

Progress of work in 2011

The research proceeded according to the original plan, with only minor modifications. In particular, the research on the *reasoning use case* started several months earlier than planned, and was nearly completed by end 2011. On the other hand, the research on *logical patterns systematization* is slightly delayed, partly because the working visit to University of Manchester had to be shifted to early 2012. While the research on *naming patterns* progressed at the level of empirical experiments (using a tool created by a foreign partner), the actual theoretical work (using the apparatus of analytical linguistics) was postponed to 2012.

WP1 – Formal Model and Pattern Language

In year 2 the formal model was relatively stable. The ontology pattern language was however extended by new features reflecting the more advanced implementation: *optional axioms* and *variable number of arguments* in certain kinds of axioms. Furthermore, rich annotations reflecting ontology updates (ontology additions and removals) can now be expressed.

New research on systematic categorization of *published logical patterns* was eventually started in the end of 2011. The main motivation is to create a sound template according to which logical/structural patterns could be retrieved and indexed at the *OntologyDesignPatterns.org* portal. Submission of the results to a major European conference is envisaged for the first half of 2012.

WP2 – Metamorphic Ontologies

The detection and transformation service, already developed in year 1, was further enhanced, in particular in terms of handling so-called *additional axioms* (3 different strategies, and 3 sub-strategies) and *recursive* detection/application. Furthermore, in addition to RESTful services, the functionality was also implemented as a *Java library*, which allowed its subsequent use in the import use case (below). The new implementation was described in a demo paper [4]. In order to enable convenient authoring of transformation patterns, a *visual editor* was developed by student M. Dudáš. It is also described in [4], and available for download from <http://owl.vse.cz:8080/tpe/>.

The WP2 deliverable, *D2.3 Report on evaluation of style transformation*, had the character of a study in a particular WP4 use case, namely the content pattern import use case, which was identified as topical in 2011. The study has the form of a section in a working paper,¹ describing the pattern detection tests on three ontologies; it will be extended in the first half of 2012 towards an IF journal submission (joint with ISTC-CNR Rome). Another part of the paper corresponds to an early version of *D4.2 Report on the ontology import use case*.

The comparison of PatOMat with the Semion tool from ISTC-CNR (including a ‘data mining ontology’ case study) was not achieved in 2011 because the PhD student involved, A. Hazucha, left the group. The task will be completed in 2012, however in a reduced form, without an extensive case study.

¹ Šváb-Zamazal O., Svátek V., Daga E.: Adapting Ontologies to Content Patterns using Transformation Patterns: Method, Implementation and Case Studies. Submitted as full paper to the ISWC 2011 conference (19% acceptance rate); rejected but with rather encouraging reviews.

WP3 – Ontology Repair

The joint research on debugging ontologies based on *anti-patterns* with Universidad Politécnica de Madrid, Spain, progressed more slowly in 2011, as the lead investigator on the Spanish side moved to another workplace (CEMAGREF institute, France). More stable implementation of the detection algorithms was however built and experiments from 2010 partly redone. The evaluation of the repair system was described in a paper.²

On the other hand, a new cooperation was established with University of Freiburg, Germany, which focuses on *naming pattern* analysis for the purpose of ontology repair. Our group has contributed to the tuning of principles and to the empirical analysis of naming (labelling) in ontological taxonomies using their *OntoCheck* tool. The result of this cooperation was a workshop paper [3], an extended version of which was invited to an international journal [6]. In parallel, a simple form of naming-pattern-based ontology repair (leveraging on WordNet) was also implemented in terms of PatOMat transformation patterns.

The third thread in the ontology repair activities will be the collaboration with University of Leipzig, Germany, within the EU LOD2 project, see sections *Related projects* and *Work plan for 2012*.

The fact that the team is involved in these three overlapping activities with different partners caused a shift of a consolidated report on this topic (labeled as Deliverable D3.3) from M21 to year 2012. On the other, the quality of this outcome is likely to be higher due to integration of multiple independent input sources.

WP4 – Use Cases

The *ontology mapping* use case is mainly undertaken in cooperation with LIRMM, University of Montpellier (joint book chapter [1]) and University of Mannheim. The research focused on the identification of comparative advantages and disadvantages of transformation-based approach to mapping (i.e. ontology style transformation followed by simple matching) vs. an approach based on complex correspondences. It seems that the transformation-based approach is suited for specific settings where a single ontology is to be matched to a network of ontologies built in the same architectural style, while for casual matching, complex correspondences are sufficient.

The *ontology import* use case, carried out in cooperation with ISTC-CNR, Rome, achieved remarkable progress in 2011. PatOMat was integrated into the ISTC-CNR's *eXtreme Design toolset* (XDTools) in the form of an *Eclipse plugin*. The content pattern import adaptation (merging) functionality and its *wizard/based interface* are described in [4]. Initial evaluation of the detection part of the functionality was carried out; this work will be, after some extension, submitted to an IF journal in the first half of 2012. End-user experiments with the combined tool are also envisaged for 2012.

The third planned use case, that of *reasoning* support, was launched during the internship of O. Zamazal at University of Mannheim. Transformation patterns were designed for the purpose of

² Roussey C., Corcho O., Zamazal O., Scharffe F.: *Antipattern detection from web ontologies: an experiment using SPARQL queries*. In: *12e Conférence Internationale Francophone sur l'Extraction et la Gestion des Connaissances (EGC 2012)*, Bordeaux, France. By incidence, the paper was not equipped by project dedication, therefore it is not listed in the official outcomes in this annual report.

ontology adaptation for reasoners, in particular for downgrading an ontology to the *EL dialect of OWL*. A joint publication at a major conference is envisaged in 2012.

WP5 – Ontology Relationship Disambiguation

The study on *disambiguation of human relationships* expressed using the RDF vocabularies FOAF and RV continued in 2011. Formal axiomatization of the examined vocabulary properties was provided. A journal submission was prepared [7] in cooperation with ISTC-CNR, Rome. However, the implementation of a tool for the disambiguation task did not progress, as no student was found for this task.

The *mutual relationships between content patterns* in the OntologyDesignPatterns.org portal were also studied during the visit of M. Vacura in Rome. ISTC-CNR is currently performing a major refactoring of the content pattern base (alignment with the DOLCE-Lite ontology); lightweight collaboration on this task is envisaged.

Related projects

In connection with the research results obtained in the first year of PatOMat, the project team also succeeded in obtaining funding for complementary projects, with which it will cross-fertilize in 2012.

The EU FP7 integrated project³ LOD2, *Creating Knowledge out of Interlinked Data* (<http://lod2.eu>) addresses among other the problem of ontology repair and enrichment in connection with large, interlinked data sources. The functionality developed in PatOMat, in particular the ontology transformation and repair based on logical and naming patterns, will be applied there, in combination with the ORE tool⁴ from the University of Leipzig. The novelty of this use case consists, among other, in the presence of large numbers of instances that have to be processed in accordance with the transformation patterns. Comparison and possibly integration with another tool, *EvoPat*, which is a component of Leipzig's *OntoWiki*,⁵ is also envisaged.

The two-year⁶ Czecho-Slovak collaborative project *LAAOS – Logical Aspects of Adaptable Ontological Schemas*, focuses on formal aspects of ontology transformation patterns, in particular on their composition operation. By the preliminary information on the web of the Ministry of Education,⁷ the project was retained for funding. It will (thanks to the expertise of the partner, the Comenius University in Bratislava – group of Dr. Martin Homola) supply the formal-logical view of the problems that are mainly addressed at syntactical, implementation and experimental levels in the current project.

Publications and web presence

Overview of publications

The overall 2011 publication record (one book chapter and four papers at workshops / demo session / local conference) is less impressive than in 2010. The reasons are the following:

³ UEP joint the project as additional partner from 1 September, 2011, i.e. after one year of the project's duration, with three more years remaining. V. Svátek is Participant Coordinator and Workpackage Leader.

⁴ <http://ore-tool.net/Projects/ORE>

⁵ <http://ontowiki.net/Projects/OntoWiki>

⁶ Expected start in January 2012.

⁷ http://www.msmt.cz/file/18866_1_1/

- The publication activity in 2011 was focused on longer-term targets, especially journal articles. As two major conference papers (of which especially the EKAW 2010 one concerned the central thread of the project) already appeared in the end of 2010, collection of material for sufficient 'added value' with regard to these publications takes time.
- A significant part of researcher capacity was also devoted to necessary implementation tasks, which however pave the way to future experiments and publications based on them in the ultimate year.
- A part of the capacity of the coordinator was devoted to the harmonization of the CSF project goals with those of the new participation in the complementary EU FP7 LOD2 project.
- The key researcher and developer, O. Zamazal, spent the second half of the year at a post-doctoral internship, co-funded by German DAAD agency, at the University of Mannheim. While the content of the research undertaken there was in perfect accordance with the CSF project, the preference of the German partners was to start with the 'reasoning scenario' sooner, which caused a delay in producing publications from the already running threads of the project.

Nevertheless, by the end of 2011, the total count of accepted refereed publications reached 15. This indicator, as declared for the whole project duration, is thus already fulfilled – citing from the original proposal: “The total number of peer-reviewed publications (incl. papers at relevant workshops and local conferences) over the project lifetime is expected to reach 12-15.”

The following indicator is also fulfilled: “...2-3 papers will be published in proceedings of top-class conferences such as ISWC, ESWC, EKAW, K-CAP or FOIS (category D) or in books (category C)”; there was one paper published at ISWC and one at EKAW as early as in 2010, and there is a new book chapter [1] published in a Springer series.

The prime target for the last year thus remains “at least one but preferably two publications” in a recognized international journal with impact factor “to come out during or directly following the project”. The project team is working towards this goal as well. At the very beginning of 2012, two journal papers were submitted: one [6] (an invited extended version of the workshop paper [3]) to the open-access *Journal of Biomedical Semantics* (by BioMed Central, a subsidiary of Springer), indexed in PubMed, and one [7] to *Applied Ontology* (by IOS Press) – the lead journal in its respective field, which has been indexed by WoS and Scopus since May 2011. There is a concrete plan for 2-3 more journal submissions (at least one to a journal with high IF) by summer 2012 at the latest.

The research again leveraged on collaboration with leading European research centers. Paper [1] arose in collaboration with LIRMM, University of Montpellier, France; papers [3,6] with University of Freiburg, Germany; and papers [4,7] with ISTC-CNR, Rome, Italy.

Web presence of the project

The publications as well as other information about the *PatOMat* project are gradually made public at the *general website* of the project, <http://patomat.vse.cz>.

The implemented services are available from the user-oriented *implementation website*, <http://owl.vse.cz:8080/>. A substantial new feature in 2011 is a comprehensive *tutorial* available from

<http://owl.vse.cz:8080/tutorial/>. This allows an external user to start using the software even without communication with the developer.

Finally, the source codes of the application/s are at <http://sourceforge.net/projects/patomat/>.

Paper with RIV category C – Book chapter

1. Šváb-Zamazal, O., Svátek, V., Scharffe, F., David, J.: Detection and transformation of ontology patterns. In: Knowledge Discovery, Knowledge Engineering and Knowledge Management, IC3K 2009, Funchal, Madeira, Portugal, Revised Selected Papers, Springer CCIS, Vol. 128, 2011.

Papers from an international conference without ISI listing

2. Svátek, V., Šváb-Zamazal, O., Vacura, M.: Adapting Ontologies to Content Patterns using Transformation Patterns. In: 2nd Workshop on Ontology Patterns - WOP2010. Appeared in 2011 as <http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-671/>.
3. Schober D., Tudose I., Svátek V., Boeker M.: OntoCheck: Verifying ontology Naming Conventions in Protégé 4. In: Workshop OBML 2011, Berlin : IMISE 2011.
4. Zamazal O., Daga E., Dudáš M., Svátek V.: Tools for Pattern-Based Transformation of OWL Ontologies. In: ISWC 2011 - Demo session papers, Bonn 2011.

Papers from a local (Czecho-Slovak) conference

5. Šváb-Zamazal, O., Svátek, V.: OWL Matching Patterns Backed by Naming and Ontology Patterns. In: Znalosti 2011, Stará Lesná, Vysoké Tatry, 218-229. VŠB-FEI, 2011.

Refereed journal submissions

6. Schober D., Tudose I., Svátek V., Boeker M.: OntoCheck: Verifying ontology naming Conventions in Protégé 4. Submitted to *Journal of Biomedical Semantics*, BioMed Central.
7. Vacura M., Svátek V., Gangemi A.: Ontological Analysis of Human Relations for Semantically Consistent Transformations of Linked Data. Submitted to *Applied Ontology*, IOS Press.

Exploitation of equipment purchased from project resources

No equipment has been purchased.

Report on travel related to the project

- O. Zamazal spent 6 months (July-December) at University of Mannheim, Germany. The cooperation with the German partners focused on research corresponding to the goals of the CSF project. Most expenses were covered by a grant provided by the German DAAD agency; the CSF project only covered the transit costs for the initial part of the stay.
- V. Svátek,⁸ O. Zamazal and student M. Dudáš attended the International Semantic Web Conference, Bonn, Germany. They presented a joint contribution to the demo section, which

⁸ Part of the travel expenses of V. Svátek were covered by another CSF project, in connection with a different publication.

included a short plenary oral presentation, a poster presentation and online demo of three software components (PatOMat RESTful services, the XDTools module and the TPE) there was also informal discussion with colleagues from ISTC-CNR Rome, University of Manchester, University of Mannheim, FBK Trento and INRIA Rhone-Alps.

- M. Vacura visited ISTC-CNR, Rome. The visit primarily focused on joint preparation of a journal paper.
- V. Svátek attended, as lecturer, the Indian-summer school on Linked Data (ISSLOD-11) in Leipzig, Germany. Within the lecture, results of PatOMat were presented. In addition, cooperation with the AKSW group in Leipzig was discussed.
- V. Svátek attended the meeting of the EU LOD2 project⁹ in Leuven, Belgium. The main benefit for the PatOMat project was familiarization with the ORE tool, which was presented at the meeting by the partners from Leipzig.

Work plan for 2012

Most of the effort will concentrate on completing the research already started in the previous years and in materializing it in the form of prestigious publications (journals and major conferences). Although the WPs1-3 were, according to the original proposal, assumed to terminate during the (1st and) 2nd year of the project, the research cooperations established with foreign partners oblige us to deepen the research beyond the originally expected level even in the 3rd year. The high level of interest of third parties in using the generic transformation tools also bring the necessity to improve their (currently rather laconic) user interface.

As the 'desire' list below is rather large, we anticipate that some of these planned tasks might be sacrificed (to say, shifted to a follow-up project) to successful completion of the key publications.

WP2 will focus on

- *Systematization* of existing logical patterns from the point of view of both theory and practical usage, especially regarding the possibility to express them in the ontology pattern *model/language* and *transform* one to another while preserving the informal semantics. Cooperation with Manchester University is assumed. The collection will be ported to the *OntologyDesignPatterns.org* portal. (This research rather extends the original WP1.)
- Subsequently, the *transformation pattern library* will be extended and restructured.
- Research on the formal, logical aspects of *composition of transformations*. This research will leverage on the cooperation with the Comenius University, Bratislava, Slovakia. The new grant for bi-lateral cooperation (the LAAOS project above) will cover the travel expenses for mutual visits of both teams.
- More extensive and principled application of *naming patterns* incl. linguistic erudition.
- Further enhancement of *rich OWL annotations* aspects of ontology patterns, in cooperation with Manchester University.
- Design of a *graphical user interface* for interaction with the PatOMat services.

⁹ As the grant agreement for this project was signed rather late, the travel costs were covered from institutional resources. The only costs covered by the CSF grant were local travel costs (approx. 15 EUR).

- Practical evaluation of usability of the transformation framework on *real ontologies* from the web. The most prominent category will be *ontologies of products and services*. Existing cooperation of the workplace with Bundeswehr University in Munich, Germany, will be exploited.¹⁰
- Completion of the comparison study of PatOMat with the *Semion* tool from ISTC-CNR, in particular in connection with an instance-data-migration scenario.

WP3 will focus on

- Integration of the *naming pattern detection* methods into Freiburg's *OntoCheck* tool, and subsequent *experiments* in this context.
- Integration of the *naming pattern detection and repair* methods into Leipzig's *ORE* (and possibly also *OntoWiki*) tool.
- *Comparison and synthesis* of the repair (and repair-oriented detection) methods in all the WP3 threads (i.e. cooperations with Madrid/ CEMAGREF, Freiburg and Leipzig) as well as other related research).

WP4 will focus on

- Harmonization of the library of *complex alignment patterns* (mostly designed by colleagues from INRIA/Montpellier and Mannheim) with the new systematization of logical patterns in WP1; this will be the final tangible outcome of the *matching use case*.
- *Empirical tests* of content pattern detection precision and recall, wrt. the *import use case*
- Integration of the PatOMat library into the version of XDTools that is part of the *NeOn toolkit* (as more widespread ontological engineering environment).
- *User tests* of the content pattern import scenario, using XDTools with integrated PatOMat (presumably, at the site of the Italian partner).
- Completion of the *reasoning* use case, in particular in terms of *scalability evaluation*.

WP5 will focus on

- Analysis of the *reuse* of the FOAF and RV vocabularies with respect to their disambiguation in new contexts.
- Creation of specifications for a *software tool* supporting the disambiguation (the actual development of such a tool/prototype depends on availability of a student programmer).
- Analysis of further *Linked Data vocabularies* with respect to possible ontoepistemic predicates present in them.

Aside research and publication activities proper, we also anticipate preparation and submission of a *follow-up proposal* to CSF, most likely in the form of a post-doctoral project by O. Zamazal. The follow-up project would concentrate on *theoretical deepening* as well as *large-scale empirical evaluation* of the results of the current project (leveraging on functional implementation and international collaborations achieved in the current project).

¹⁰ This cooperation is primarily carried out under the label of a CSF project running in parallel, „Web Semantization“ (P202/10/0761). While „Web Semantization“ also focuses on product and service ontologies, it is mainly concerned by their construction from non-ontological resources and population from textual data. This research is orthogonal to that of transformation of such ontologies within the OWL language itself.