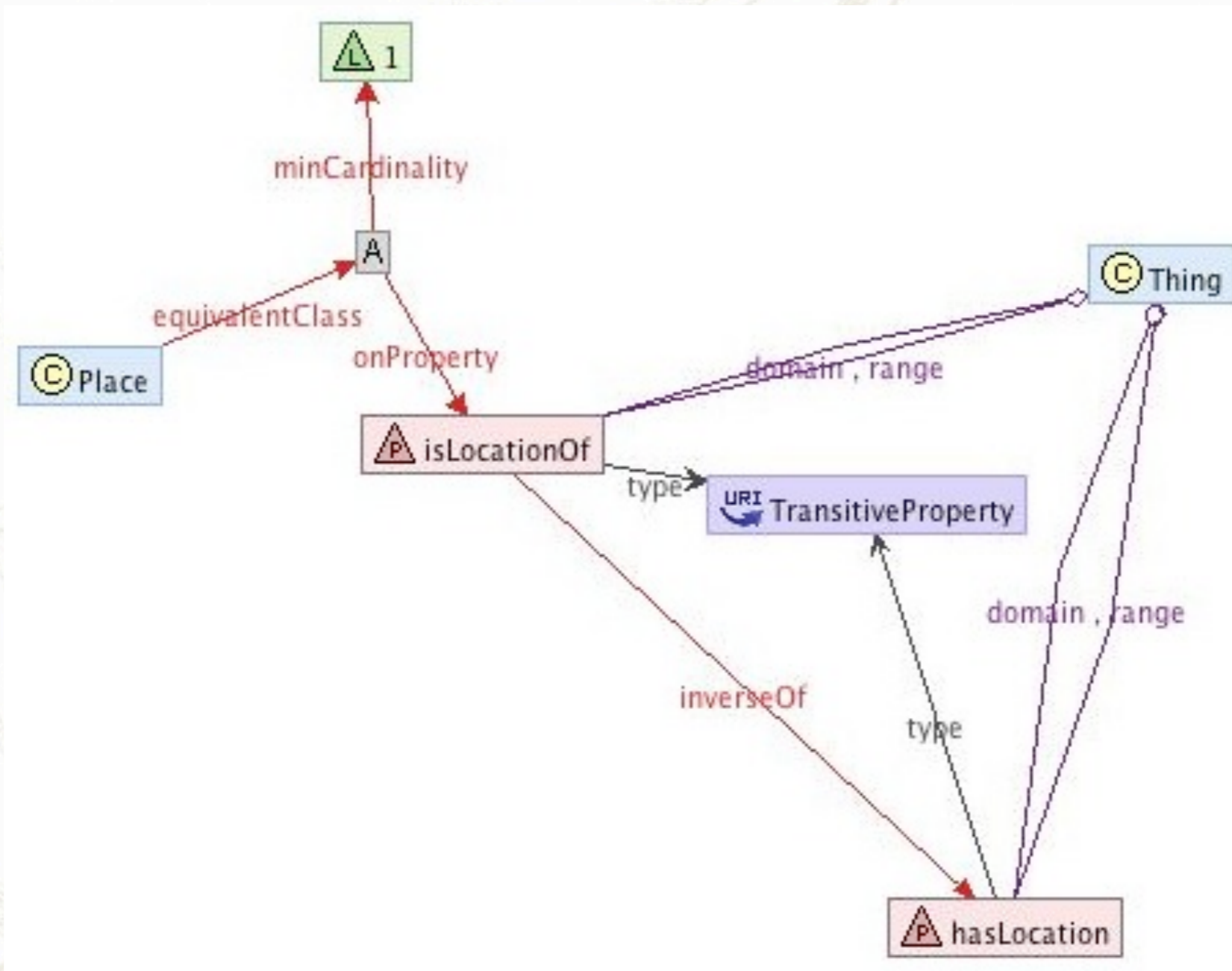
“of course, my confessions probably aren’t nearly as interesting as yours”



The *Discussion* frame from FrameNet

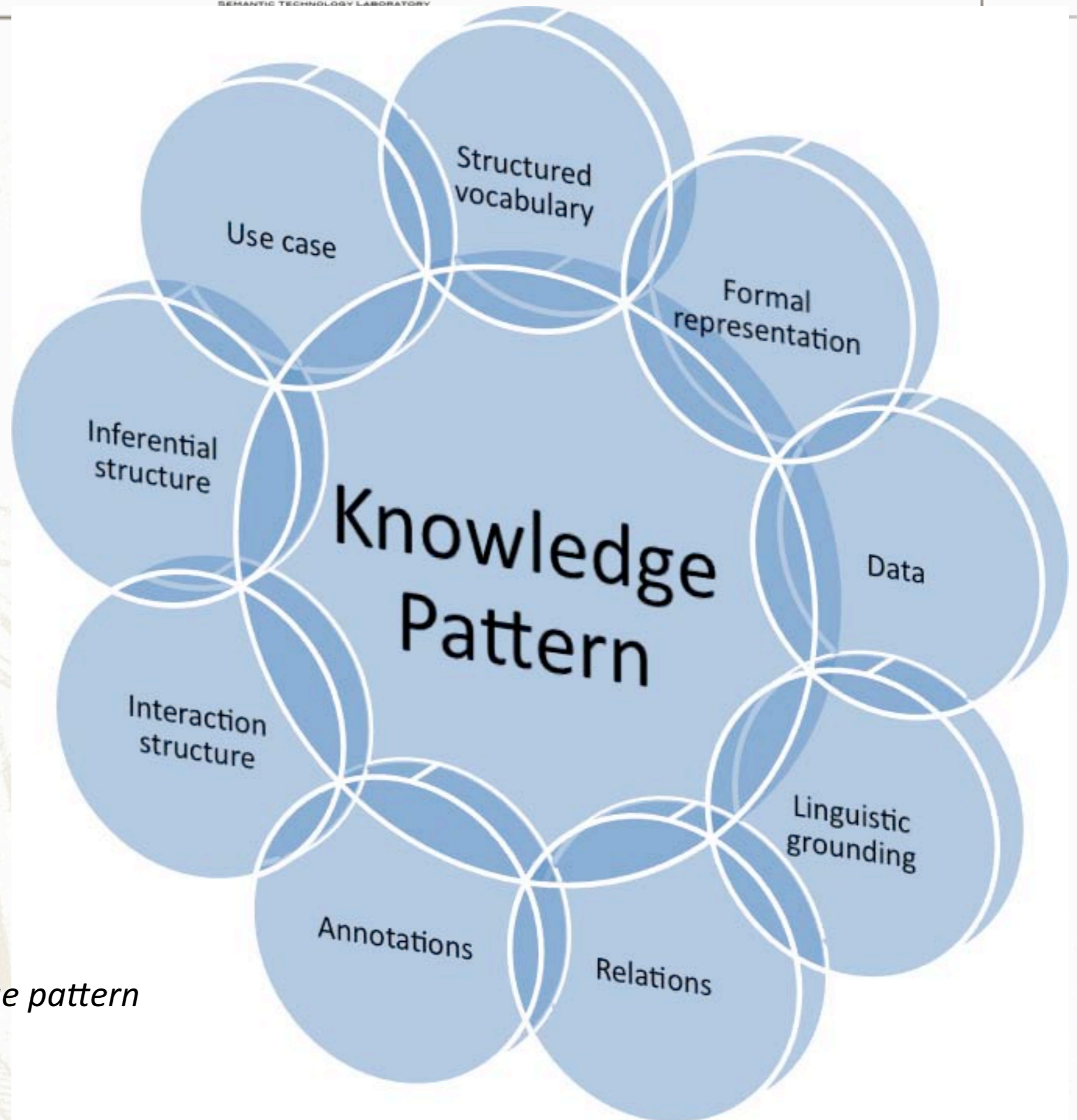


The *Place* content pattern



Sample frame-like structures

- Lexical frames
- Microformats, Infoboxes
- Query types
- Competency questions
- N-ary relations
- OWL/RDFS classes with proper (and locally complete?) sets of restrictions or properties
- Content ontology design patterns
- Noshir's structural signatures
- Data and conceptual models (if modularized according to requirements)
- Some HTML structures
- Some XML templates



The façades of a knowledge pattern

Practicality of knowledge patterns

- Easier ontology engineering
- Direct correspondence between requirements (problems), ontologies (solutions), and interaction (presentation)
- Advanced linked data consumption (exploration, discovery, lenses, matching, etc.)
- Easier schema and data reengineering and refactoring
- Reasoning modularization
- Query patterns
- Linguistic grounding of ontologies

Isn't KR enough?

- Knowledge representation has progressed a lot since the seventies
- Its original motivations, as e.g. described by Minsky and Brachman, have been drastically simplified
 - concentration on representation and complexity
- A lot of different KR languages can represent frames: DL, FL, FOL, CG, KM, etc.
 - but: which frames should be used, for what purpose, and with what pros and cons
- We need some *design*



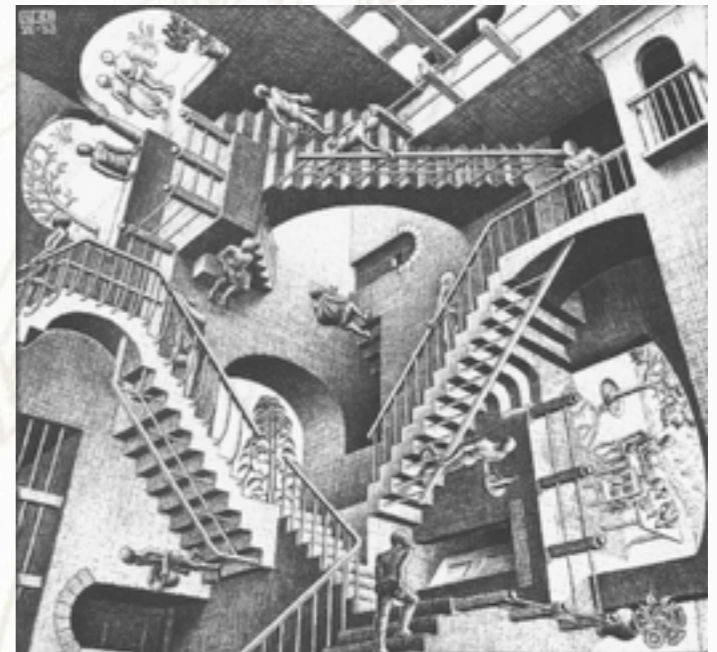
Design patterns

- Christopher Alexander, then Gang of Four
- Architecture ... Software engineering ... Knowledge engineering
- Focus on natural evolution of good practices
- Rich vocabulary to *talk* about problem-solving dynamics in actual people doing applied stuff
- Quite strict limits on the size of the problem-solution pair



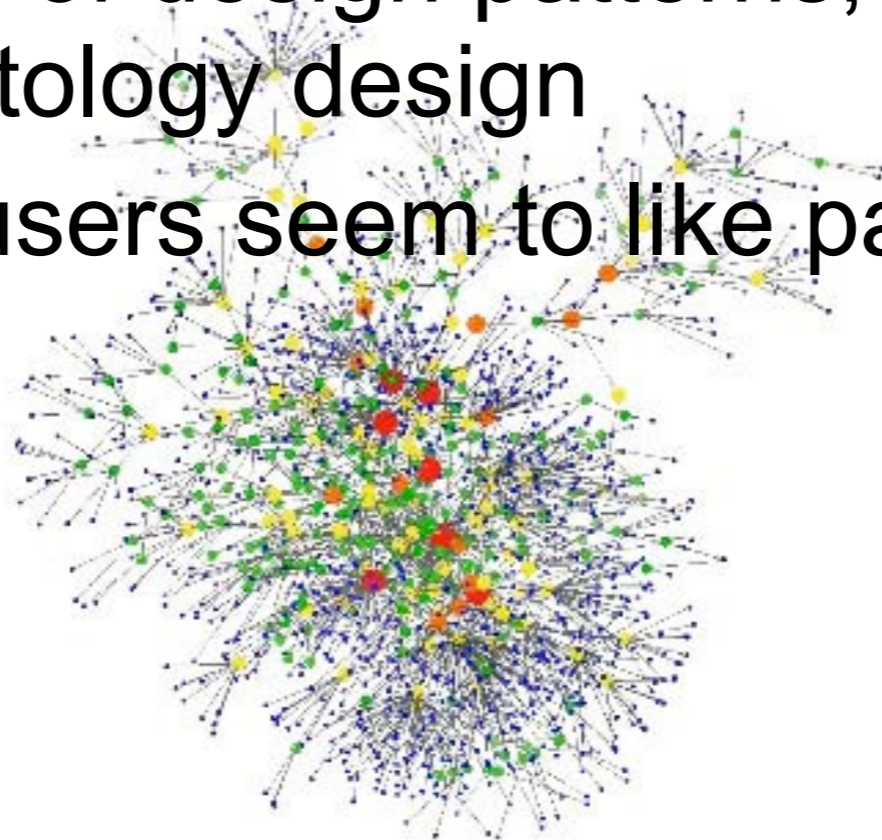
Knowledge patterns

- Clark & Porter 2000
- Address one aspect (i.e. schematic) of the knowledge engineering domain
- Reusable axiom schemata, signatures can be morphed while keeping the logical and reasoning properties
- Implemented in a language (KM) that needs (at least) full-fledged OWL2 and RIF to be reengineered
- Related to *competency questions* (Grüninger & Fox)



Ontology patterns

- Staab, Svatek, Gangemi, Rector, ...
- Put together the schematic and pragmatic approaches to design patterns
 - Good practices, competency questions, unit tests
 - Several types of design patterns, addressing most aspects of ontology design
 - Good news: users seem to like patterns; quality improves





Ontology Design Patterns . org (ODP)

OntologyDesignPatterns.org is a Semantic Web portal dedicated to ontology design patterns (OPs). The portal started under the NeOn project that still partly supports its development (<http://www.neon-project.org>).



navigation

- Main page
- Modeling Issues
- Reviews
- Feedback
- Domains
- Training
- Events

help

- Request an ODP account
- How to post a Content OP

contribute

- Post a Content OP
- Add a Domain
- Post your Feedback
- Post a modeling issue
- List of Categories
- List of Properties

catalogues

Contribute...

- Submit Pattern
- Post Review About a Pattern
- Add Domain of Interest
- Post Modeling Issue
- Post Your Feedback
- Request Account

View...

- List of Patterns
- Reviews
- List of Domains
- Modeling Issues
- Users' Feedback

More information...

- How to Post a Pattern
- Training Area
- Events
- Partners
- List of Categories
- List of Properties

Latest ODP News!

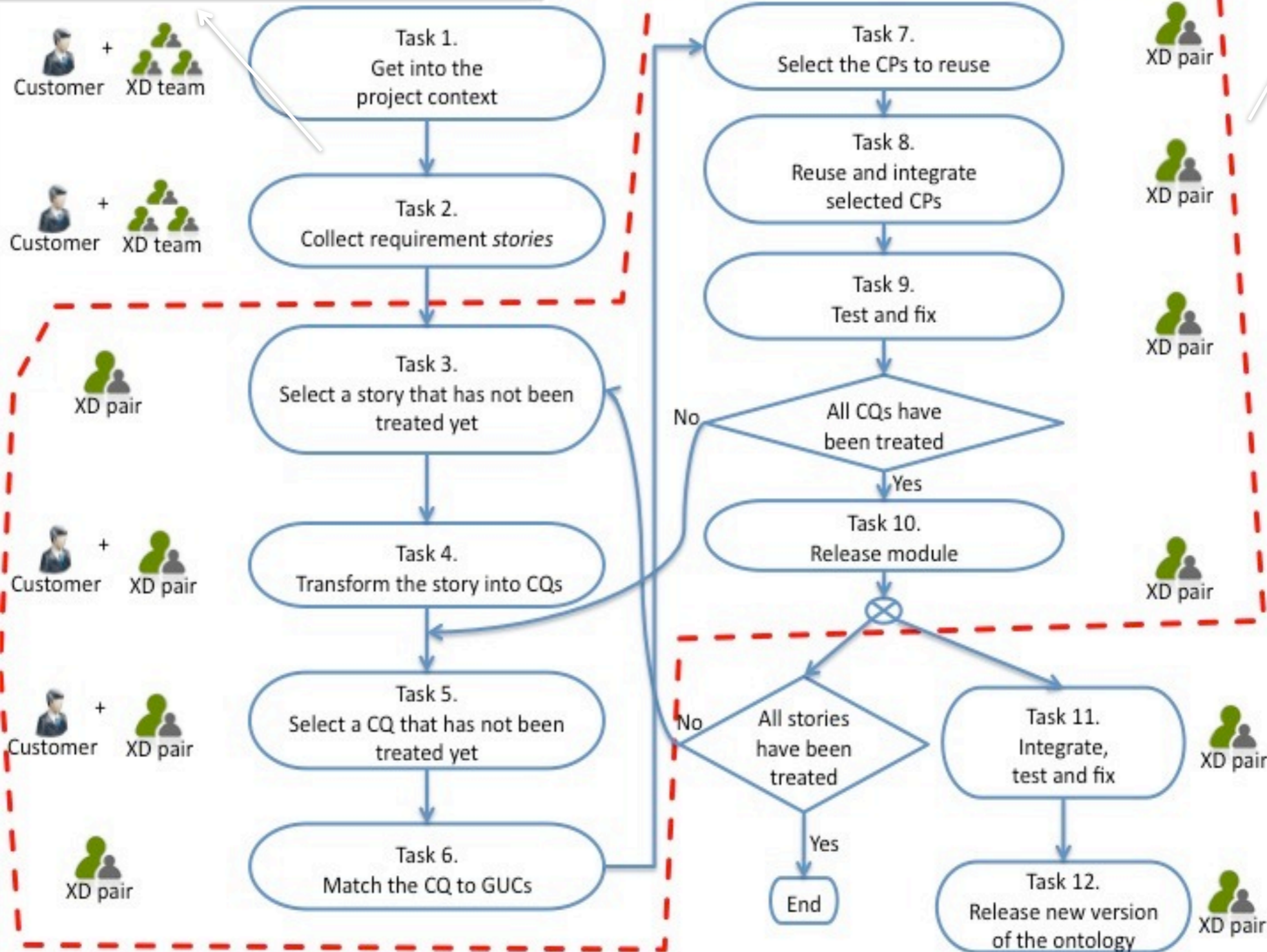
- 2 April 2009 10:10:52 The Loreley of Ontology Design Patterns [↗](#) (by VioletaDamjanovic)
- 21 October 2008 12:12:59 EvalWF has been released [↗](#) (by EnricoDaga)
- 5 June 2008 11:11:54 News at ODP portal! [↗](#) (by EnricoDaga)

There is much more... Check out the [about ODP](#) page for more information on the portal content and structure.



If you have no idea what we are talking about, visit the "What is a pattern?" page.

Title	Tuna observation
Depends on	Exploitation values, Tuna areas
Description	In 2004 the resource of species "Tuna" in water area 24 was observed to be fully exploited in the tropical zone at pelagic depth.
Priority	High



eXtreme Design - NeOn Toolkit - C:\Programmi\NeOn\NeOnToolkit-2.3.1-B264-base-win32\NeOnToolkit\workspace

File Edit Navigate Search Project Window Help

Ontology Navigator

- cpannotationschema.owl
- objectrole.owl
- http://www.NewOnto1.org/ontology1
 - Classes
 - <http://www.NewOnto1.org/ontology1/class250 (class)
 - <http://www.NewOnto1.org/ontology1/class328 (class)
 - <http://www.NewOnto1.org/ontology1/class250 (class)
 - <http://www.NewOnto1.org/ontology1/class140 (class)
 - Object Properties
 - <http://www.NewOnto1.org/ontology1/prop734 (object property)
 - Data Properties
 - Annotation Properties

Entity Properties

OWL Ontology

URI: http://www.NewOnto1.org/ontology1

Location: file:/C:/Programmi/NeOn/NeOnToolkit-2.3.1-B264-base-win32/NeOnToolkit/workspace/O

Imports

http://www.ontologydesignpatterns.org/cp/owl/classification.owl Remove

Add to imports:

ontology to import Add

Imports and Namespaces

Prefix	URI	Edit	Remove
owl	http://www.w3.org/2002/07/owl#	Edit	Remove
xsd	http://www.w3.org/2001/XMLSchema#	Edit	Remove
swrl	http://www.w3.org/2003/11/swrl#	Edit	Remove
swrlb	http://www.w3.org/2003/11/swrlb#	Edit	Remove

[OntologyProject] http://www.NewOnto1.org/ontology1

Project: OntologyProject

Ontology: http://www.NewOnto1.org/ontology1

Time: Mon Feb 22 12:39:24 CET 2010

Description	About	Message
Domain or range intersection (1)	All locally defined property ...	Domain or range of a property cor
Domain or range intersection	The relation ontology1:pro...	You defined an intersection of cla
Missing label (5)	This analyzer applies to all l...	All entities should have at least o
Missing label	ontology1:class328 (class)	The class ontology1:dass328 nee
Missing label	ontology1:class250 (class)	The class ontology1:dass250 nee
Missing label	ontology1:class140 (class)	The class ontology1:dass140 nee
Missing label	ontology1:prop734 (object...	The object property ontology1:pr
Missing label	<http://www.NewOnto1.o...	The ontology itself needs at least
Missing inverse property	This analyzer applies to all l...	Each object property should have
Missing comment (5)	This analyzer applies to all l...	All entities should have at least o
Missing comment	ontology1:class328 (class)	The class ontology1:dass328 nee
Missing comment	ontology1:class140 (class)	The class ontology1:dass140 nee
Missing comment	ontology1:prop734 (object...	The object property ontology1:pr
Missing comment	ontology1:class250 (class)	The class ontology1:dass250 nee
Missing comment	<http://www.NewOnto1.o...	The ontology itself needs at least
Missing domains or ranges (1)	This analyzer applies to all l...	Each property should have its don
Missing domains or ranges	ontology1:prop734 (object...	Range for property ontology1:pr
Architectural import notice (0)	This analyzer applies to th...	Most of the locally defined entities
Missing type (0)	This analyzer applies to all l...	Each entity must be the instance c
Isolated entity (0)	This analyzer applies to all l...	Each entity must be related at lea
Unused imported ontology (0)	This analyzer applies to th...	All imported ontologies should hav

ODP Registry

Content ODP Submissions

- Submissions
 - AgentRole
 - Biological_Entities
 - Classification
 - Co-participation
 - CollectionEntity
 - Componency
 - Constituency
 - Description
 - GO_Top
 - HasPest
 - Information_realization
 - Invoice
 - Metonymy-species-commodity

ODP Details

URI: http://aims.fao.org/aos/v1.0/

Location: http://www.aims.fao.org/aos#hasPest

Property	Value	Lang
versionInfo	"20091021"	
comment	"Last revised by KCEW 20080508: removed all ..."	en
comment	"Last revised by KCEW 2008: set the correct n..."	en
comment	"Last revised by KCEW 20090730: added some..."	en
comment	"Last revised by KCEW 20090821: " / " URI inst..."	en
comment	"Last revised by KCEW 20090828: edited with ..."	en
comment	"Last revised by KCEW 20090924: changed na..."	en
comment	"Last revised by KCEW 20090929: added hasS..."	en
comment	"Last revised by KCEW 20091002: added all rel..."	en
comment	"Last revised by KCEW 20091021: completed r..."	en
comment	"TODO: remove chemical and geographical con..."	en

XD Selector

Indexing (7)

- /agentrole.owl [http://www.ontologydesignpatterns.org/cp/owl/agentrole.owl]
 - Physical URI: http://www.ontologydesignpatterns.org/cp/owl/agentrole.owl
 - Score: 0.41632402
- /objectrole.owl [http://www.ontologydesignpatterns.org/cp/owl/objectrole.owl]
- /taskexecution.owl [http://www.ontologydesignpatterns.org/cp/owl/taskexecution.owl]
- /taskrole.owl [http://www.ontologydesignpatterns.org/cp/owl/taskrole.owl]
- /classification.owl [http://www.ontologydesignpatterns.org/cp/owl/classification.owl]
- /participation.owl [http://www.ontologydesignpatterns.org/cp/owl/participation.owl]
- /timeinterval.owl [http://www.ontologydesignpatterns.org/cp/owl/timeinterval.owl]
- Latent indexing (37)
- /agentrole.owl [http://www.ontologydesignpatterns.org/cp/owl/agentrole.owl]
- /objectrole.owl [http://www.ontologydesignpatterns.org/cp/owl/objectrole.owl]
- /taskrole.owl [http://www.ontologydesignpatterns.org/cp/owl/taskrole.owl]

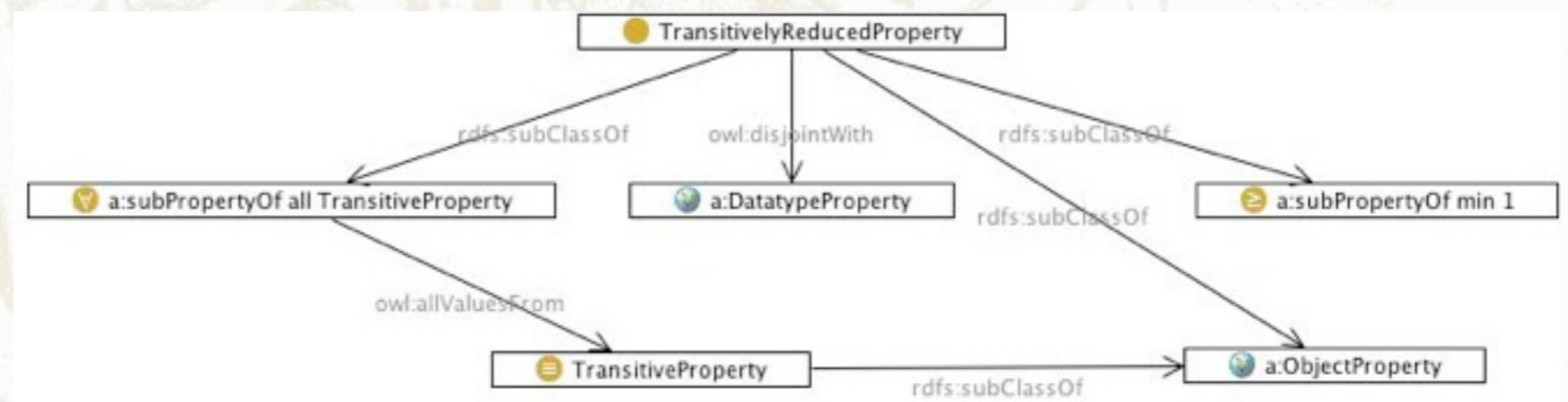
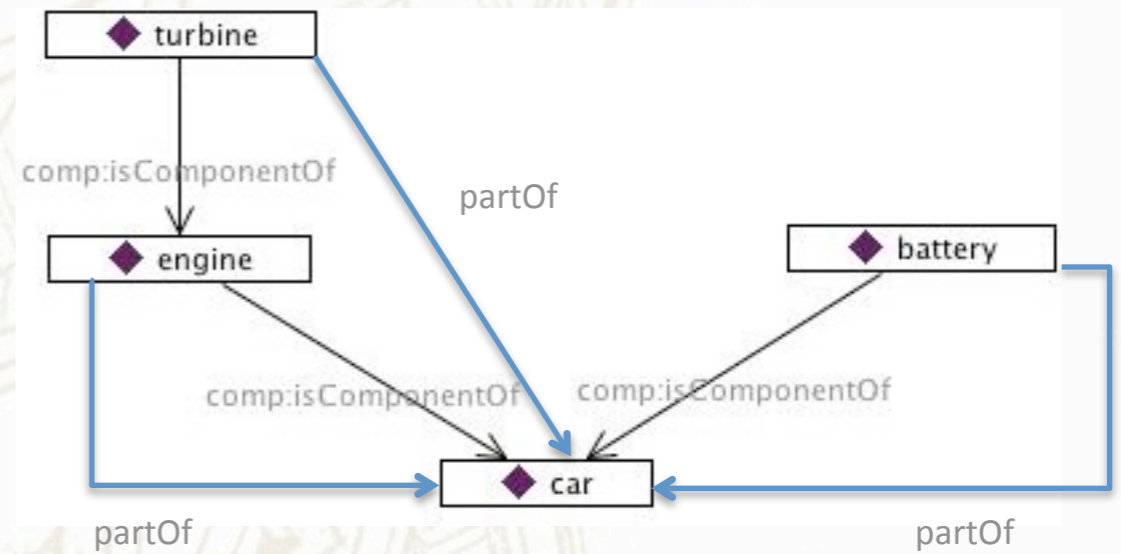
Specialize, compose, annotate
ODPs and ontologies

Analyze your ontology against
good practices and patterns

Browse, search, and get Content ODPs

Logical patterns

- Macros: subClassOf Class AND subClassOf Restriction
- N-ary relation reification
- Classes as values
- Exhaustive partition
- Transitive reduction
- ...



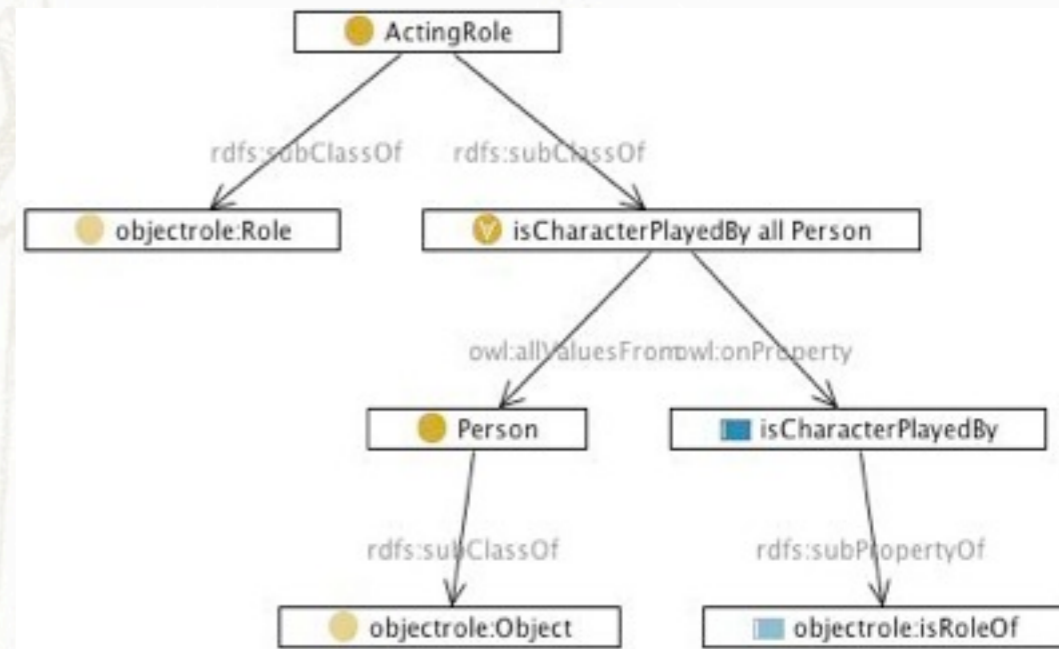
Reasoning patterns

- Materialize and query
- Classify and materialize
- Assert constructs, check consistency, and classify
- Learn disjointness and check consistency
- Make NER, populate, and classify



Content patterns

- Transitive part-of
- Object vs. Event
- Situation
- Judgment communication
- Structure, Function, Process (GO)
- Linnaean taxonomy
- Invoicing
- Resource abundance observation



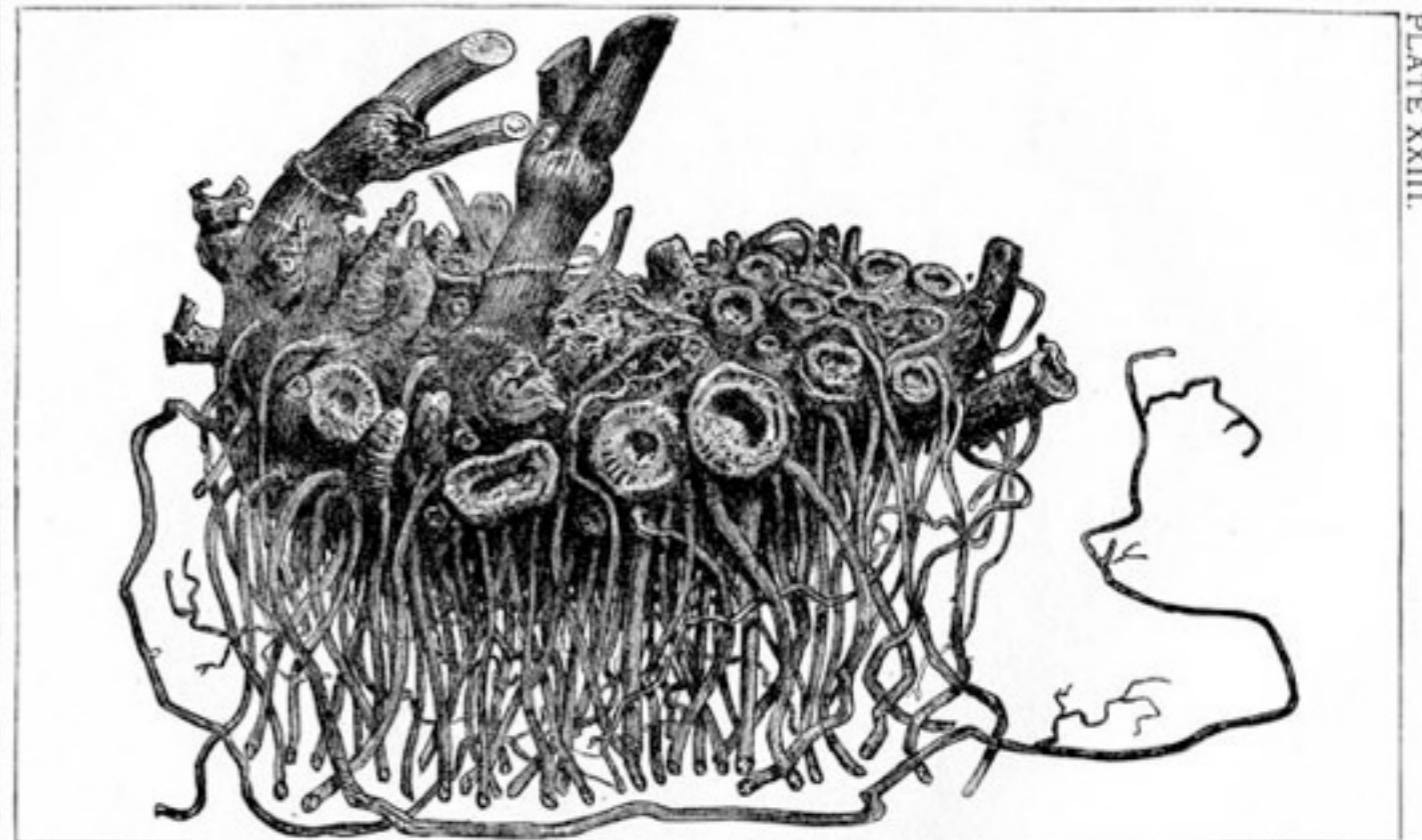
Reengineering patterns

- Thesaurus to SKOS
- Relational DB to OWL
- XML to RDF
- NER entities to ABox
- WordNet DB to OWL
- Microformat to RDF



Correspondence patterns

- *Class to Property+RangeRestriction*
- *DatatypeProperty+DatatypeRestriction to ObjectProperty+RangeRestriction*
- *Class to Individual*



A FRESH RHIZOME OF CIMICIFUGA RACEMOSA.
(Natural Size.)

Naming patterns

- Subclass name specializes superclass name
- Property prefixes nouns with *has**
- Capitalized class name
- Class name describes *one* of its instances
 - Antipattern: Plural class names

Architectural patterns

- DL *species/profiles*
- Recipes on exclusive structures to be used for a particular application
- Modular organization of ontology networks

Anti-patterns

- Partonomies as subsumption hierarchies
- Linguistic disjunction as class disjointness
- Linguistic conjunction as class disjunction
- Causality as entailment
- Expressions as instances of the class representing their meaning

Experimental evidence (I)

- Experiments conducted on master and PhD students show that the usage of content patterns improve the quality of ontologies (measured in terms of task-coverage and error-freedom), and the subjective perception of smooth and good design
- *KCAP09*

Experimental evidence (II)

- eXtreme Design methods applied on a test set of 7 design pairs and 1 integration pair
- Experiments on PhD students show that XD further improves quality and error-freedom, but also improves coverage of the proposed requirements
- *EKAW2010*

eXtreme Design camp in Bologna

- **Setting**

- After a PhD course (~ 30 hours training)
- 12 people – 6 pairs + 1 person acting as customer
- One day - around 6 hours of modeling
- 15 stories

- **Result**

- All stories addressed
- Four integration iterations
- The resulting ontology is fully documented and highly modularized
- Full documentation on the ODP wiki (training):

- **As additional to the “unconference” approach of VoCamp**

- The ontology is produced
- Documentation is available immediately
- The ontology is modular and well designed

Experimental evidence (III)

- OntoCase applied to Text2Onto ontology learning results makes ontologies better in terms of cohesion, consistency, etc.
- *ISWC2009*

Experimental evidence (IV)

- Frame detection and discovery has been performed based on a hybrid NLP-KE method
- FrameNet frames have been detected in a large corpus, they have been specialized with local corpus terminology, and represented in OWL, associating the corpus material as linguistic grounding
- *ESWC2009*

CommerceScenario

- Buyer: People_noun.group_Instance who read_verb.cognition this warning_noun.communication
- Target: buy_verb.possession
- Goods: these toys_noun.artifactInstance
- Recipient: for their children_noun.person_Concept **or** grandchildren_noun.person_Concept

Experimental evidence (V)

- Semantic Scouting
 - Organizational knowledge management project
 - Starting from XML templates as module/pattern drafts
 - Reengineering XML and scraped templates
 - Reengineering DB schemata
 - Obtained modular, pattern-based, task-based ontology
 - Query patterns from knowledge patterns: expert finding, etc.

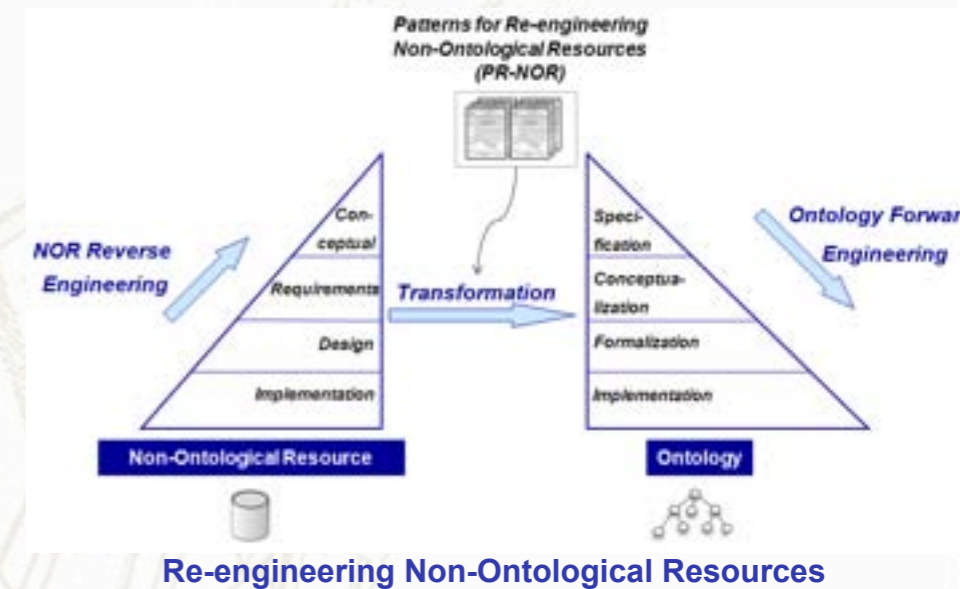


Methods, tools and initiatives

- STLab's **eXtreme Design** method
- INRIA's **correspondence pattern**-based alignment methods
- UPM's **NOR** and **lexico-syntactic patterns** methods
- Tools for pattern-based design
 - Vinay Chaudry's **AURA**
 - Luigi Iannone's Protégé4 **OPPL** plugin
 - STLab's NeOn Toolkit **XD** plugin
 - Eva Blomqvist's **OntoCase**
 - Vojtech Svatek's and François Scharffe's **PatOmat**
 - Clark&Porter's **KM**
 - **beInformed**
 - ...
- Workshop on Ontology Patterns (WOP) @ISWC2009, @ISWC2010
 - WOP2010 will include a VoCamp with XD support
- KIELD workshop @EKAW2010
 - Matching, extraction, enrichment, focused on linked data

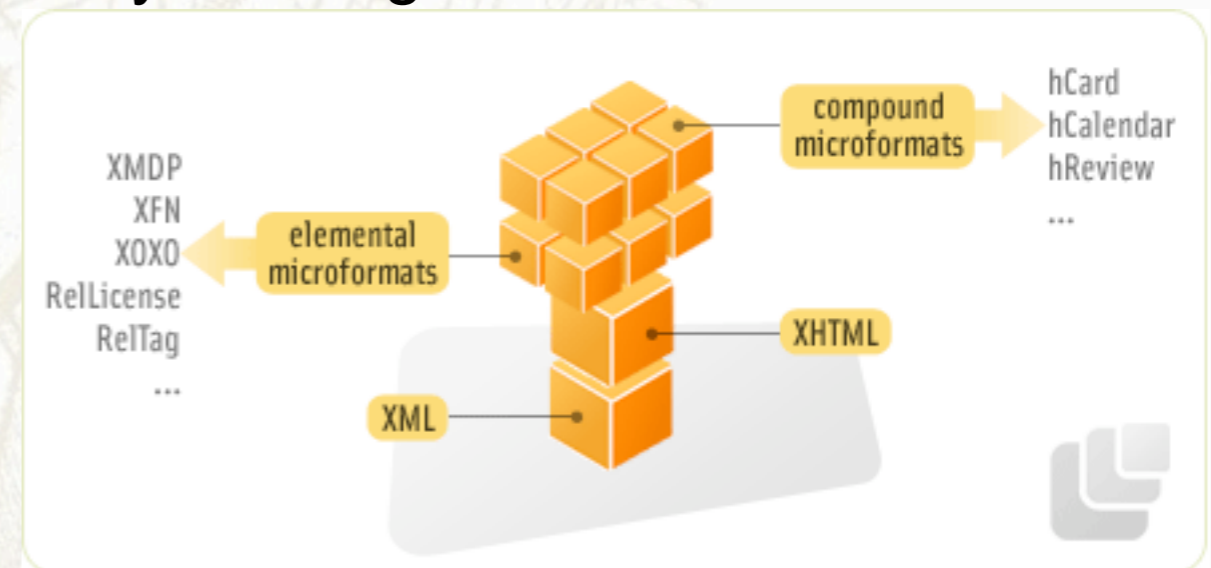
Methods, tools and initiatives

- STLab's **eXtreme Design** method
- INRIA's **correspondence pattern**-based alignment methods
- UPM's **NOR** and **lexico-syntactic patterns** methods
- Tools for pattern-based design
 - Vinay Chaudry's **AURA**
 - Luigi Iannone's Protégé4 **OPPL** plugin
 - STLab's NeOn Toolkit **XD** plugin
 - Eva Blomqvist's **OntoCase**
 - Vojtech Svatek's and François Scharffe's **PatOmat**
 - Clark&Porter's **KM**
 - **beInformed**
 - ...
- Workshop on Ontology Patterns (WOP) @ISWC2009, @ISWC2010
 - WOP2010 will include a VoCamp with XD support
- KIELD workshop @EKAW2010
 - Matching, extraction, enrichment, focused on linked data



Research directions

- Pattern discovery and detection
 - Integrate symbolic and knowledge patterns
 - Linked data pattern harvesting
 - The *boundary* problem
- Bootstrapping from reuse
 - Public, editable inventories and community building
 - Component library
 - Lexical frames
 - Data Modelling
 - XML schema inventories
 - Microformats
 - Recipe collections (cf. Good Relations, LOD design patterns)



Conclusions

- Semantic Web needs meaningful, relevant ontologies (vocabularies+data) for tasks and applications
- Relevance is associated with cognitive frames for actions, plans, expectations
- Knowledge patterns should be the target units of ontology design
- The ODP project grows significantly in that direction
- **A lot of work** to be done