Semantic Technologies for Decision Support - A Pattern-based Approach

Final report

Project leader: Eva Blomqvist <eva.blomqvist@liu.se>

February 27, 2018

Abstract: This project aimed at the adoption of Semantic Web technologies in Decision Support Systems (DSS), based on the use of Ontology Design Patterns (ODPs), with specific focus on industrial actors. Additionally, the project has created a (virtual) research group on semantic technologies at IDA that was previously missing. Important contributions include, in addition to a set of methods, software components, and ODPs: 29 scientific publications (including 5 journal articles, one PhD and one Lic thesis, and three book chapters), collaboration with five new university research groups in related areas, collaboration with several new industry partners on concrete methods and tools as well as introducing a domain focus both on the security and the e-health areas, acquisition of 6 national grants (e.g., from Vinnova and Energimyndigheten), one EU FP7-Sec grant (VALCRI) and one Horizon 2020 grant (SPIRIT) also in the security area, involvement in standardisation through W3C groups, arranging several public workshops, and a seminar series at LiU.

1 Short Project Summary

The Semantic Web has been researched for more than fifteen years; nevertheless, the techniques have only to a certain extent been applied to Decision Support Systems (DSS). However, this situation is about to change since (i) Semantic Web standards are in place, leading to the availability of stable and scalable software frameworks base on those standards, and (ii) methods and semantic technologies specifically relevant for DSS are also emerging, e.g. distributed Web-scale reasoning, and RDF stream processing [3]. Still, there is a need for methods allowing industry to more easily adopt and adapt semantic technologies to DSS.

Design Patterns (DPs) are well-tested and consensually agreed solutions to recurrent problems, which have proven effective in other fields, e.g., software engineering. DPs for semantic technologies are at the very forefront of Semantic Web research. The term Ontology Design Pattern (ODP) was coined simultaneously by the applicant, and Gangemi, in 2005 [1, 2], and now ten years later we have summarised the state of the area in some recent publications (c.f. project publications 5-8 listed in the appendix). Applying ODPs in DSS is a completely novel approach, which facilitates the wider industry adoption of Semantic Web technologies, and is a means to tailor technologies for use in DSS. Therefore the focus of this project has been on supporting various DSS-related needs and challenges by means of ODPs, as well as methods and tools related to them.

2 Scientific Results

Early in this project a literature survey and a set of interviews were performed, to establish the main challenges in the area of DSS where Semantic Web technologies could make an

important contribution. We identified the following focus areas (c.f. project publication 21 in the appendix):

1. *Information integration* - in particular the integration between information provided by human actors and existing structured data.
2. *Flexible information filtering* - in particular the opportunity to have context- and user-specific filtering that evolves along with user needs.
3. *Information aggregation and event detection* - beyond the plotting of raw information on maps or in graphs, i.e., what we call semantic Complex Event Processing (semantic CEP), where timeliness and flagging of potentially important situations is essential, as well as information summarization together with the possibility of drill-down into more detailed information.
4. *Model evolution* - to be able to handle changes and concept evolution in the real world, without extensive human intervention.
5. *Decision sharing* - conveying the meaning of decisions to external parties/systems.

The project has produced several *methods and software components* within focus areas 1-3, e.g., for addressing the information filtering and semantic CEP tasks (on top of existing state of the art RDF stream processing engines, as reviewed in project publ. 15, 26 and 28, and results presented in project publ. 2, 9, 10, 16, and 17, covering results from the VALCRI and E-care@home projects), and showcasing information integration through linked data (results from the LOD-Sweden, DEFRAM and DEFRAM-2 projects, c.f. publication 18). The semantic CEP software has been used in several demonstrators, e.g. in the agriculture, traffic, and criminal intelligence domains[1], and in e-health demonstrators such as influenza tracking (c.f. project publication 19), home care monitoring (c.f. project publication 1) as well as ongoing work on symptomic surveillance for animal health (current work together with SVA, not yet published). By demonstrating the benefits of Semantic Web and linked data technologies in several industry fields, all related to DSS, these results provide both good examples and experiences that benefits our industry partners when integrating these technologies into their daily business.

Additionally, several *ODPs* were produced earlier in the project, which are now used in the demonstrators, where one is targeting semantic CEP[2] (c.f. publ. 24-25), and a set of ODPs, so-called Statistical Knowledge Patterns (SKPs), are targeting information integration and filtering of linked data (associated with an *automated method and software components* for detecting and formalising such SKPs - c.f. publ. 4, 22, 23 and 27). In VALCRI additional ODPs have been emerging, together with reusable ontologies, through the development of a project specific ODP repository, but are so far not publicly available. In addition, a library of ODPs for the e-health domain has been developed, based on the W3C standard SSN[3], where a journal publication is still under review, but ODPs are publicly available online[4]. All together this has meant a considerable increase in the availability of ODPs for common DSS tasks, which in turn supports adoption of the related technologies by industry.

Also, work has been carried out on methodology improvements for ODP usage (c.f. publ. 3, 6, 7, 11, 12, 13, 14, 20 and 29), and better tooling[5] for collaborative ontology

[1]The queries and datasets used in the criminal intelligence demo, and a recording of the running demo is available at http://valcri.ida.liu.se:8080/eswc2015/, and an online demo of query templates at http://ontology.ida.liu.se:8680/iswc/rsp-spin

[2]http://ontologydesignpatterns.org/wiki/Submissions:EventProcessing

[3]https://www.w3.org/TR/vocab-ssn/

[4]See ODPs with a prefix "SmartHome" submitted by Marjan Alirezaie to the ODP portal at: http://ontologydesignpatterns.org/wiki/Submissions:ContentOPs

[5]The tool will be released as a WebProtégé modelling plugin, and is already available as open source code (https://github.com/hammar/webprotege), together with an online demo (http://wp.xd-protege.com/),

engineering with ODPs (c.f. publ. 3, 11, 13). During this project we have found that many of the barriers to adoption is actually on the organisational and methodological level, where simplified and high-quality tools and clear methodologies can make a considerable difference in bringing these technologies out to industry.

In addition, during the VALCRI project work also started towards covering the focus areas 4-5, although in the end they are only partly covered by the end of this project. In particular, the PhD student Zlatan Dragisic has been engaged in the context of VALCRI for developing solutions in ontology evolution, and another focus area of VALCRI (mainly covered through a collaboration with the project partner Middlesex University) is the recording and sharing of hypotheses and decisions made by police analysts within the system, i.e., contributing to focus area 5.

3 Organisational Achievements, Promotions and Theses

This project was initially proposed for supporting the project leader and employing a first PhD student. Partly through support from this project the applicant was in 2016 promoted to Universitetslektor (from Forskarassistent) and in 2017 she was accepted as Docent. The PhD student that was employed, Robin Keskisärkkä, successfully presented his Licentiate thesis in 2017 (c.f. project publication 2 in the appendix), and has been financially supported by this projects since starting his research work[6]. In summary, the persons directly funded by the project have been:

- Eva Blomqvist, project leader (Universitetslektor, Docent since Sept. 2017).
- Robin Keskisärkkä, PhD student (Lic presented in Nov. 2017)

Although not directly financially supported by this project, a second PhD student supervised by the project leader, Karl Hammar, completed his PhD in 2017 (c.f. project publication 3 in the appendix). This project has contributed to his PhD in terms of providing opportunities for the project leader to acquire and engage in related projects, e.g., VALCRI and E-care@home, where the student then performed studies and experiments included in the PhD thesis. Other affiliated PhD students include Zlatan Dragisic (supervised by Patrick Lambrix, and former member of the Semantic Web group described below) who completed his PhD in 2017, and Valentina Ivanova (again supervised by Patrick Lambrix, and member of the Semantic Web group described below) who completed her PhD in early 2018. Both these students were involved in the VALCRI project and worked on topics directly contributing to the aim of this CENIIT funded project.

The project supported three Master's Theses in the area during 2012-2014. One of the students studied the relation between current XML standards for message exchange in the emergency management domain and Semantic Web standards, while another student studied automated methods for ontology engineering based on texts expressed in Swedish.

In addition, the project has supported the creation of a virtual research group at IDA, focusing on Semantic Web technologies, which was completely missing at the start of the project. This group has currently seven members (4 senior researchers and 3 PhD students) from both the HCS and ADIT divisions of IDA, as well as external affiliated persons, such as Karl Hammar at Jönköping University. The group has regular meetings, runs a seminar series on Semantic Technologies at IDA with monthly seminars, collaborates on both project applications and work within current projects (such as VALCRI), and is giving a joint PhD course on Semantic Web topics.

and short video (https://www.youtube.com/watch?v=ZRH6vGXocqU&feature=youtu.be).

[6]Due to periods of parental leave, the student has taken longer than usual, calendar-wise, to complete his Lic.

The project has had specific connections with three related CENIIT projects, previously "Stream-Based Reasoning Grounded Through Sensing" and currently mainly "Semantic Parsing for Text Analytics", and "Integration and Interoperability of Graph-Data Systems". The first project, led by Fredrik Heintz, contributed mainly a discussion on stream reasoning aspects, that has continued since then, mainly between the involved PhD students (Daniel de Leng on one hand and Robin Keskisärkkä related to this project), for instance through seminars in the Semantic Technologies seminar series and joint participation in international workshops, such as the annual Stream Reasoning workshop[7]. The second project, led by Marco Kuhlmann, has led to Marco participating in the VALCRI project and becoming a co-supervisor of one affiliated PhD student (Zlatan Dragisic), mainly concerning the ontology evolution aspects addressed in VALCRI. Finally, the third project, led by Olaf Hartig, has spurred a collaboration also in the context of VALCRI, concerning different graph data format for data integration and provenance management of data in this project. Olaf is also one of the members of the Semantic Web virtual research group, mentioned above.

4 Industrial Collaboration

The initial *industry partners* of this project that still remain active and relevant include VSL Systems AB and FOI, which are involved in research and development of different aspects of DSS, but unfortunately we do not at this point have any joint projects. There have during the past years been attempts to extend our collaboration with Saab AB, and this is still ongoing. We have had several meetings discussing joint project applications, but at the moment the collaboration is at the level of joint seminars and workshops (latest meetings were held in November 2017).

In our previous VINNOVA projects we have collaborated with several Swedish SME:s, such as Novogit AB (Anders Östman) and Metasolutions AB (Hannes Ebner) in Stockholm, in particular concerning technologies for data integration and publishing, as well as a number of public bodies, e.g. MSB, SGU, Naturvårdsverket, Lantmäteriet and SVA. The collaboration with SVA has also led to a new VINNOVA-funded project on their part, where the project leader is currently acting as expert consultant, and Karl Hammar is now engaged as a postdoc. The project is focused on ontologies for syndromic surveillance of animal diseases, and applies the principles, methods and patterns developed in this CENIIT-funded project to support their ontology engineering process and DSS development.

Other active industry contacts consist of the industry partners in the VALCRI project, i.e., Space Applications Services in Belgium (Nick Evers), Object Security in the UK (Rudolf Schreiner), and AE Solutions in the UK (Rick Adderley). Together with the industry partners in VALCRI we have performed active software, and ontology, development, as well as evaluating our approaches and tools together with end-users in the criminal intelligence domain. The new EU-funded project SPIRIT, which is still under negotiation but expected to start in autumn 2018, is acquired in collaboration with AE Solutions, as a result of our collaboration in VALCRI.

The E-care@home project (where the project is represented through partial employment at RISE SICS East AB) has provided us with some opportunities to extend our industry collaboration to a new domain, i.e., the e-health field. The project as such does not consist of any industry partners, but there are ongoing discussions with several companies in the domain for acquiring new related projects. However, the collaborations in the e-health area are still in early development.

[7] http://www.ifi.uzh.ch/en/ddis/events/streamreasoning2018.html

The project also contributed to three W3C community groups, where the most prominent and currently active is concerned with RDF stream processing[8] (c.f. focus area 3).

References

[1] E. Blomqvist and K. Sandkuhl. Patterns in ontology engineering: Classification of ontology patterns. In *Proc. of ICEIS 2005, Miami, USA, May 25-28, 2005*, pages 413–416, 2005.

[2] A. Gangemi. Ontology Design Patterns for Semantic Web Content. In *The Semantic Web ISWC 2005*, volume 3729 of *LNCS*. Springer, 2005.

[3] E. Della Valle, S. Ceri, F. van Harmelen, and D. Fensel. It's a Streaming World! Reasoning upon Rapidly Changing Information. *IEEE Intelligent Systems*, 24(6):83–89, 2009.

5 Appendix - Project Publications

Project publications include (project participants in bold, journal articles marked with a *):

1. * Alirezaie, M., Renoux, J., Köckemann, U., Kristoffersson, A., Karlsson, L. **Blomqvist, E.**, Tsiftes, N., Voigt, T., Loutfi, A.: An Ontology-based Context-aware System for Smart Homes: E-care@home. *Sensors* 17(7): 1586, 2017.
2. **Keskisärkkä R.**: *Towards Semantically Enabled Complex Event Processing.* Licentitate thesis, Linköping University Electronic Press, 2017.
3. **Hammar, K.**: *Content Ontology Design Patterns - Qualities, Methods and Tools.* Dissertation, Linköping University Electronic Press, 2017.
4. * Zhang Z., Gentile A. L., Augenstein I., **Blomqvist E.**, Ciravegna F.: An Unsupervised Data-driven Method to Discover Equivalent Relations in Large Linked Datasets. In *Semantic Web* 8(2): pp.197-223, IOS press, 2017.
5. **Blomqvist E.**, Hitzler P., Janowicz K., Krisnadhi A., Narock T., and Solanki M. Considerations regarding Ontology Design Patterns (Editorial). In *Semantic Web* 7(1): pp.1-7, IOS press, 2016.
6. **Blomqvist E.**, **Hammar K.**, and Presutti, V. Engineering Ontologies with Patterns – The eXtreme Design Methodology. In *Ontology Engineering with Ontology Design Patterns: Foundations and Applications*, Hitzler et al. (Editors), Chapter 2, IOS Press, 2016.
7. **Hammar, K.**: Quality of Content Ontology Design Patterns. In *Ontology Engineering with Ontology Design Patterns: Foundations and Applications*, Hitzler et al. (Editors), IOS Press, 2016.
8. **Hammar, K.**, **Blomqvist, E.**, Carral, D., van Erp, M., Fokkens, A., Gangemi, A., van Hage, W. R., Hitzler, P., Janowicz, K., Karima, N., Krisnadhi, A., Narock, T., Segers, R., Solanki, M., and Svátek, V.: Collected Research Questions Concerning Ontology Design Patterns. In *Ontology Engineering with Ontology Design Patterns* pp. 189-198, IOS Press, 2016.
9. **Keskisärkkä R.** Representing RDF Stream Processing Queries in RSP-SPIN. In *Proceedings of the ISWC 2016 Posters & Demonstrations Track co-located with the 15th International Semantic Web Conference (ISWC-2016), Kobe, Japan, October 17-21, 2016* (to appear), CEUR Workshop Proceedings, 2016.
10. **Keskisärkkä R.** Query Templates for RDF Stream Processing. In *Proceedings of Stream Reasoning Workshop 2016 October 17th–18th, 2016, Kobe, Japan. Collocated with the 15th International Semantic Web Conference (ISWC 2016)* (to appear), CEUR Workshop Proceedings, 2016.
11. **Hammar K.**, and Presutti, V. Template-Based Content ODP Instantiation. In *Proceedings of the 7th Workshop on Ontology and Semantic Web Patterns (WOP 2016)*, CEUR Workshop Proceedings, 2016.
12. Karima N., **Hammar K.**, and Hitzler P. How to Document Ontology Design Patterns. In *Proceedings of the 7th Workshop on Ontology and Semantic Web Patterns (WOP 2016)*, CEUR Workshop Proceedings, 2016.
13. **Hammar, K.**: Ontology Design Patterns in WebProtégé. In *Proceedings of the ISWC 2015 Posters & Demonstrations Track co-located with the 14th International Semantic Web Conference (ISWC-2015), Betlehem, USA, October 11, 2015*, CEUR Workshop Proceedings, Vol.1486, 2015.
14. **Dragisic Z.**, Lambrix P., and **Blomqvist E.**: Integrating Ontology Debugging into the eXtreme Design Methodology. In *Proceedings of the 6th Workshop on Ontology and Semantic Web Patterns (WOP 2015)*, CEUR Workshop Proceedings, Vol.1461, 2015.
15. **Keskisärkkä R.** and **Blomqvist E.**: Sharing and Reusing Continuous Queries – Expression of Interest. In online *Proceedings of the RDF Stream Processing Workshop at ESWC2015*, 2015.
16. **Keskisärkkä R.** and **Blomqvist E.**: Supporting Real-Time Monitoring in Criminal Investigations. In *The Semantic Web: ESWC 2015 Satellite Events - ESWC 2015 Satellite Events, Portoroz, Slovenia, May 31 – June 4, 2015, Revised Selected Papers*, Springer, LNCS Vol.9341, 2015.
17. **Keskisärkkä R.** and **Blomqvist E.**: Towards the Use of RDF Stream Processing Engines for Event Enrichment from Social Media Streams. In online proceedings of the *Workshop on Semantics and Analytics for Emergency Response (SAFE2015) collocated with the The 12th International Conference on Information Systems for Crisis Response and Management (ISCRAM2015)*, 2015.
18. * **Blomqvist E.** and Thollander P.: An Integrated Dataset of Energy Efficiency Measures Published as Linked Open Data. *Energy Efficiency*, Vol. 8, Issue 6, 2015.

[8] http://www.w3.org/community/rsp/

19. * Timpka, T., Spreco, A., Dahlström, Ö., Eriksson, O., Gursky, E., Ekberg, J., **Blomqvist, E.**, Strömgren, M., Karlsson, D., Eriksson, H., Nyce, J., Hinkula, J., and Holm, E.:Performance of eHealth Data Sources in Local Influenza Surveillance: A 5-Year Open Cohort Study. In: *Journal of Medical Internet Research*, 16(4):e116, 2014.
20. **Hammar, K.**: Ontology Design Pattern Property Specialisation Strategies. In *Knowledge Engineering and Knowledge Management - 19th International Conference, EKAW 2014, Linköping, Sweden, November 24-28, 2014. Proceedings*, Springer LNCS, 2014.
21. * **Blomqvist, E.**: The Use of Semantic Web Technologies for Decision Support - A Survey. In: *Semantic Web Journal* (online preprint available 2012), 5(3): 177-201, IOS Press, 2014.
22. Zhang, Z., Gentile, A. L., **Blomqvist, E.**, Augenstein, I., and Ciravegna, F.: Statistical Knowledge Patterns: Identifying Synonymous Relations in Large Linked Datasets. In: *Proc. of the 12th International Semantic Web Conference, 21-25 October 2013, Sydney, Australia.* LNCS, Springer, 2013.
23. **Blomqvist, E.**, Zhang, Z., Gentile, A. L., Augenstein, I., and Ciravegna, F.: Statistical Knowledge Patterns for Characterizing Linked Data. In: *Proceedings of the Workshop on Ontology and Semantic Web Patterns (4th edition) - WOP2013*, CEUR, 2013.
24. Rinne, M., **Blomqvist, E., Keskisärkkä, R.**, and Nuutila, E.: Event Processing in RDF. In: *Proceedings of the Workshop on Ontology and Semantic Web Patterns (4th edition) - WOP2013*, CEUR, 2013.
25. Rinne, M., **Blomqvist, E.**: The Event Processing ODP. In: *Proceedings of the Workshop on Ontology and Semantic Web Patterns (4th edition) - Pattern track - WOP2013*, CEUR, 2013.
26. **Keskisärkkä, R. and Blomqvist E.**: Event Object Boundaries in RDF Streams: A Position Paper. In: *Proceedings of OrdRing 2013 - 2nd International Workshop on Ordering and Reasoning*, CEUR, 2013.
27. Zhang, Z., Gentile, A. L., Augenstein, I., **Blomqvist, E.**, and Ciravegna, F.: Mining Equivalent Relations from Linked Data. In: *Proc. of the annual meeting of the Association for Computational Linguistics (ACL) 2013.*
28. **Keskisärkkä, R. and Blomqvist E.**: Semantic Complex Event Processing for Social Media Monitoring - A Survey. In: *Proc. of Social Media and Linked Data for Emergency Response (SMILE) Co-located with ESWC - May 26-30, 2013 at Montpellier, France*, CEUR, 2013.
29. **Blomqvist E.**, Seil Sepour A., and Presutti V.: Ontology Testing - Methodology and Tool. In: *Knowledge Engineering and Knowledge Management - 18th International Conference, EKAW 2012.* LNCS, Vol. 7603, Springer, 2012.