



Review of the Doctoral Thesis „XML Formats Evolution and Integration“ from Jakub Klímek

Jakub Klímek developed methods for XML schema evolution and for XML schema integration in his doctoral thesis. Both tasks had been realized with the same model-driven approach. In the thesis, the different kinds of conceptual models and transformation between these models are defined and their applications and extensions for evolution are represented. Complex evolution operations are introduced and propagated to the different models. The same model driven approach had been applied for an XML-schema integration method.

XML has established in the last years as a standard technology for data exchange and representation of semi structured data. The evolution of XML schemas is needed in all applications that are used for a long time. Till now, there exist only a few publications on that field. The Prague XML and Web Technologies Research Group developed a model driven design method for XML schema design. Jakub Klímek applied this method onto the problems of XML schema evolution as well as integration of different XML schemas in his thesis. Both fields are very important: the approach of using a Platform Independent Model (PIM) and Platform Specific Models (PSM) for supporting the schema re-design is a novel approach and it is formal correct defined in the submitted doctoral thesis.

The submitted doctoral thesis makes the following contributions:

- The complete process of XML schema evolution with a model-driven approach had been developed, formal defined, all subtasks are considered and the evolution process is evaluated with several real-life sample applications.
- The formal foundations for the reverse-engineering of conceptual models from XML schemas had been defined.

- The model driven approach had also been applied for the XML schema integration process. Jakub Klímek developed a semi-automatic method for deriving a conceptual model from different local schemas. He used an established similarity function for deriving of an initial conceptual model and developed a method for a continuous improvement of this model.
- The doctoral thesis also covers the problem of integration reusable schema parts: this problem is relevant for a novel design method and describes the foundation for designing and evolving of schemas with a library of design patterns.
- The conceptual model had been extended with two different kinds of inheritance hierarchies. The UML class model defines inheritance hierarchies with generalization relationships. In XML schema, a so-called structural inheritance is supported. Both kinds of inheritance relationships are introduced in the developed conceptual models and it is shown, how the different kinds of inheritance are translated into XML schemas.
- The developed models are also translated into Schematron - a schema representation method which can represent additional semantic constraints in contrast to XML schema.

Jakub Klímek demonstrates with his doctoral thesis and his publications in the last years his ability for scientific work. His work is embedded into the research of the Prague XML and Web Technologies Research Group. Jakub Klímek showed with his publications his ability to cooperate with other scientists as well as his ability to conduct his own research.

All chapter of the thesis had been written carefully and contain an introduction, an embedding of the work into the whole project, concept, formalization, related work and evaluation. The formal parts of the thesis are correct and complete.

The evaluations of the approach are another positive aspect of the thesis. The main parts of the thesis had been evaluated and tested with different real-world examples.

All chapters of the thesis have been published on international peer-reviewed workshops and conferences. Jakub Klímek is first author of several publications and contributed own subtask to the work of the Prague XML and Web Technologies Research Group.

Summarizing, I strongly recommend the acceptance of the submitted thesis. Jakub Klímek demonstrated with this PhD and his publications in the field of integration and evolution of XML schemas his ability for scientific work. He is able to develop solutions and to define the formal foundations for it, to implement his developed methods and to evaluate his solutions.



Meike Klettke

Rostock, 13. August 2013