

PatOMat project and  
beyond:  
can **patterns** support  
OWL's dorsal muscles  
to bear the **real world**?

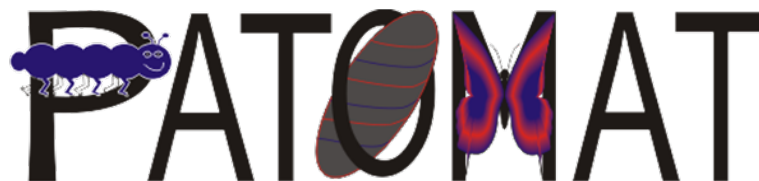


*Vojtěch Svátek, Ondřej Zamazal*  
University of Economics, Prague  
Knowledge Engineering Group

# Part I. PatOMat project

# PatOMat project

- 2010-2012, funded by CSF, <http://patomat.vse.cz>
- Developing a generic infrastructure for OWL ontology transformation
  - based on **transformation patterns**
  - includes **naming** transformation and use of **annotations**
  - primarily focused on ‘modelling style’ change
  - preserving the underlying real-world setting but typically not the logical model (and immediate outlook)



# Usage scenarios

- Generic style transformation for a different setting
  - e.g. OWL  $\Leftrightarrow$  SKOS
- Ontology matching
  - One ontology adapts its style to the other for easier match
- Ontology import (with adaptive merging)
  - Importing ontology adapts itself to the imported one, or vice versa
- Ontology complexity downgrading
  - e.g. to OWL-EL; typically most careful kind of transformation
- Ontology pattern education
  - e.g. demonstrating the transition between alternative logical patterns such as W3C CPV ‘approaches’

# Some external collaborations

- Infrastructure – *Manchester, Rome*
  - use of (OWL-API and) OPPL; XDTools framework
- Matching – *Mannheim, Montpellier*
  - cross-feeding with complex matchings / alignment patterns
- Import – *Rome*
  - special case of importing (ODP portal) content patterns
- Downgrading – *Mannheim*
  - OWL EL use case
- Naming patterns – *Freiburg*

# Implementation

- See <http://owl.vse.cz:8080>
- Self-standing *RESTful services* for
  - source pattern detection
  - transformation instructions generation
  - actual transformation
- *Java library*; used in Rome's *XDTTools* (Eclipse/NeOn)
- *Graphical editor* for transformation patterns (TPE)
- Very sparse *collection* of transformation patterns (cca 20)
- Prototype framework, not yet tested on large data
- *Demo can follow*

# PatOMat follow-up

- EU LOD2 (joined from September 2011)
  - part of WP3 “Knowledge Base Creation, Enrichment and Repair” (DL Learner, ORE...)
- CZ-SK LAAOS
  - „Logical aspects of adaptable ontological schemes“ – focus on compositionality of transformation patterns
- Possibly CSF postdoc project by Ondřej

# Part II. „Beyond“



# Is OWL sufficiently rich?

- Opinion 1: Definitely! We struggle to **devise reasoners** even for its simpler fragments!
- Opinion 2: It's very restricted. We need HOL, even FOL is not enough to **model the world!**
- Not much in between... ☹️

# Proposal

- Let's stick to **OWL**
  - Because it's tractable
  - **Because it's widely known and used!**  
(even if mostly erroneously)
- Let's find what deserves to be modelled and add it to the models... even if **out of sight for reasoners!**
  - **Annotations** and coherent **naming** are preserved during serialization; naming is (mostly) rendered at first place
- Ampler real-world model – available at least to human **users** and to processing **tools** (other than classical reasoners) with adequate capability

# Where to start from?

- The real world is primarily ‘about’ **individuals**
- But A-box modelling in OWL does not offer much discernment
  - 2 types of axioms... plus punning
- OWL patterns thus do not unambiguously match (even the important!) **real-world patterns**
- If we build T-boxes on the top of an incoherent U, reasoning may not return plausible results

# Web of Entities or Web of Data?

- **Entities** and relationships exist in the real world
- RDF/OWL is designed to work with them... but:
  - There are different ways (modelling styles) to express the same real-world setting of entities
  - The same way of modelling may correspond to multiple real-world settings
- Growing number of open and (sometimes) linked **datasets**
- Abstraction layer: vocabularies with often vague real-world semantics
- Style heterogeneity is thus obscured...

# ‘Onto-LD Gap’ Hypothesis

- The low adoption of ontologies on the WoD is not due to their complex T-box semantics (only)
  - (LD folks are curious and not dumb)
- ...but to the unclear mapping of datasets to the underlying real-world setting (WoE)
  - (This makes them believe that only the WoD works)
- Allowing the U to capture the common **real-world patterns** is needed in order to link the LD world to the ‘T-boxy’ world

# What is common enough to deserve (re-)modelling?

- Universals vs. particulars
- Entities vs. relationships
- Concrete individuals vs. collections
- Temporality – 3D vs. 4D entities?
- other?
- *...while not dragging the whole (exciting!) foundational ontology inventory back again*
- *... only what is intuitive and frequent*

# Material for empirical studies

- (Linked Open) Vocabularies
  - and associated datasets
  - practical impact, but harder to analyse
- Ontology Pattern Libraries
  - Manchester, Rome, W3C, Working Ontologist, ...
  - easy to analyse (well described), but may be contrived

# Universals vs. particulars

- Universals can be expressed via OWL classes... or OWL individuals
  - Examples: W3C VP, ProductOrServiceModel in GoodRelations, etc.
  - Often just a matter of guess when adopting an ontology!
  - *BTW the logic crowd uses such meta-modelling as reasoning workaround...*



# Entities vs. relationships

- Particular relationships can be expressed via OWL property assertions... or via a set of property assertions linked to the same OWL individual
  - Examples: W3C N-ary, Relationship in (FOAF) RV, etc.

# Concrete individuals vs. collections

- ‘Determinate’ collections can be represented via class expressions
- ‘Indeterminate’ collections can be only approximately modelled via existential restriction
  - Examples: ‘some lions’ in W3C CPV, SomeItems in GoodRelations, etc.
  - Explicit modelling desirable, different usage...

# 3D vs. 4D entities

- 4D entities are projections of normal (3D) entities into time intervals
- Example: KEG LD – Presenter entity
  - changed affiliation → another entity
  - due to minimalist modelling
  - but surely common at least in LD extracted from legacy sources

# How to encode into OWL?

- Explicit declaration of real-world setting
  - ...*analogy to OntoClean metaproperties or the FMA ‘boolean quadruple’*
  - Using the OWL annotation space
  - Using the entity naming ‘space’
- Linking between different styles for same real-world setting
  - Transformation patterns (e.g. PatOMat ‘educative’ scenario)

# Next steps

- Detailed empirical analysis of real-world settings behind existing resources
  - Straightforward for pattern libraries
  - More subtle for LD vocabularies
- Clean alignment with the FO stuff (Aldo...)
- Proposal for canonical modelling in OWL
  - Start a W3C IG on Annotations, or another channel?
- Prototype tool support (editor plugins?)
- Logical means for re-modelled real world
  - Some within the DL family, some beyond?