Rapid Development of Executable Ontology for Financial Instruments and Trading Strategies

Dejan Lavbič and Marko Bajec

University of Ljubljana, Faculty of Computer and Information Science,
Tržaška cesta 25, 1000 Ljubljana, Slovenia
{Dejan.Lavbic,Marko.Bajec}@fri.uni-lj.si

Abstract. In this paper we employ Rapid Ontology Development approach (ROD) with constant evaluation of steps in the process of ontology construction for development of Financial Instruments and Trading Strategies (FITS) ontology. We show that ontology development process does not conclude with successful definition of schematic part of ontology but we continue with post development activities where additional axiomatic information and instances with dynamic imports from various sources are defined. The result is executable ontology as part of Semantic Web application that uses data from several semi structured sources. The overall process of construction is suitable for users without extensive technical and programming skills and those users are rather experts in the problem domain.

Keywords: ontology, semantic web, financial instruments, trading strategies, rapid ontology development.

1 Introduction

Semantic Web technologies are being adopted less than expected and are mainly limited to academic environment, while we are still waiting for greater adoption in industry. The reasons for this situation can be found in technologies itself and also in the development process, because existence of verified approaches is a good indicator of maturity. There are various technologies available that consider different aspects of Semantic Web, from languages for capturing the knowledge, persisting data, inferring new knowledge to querying for knowledge etc. Regarding the development process, there is also a great variety of methodologies for ontology development, as it will be further discussed in section 2, but simplicity of using approaches for ontology construction is another issue. Current approaches in ontology development are technically very demanding and require long learning curve and are therefore inappropriate for developers with little technical skills and knowledge. Besides simplification of the development process ontology completeness is also a very important aspect. In building ontology, majority of approaches focus on defining common understanding of a problem domain as a schematic model of the problem and conclude the development after few successful iterations. Post development

J.M. Zain et al. (Eds.): ICSECS 2011, Part I, CCIS 179, pp. 232–243, 2011.

activities that deal with defining instance data and employing developed ontology in Semantic Web application are usually omitted.

In this paper we apply Rapid Ontology Development (ROD) approach to construct Financial Instruments and Trading Strategies (FITS) ontology. The goal was to develop ontology by constructing schematic part of ontology including axiomatic information to fully support trading by employing reasoning. Furthermore this TBox part of ontology was combined to instance data (ABox) to construct knowledge base and therefore build mash up Semantic Web application to support financial instruments trading by applying various trading strategies. Target users of this approach are ones without extensive technical knowledge of data acquisition and ontology modeling but experts in financial trading. The main guideline in constructing ontology was to develop it to the level that enables direct employment in an application, which differs from majority of existing approaches where ontologies are mainly developed only to formally define the conceptualization of the problem domain.

The remainder of this paper is structured as follows. First we present some related work in section 2 with emphasis on ontology development methodologies and applications of financial ontologies. Next, in section 3, we introduce our approach for facilitating Semantic Web applications construction. The details of case study from the domain of financial instruments and trading strategies is further presented in section 4. First FITS ontology is presented, followed by semantic integration of data sources and then technological details about the prototype are depicted. Finally in section 5 conclusions with future work are given.

2 Related Work

Ontologies are used for various purposes such as natural language processing [1], knowledge management [2], information extraction [3], intelligent search engines [4], business process modeling [5] etc. While the use of ontologies was primarily in the domain of academia, situation now improves with the advent of several methodologies for ontology manipulation. Existing methodologies for ontology development in general try to define the activities for ontology management, activities for ontology development and support activities. Several methodologies exist for ontology manipulation and will be briefly presented in the following section. CommonKADS [6] is focused towards knowledge management in information systems and puts emphasis to early stages of software development for knowledge management. Enterprise Ontology [7] is groundwork for many other approaches and is also used in several ontology editors. METHONTOLOGY [8] enables building ontology at conceptual level and this approach is very close to prototyping. TOVE [9] is oriented towards using questionnaires that describe questions to which ontology should give answers. HCONE [10] is a decentralized approach to ontology development by introducing regions where ontology is saved during its lifecycle. OTK [11] defines details steps in two processes – Knowledge Meta Process and Knowledge Process. UPON [12] is based on Unified Software Development Process and is supported by UML language. DILIGENT [2] is focused on different approaches to distributed ontology development.

In the domain of finance several ontologies and implementations of Semantic Web based application exits. Finance ontology [13] follows ISO standards and covers several aspects (classification of financial instruments, currencies, markets, parties involved in financial transactions, countries etc.). Suggested Upper Merged Ontology (SUMO) [14] also includes a subset related to finance domain, which is richly axiomatized, not just taxonomic information but with terms formally defined. There are also several contributions in financial investments and trading systems [15-17]. Several authors deal with construction of expert and financial information systems [18-21].

3 Facilitating Semantic Web Applications Construction

3.1 Problem and Proposal for Solution

This paper describes semantic mash up application construction based on ontologies. The process is supported by continuous evaluation of ontology where developer is guided throughout the development process and constantly aided by recommendations to progress to next step and improve the quality of the final result. Our main objective is to combine dynamic (Web) data sources with a minimal effort required from the user. The results of this process are data sources that are later used together with ontology and rules to create a new application. This final result includes ontology that not only represents the common understanding of a problem domain but is also executable and directly used in the semantic mash up application.

Existing approaches for ontology development and semantic mash up application construction are complex and they require technical knowledge that business users and developers don't possess. As mentioned in section 2 vast majority of ontology development methodologies define a complex process that demands a long learning curve. The required technical knowledge is very high therefore making ontology development very difficult for non-technically oriented developers. Also majority of reviewed methodologies include a very limited evaluation support of developed ontologies and if this support exists it is limited to latter stages of development and not included throughout the process as is the case with our approach. Another problem that also exists is that the development process of ontology is completed after the first cycle and not much attention is given to applicability of ontology in an application.

3.2 Rapid Ontology Development

The process for ontology development ROD [22] that we follow in our approach is based on existing approaches and methodologies but is enhanced with continuous ontology evaluation throughout the complete process.

Developers start with capturing concepts, mutual relations and expressions based on concepts and relations. This task can include reusing elements from various resources or defining them from scratch. When the model is defined, schematic part of ontology has to be binded to existing instances of that vocabulary. This includes data from relational databases, text files, other ontologies etc. The last step in bringing ontology into use is creating functional component for employment in other systems.

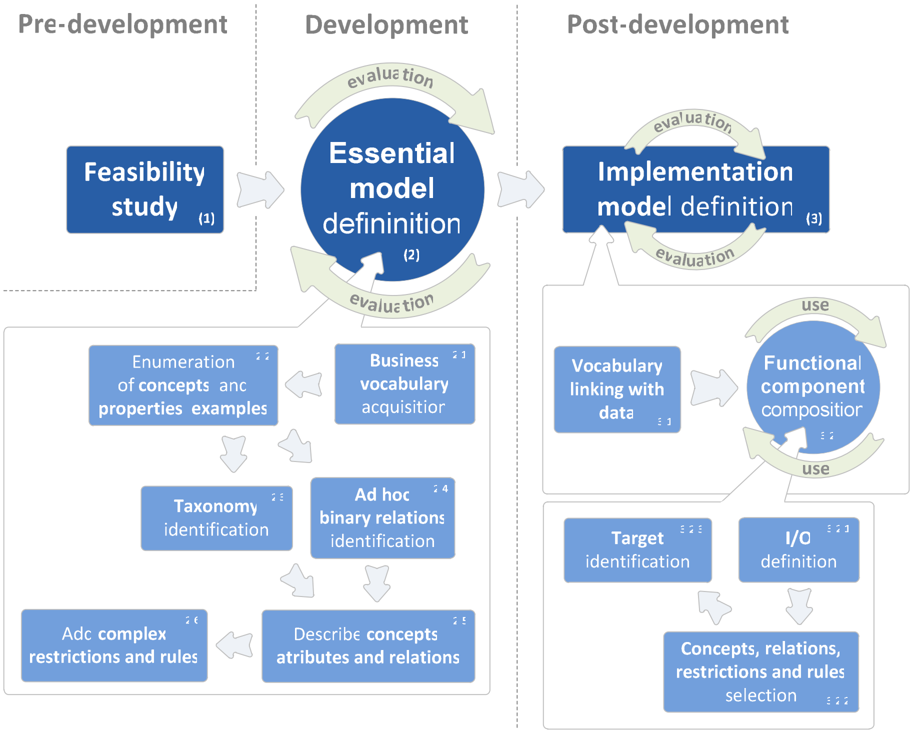


Fig. 1. Process of Rapid Ontology Development (ROD)

The ROD development process can be divided into the following stages: *pre-development*, *development* and *post-development* depicted in Fig. 1. Every stage delivers a specific output with the common goal of creating functional component based on ontology that can be used in several systems and scenarios.

The role of constant evaluation as depicted in Fig. 1 is to guide developer in progressing through steps of ROD process or it can be used independently of ROD process. In latter case, based on semantic review of ontology, enhancements for ontology improvement are available to the developer in a form of multiple actions of improvement, sorted by their impact. Besides actions and their impacts, detail explanation of action is also available (see Fig. 2). When OC measurement reaches a threshold (e.g. 80%) developer can progress to the following step. The adapted OC value for every phase is calculated on-the-fly and whenever a threshold value is crossed, a recommendation for progressing to next step is generated. This way developer is aided in progressing through steps of ROD process from business vocabulary acquisition to functional component composition. Detail presentation of ontology completeness indicator is further presented in [22].

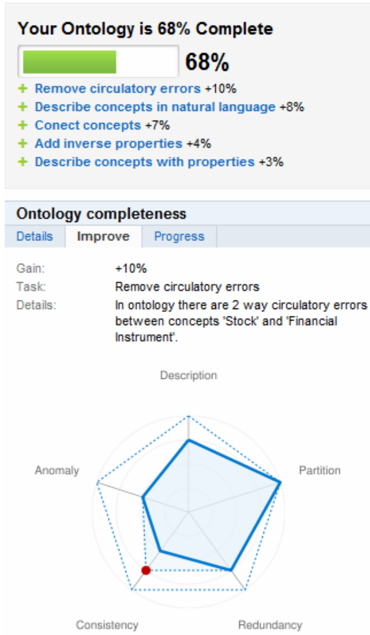


Fig. 2. Ontology completeness and improvement recommendation

4 Case Study Implementation and Discussion

4.1 FITS Ontology

The problem domain presented in this paper is financial trading and analysis of financial instruments. As already discussed in related work section there are several financial instruments ontologies already present. The purpose of our work was to extend these approaches to the information system level, couple the ontology with reasoning capabilities, define inputs, outputs, dynamic imports and build fully executable Semantic Web solution for financial instruments analysis and trading strategies. For this purpose basic Financial Instruments (FI) ontology was developed following ROD approach (see Fig. 3). The FI ontology introduces basic concepts, including financial instrument, stock exchange market, trading day and analysis. Further details in form of taxonomy are provided for financial instruments, trading day and analysis.

While FI ontology defines elementary entities from financial trading domain, are ontologies that capture trading strategies more complex, including advanced axioms and rules. In our case we have define four different trading strategies: (1) simple

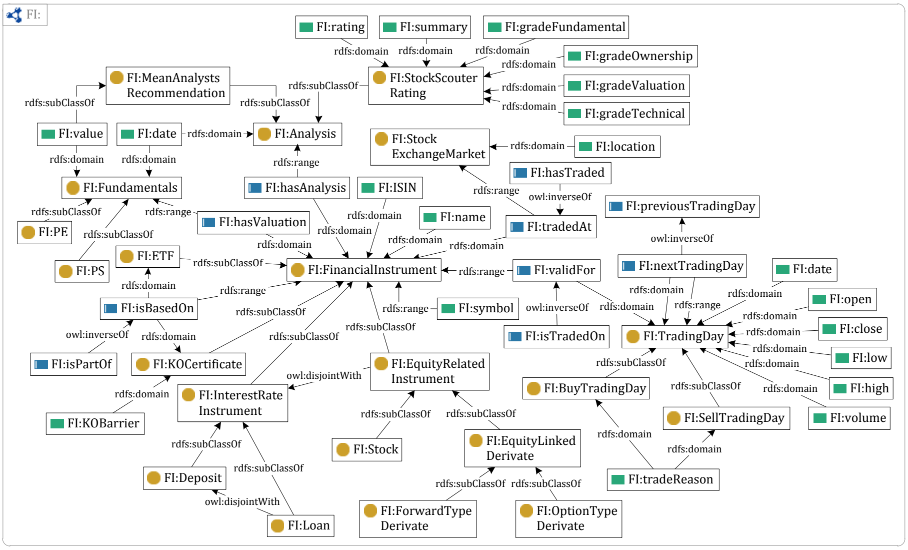


Fig. 3. Excerpt from FITS ontology

trading strategy (STs), (2) strategy of simple moving averages (SMAs), (3) Japanese candlestick trading strategy (JCTs) and (4) strategy based on fundamental analysis (FAs).

Every user has a possibility to define its own trading strategy whether from scratch or reusing existing ones. The main purpose of trading strategies is to examine the instances of *FI:TradingDay* concept and decide whether the instance can be classified into *FI:SellTradingDay* or *FI:BuyTraddingDay*. An example of this process can be found on Fig. 4 where and excerpt from JCTs is presented.

The JCTs is based on price movements which enable to identify patterns from daily trading formations. In this strategy price of a financial instrument is presented in a form of candlestick (low, open, close, high) and several patterns are identified (e.g. doji, hammer, three white soldiers, shooting star etc.). This strategy is rather complex but by following ROD approach (presented in section 3.2) domain experts can define it without being familiar with technical details of knowledge declaration and encoding.

After the selection of desired trading ontologies or composition of existing ones user can define the final ontology (see Fig. 5) which is than coupled with reasoning engine to allow the execution and performing trading analysis on real data available from several sources. At this point the schematic part of ontology (TBox component) is defined and further it still needs to be associated to instances (ABox component) by semantic integration of several data sources, which we will address in the following section 4.2.

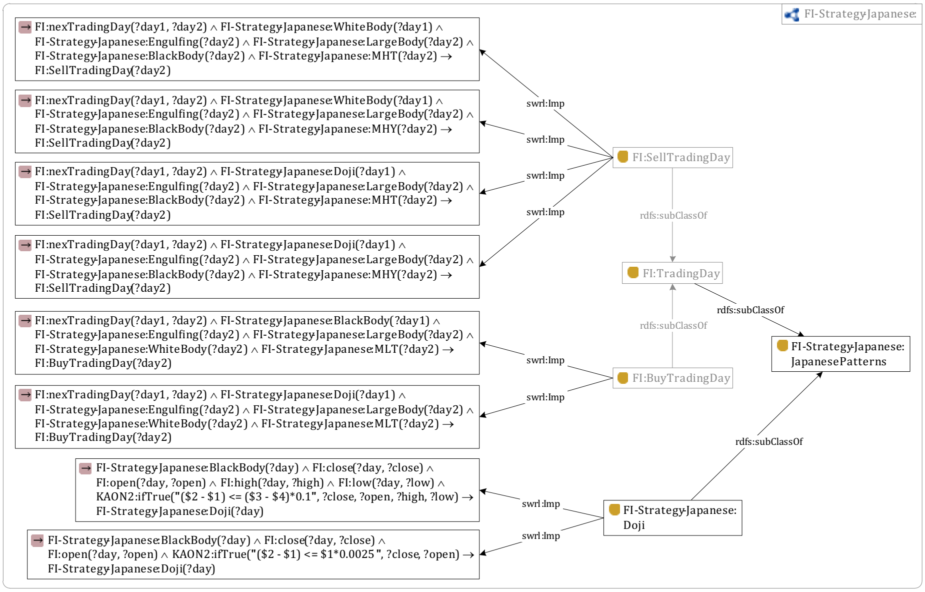


Fig. 4. Excerpt from Japanese candlestick trading strategy

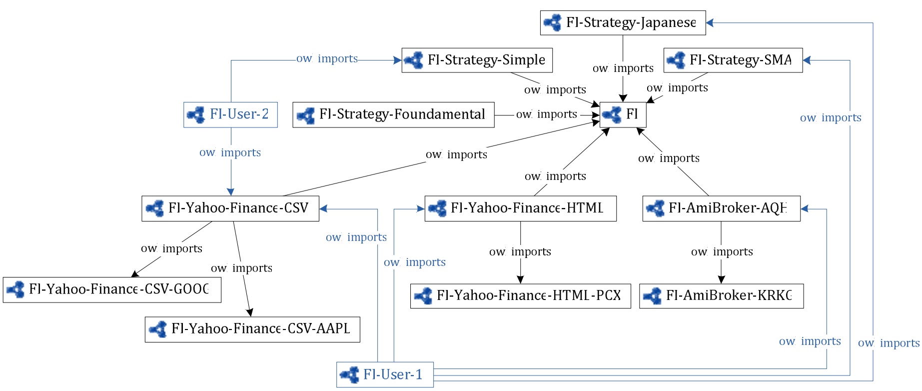


Fig. 5. Composition of final ontology for employment in Semantic Web application

4.2 Semantic Integration of Data Sources

In ROD approach there are several imports available: (1) existing ontologies, (2) relational or analytical databases, (3) CSV file and (4) semi structured data sources (e.g. HTML). In the process of creating FITS ontology the most prominent approach was reusing data from semi structured sources, mainly from HTML pages. When

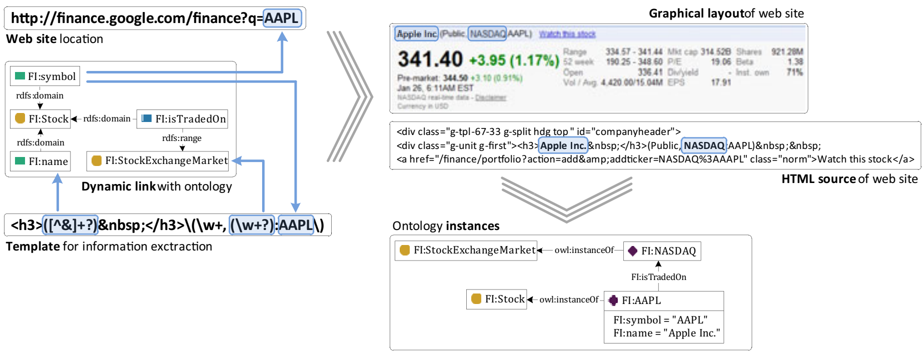
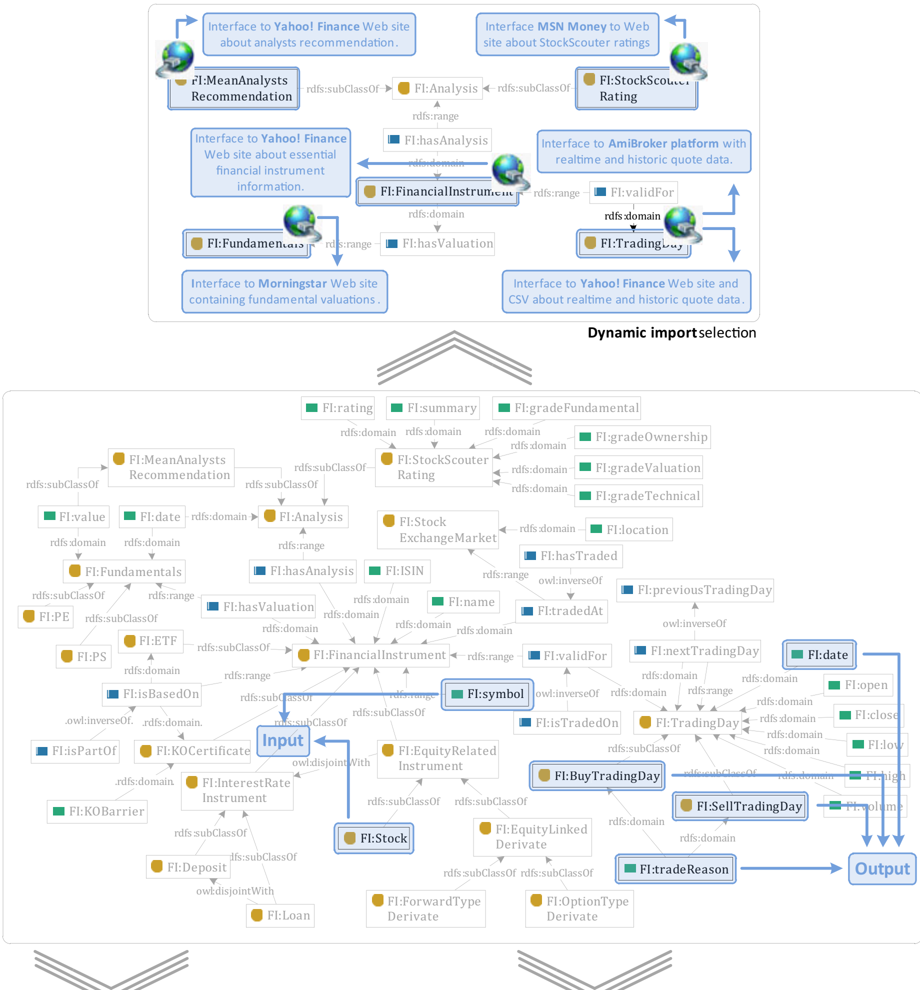


Fig. 6. Dynamic import of data property values related to financial instrument concept from Google Finance web data source

building executable ontology we relied on publicly available data about trading financial instrument, which are available on web pages and in vast majority in an unstructured form. Therefore linking wizard from ROD approach was used which incorporated the technology of regular expression and XQuery formulation for extracting data from semi structured data sources.

The role of semantic integration of data sources is to define wrapper to selected data sources and establish dynamic link between ontology entities (e.g. classes, properties etc.) and data source. An example of a simple web site wrapper is depicted in Fig. 6. This wrapper takes as an input financial instrument's symbol and uses Google Finance web page to extract information about financial instrument's name and stock exchange market where is being traded. As a result individuals are added or altered to the knowledge base with FITS ontology. These dynamic links can be defined for every selected entity as depicted in Fig. 7. For our case study there are 6 links defined. As analysis is concerned, mean analysts ratings are extracted from Yahoo! Finance web site, while stock scouter ratings are extracted from MSN money web site. All the essential data about the financial instruments are retrieved from Yahoo! Finance web site, while data about fundamental analysis are obtained from Morningstar web site. The quotes data are transferred from various sources, including historical data from Yahoo! Finance web site and real-time data from AmiBroker trading platform.

The last step in defining the Semantic Web application is to outline the input and the output component. The user can choose within the graphical interface which ontology entities will be used for input and which for output. In our case the input includes the symbol of stock that we want to trade and the outputs includes instances of trading days with buy or sell signals and trade reasons.



Input component

```
PREFIX FI: <http://www.lavbic.net/SW/FI.owl>
INSERT DATA {
    ?FI FI:symbol ?symbol .
    ?FI rdf:type FI:Stock
}
```

Output component

```
PREFIX FI: <http://www.lavbic.net/SW/FI.owl>
SELECT ?symbol ?date ?type ?tradeReason {
    { ?day rdf:type FI:BuyTradingDay } UNION { ?day rdf:type FI:SellTradingDay } .
    ?day rdf:type ?type .
    ?day FI:date ?date .
    FILTER (?date >= ?dateFrom) .
    FILTER (?date <= ?dateTo) .
    ?day FI:tradeReason ?tradeReason .
    ?day FI:validFor ?instrument .
    ?instrument FI:symbol ?symbol
}
```

Fig. 7. Dynamic import selection with input and output definition

4.3 Technology

The selected language for ontology presentation is OWL DL, since it offers the highest level of semantic expressiveness for selected case study and is one of the most widely used and standardized language that has extensive support in different ontology manipulation tools. Besides OWL logical restrictions, Semantic Web Rule Language (SWRL) rules were also employed due to its human readable syntax and support for business rules oriented approach to knowledge management [23].

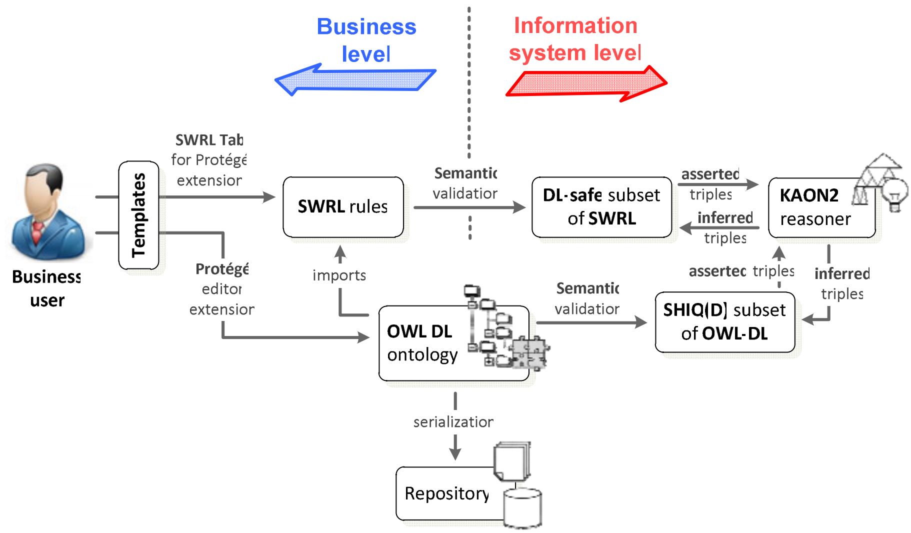


Fig. 8. Prototype of selected case study

The ontology manipulation interface for business users is based on Protégé Ontology Editor and Knowledge Acquisition System and SWRL Tab for Protégé. It enables entering OWL individuals and SWRL rules where a step further is made

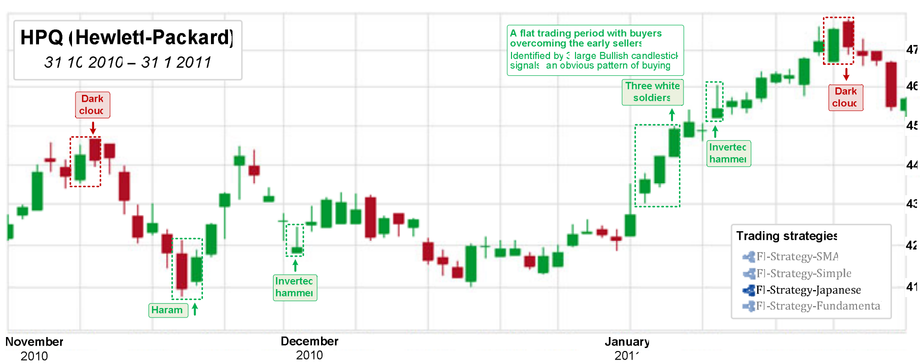


Fig. 9. GUI example of a Japanese trading strategy analysis on HPQ stock in the period from November 2010 to February 2011

towards using templates for entering information (see Fig. 8). At the information system level KAON2 inference engine is used to enable inference capabilities. Due to limitations of SHIQ(D) subset of OWL-DL and DL-safe subset of SWRL language, before inference is conducted, semantic validation takes place to ensure that all preconditions are met.

Fig. 9 depicts an example of firing trading rules on a real case scenario. The selected quote is HPQ (Hewlett-Packard) in the trading period of 3 months where several trading rules from Japanese trading strategy are being fired. From the GUI user can always select which subset of trading strategies is used (see section 4.1) and get details about the pattern found.

5 Conclusions and Future Work

Current approaches for ontology development require very experienced users and developers, while using the ROD approach for constructing FITS ontology is more appropriate for less technically oriented users. With constant evaluation of developed ontology that ROD approach offers, developers get a tool for construction of ontologies with several advantages: (a) the required knowledge for ontology modeling is decreased, (b) the process of ontology modeling doesn't end with the last successful iteration, but continues with post development activities of using ontology in a Semantic Web application and (c) continuous evaluation of developing ontology and recommendations for improvement. It has been demonstrated on a case study from financial trading domain that a developer can build Semantic Web application for financial trading based on ontologies that consumes data from various sources and enable interoperability. The solution can also be easily packed into a functional component and used in various systems. The results from using ROD approach is that the resulting artifact is executable ontology that is available in open format (e.g. OWL and SWRL language) and available for further inclusion. When reusing and building additional applications users have free selection of inference engines and also ontology manipulation tools. Added value is also defined in dynamic imports of data (instances in knowledge base) that can be acquired also at the runtime level.

The future work includes improvement of developed ontology and combining it with other approaches that mainly focus on schematic part of ontology and extend the possible use cases. One of the planned improvements is also integration with popular social networks to enable developers rapid ontology development based on reuse and therefore employ the community effort in curation process.

References

1. Staab, S., Braun, C., Bruder, I., Duesterhoeft, A., Heuer, A., Klettke, M., Neumann, G., Prager, B., Pretzel, J., Schnurr, H.P., Studer, R., Uszkoreit, H., Wrenger, B.: A system for facilitating and enhancing web search. In: Staab, S., Braun, C., Bruder, I., Duesterhoeft, A., Heuer, A., Klettke, M., Neumann, G., Prager, B., Pretzel, J., Schnurr, H.P., Studer, R., Uszkoreit, H., Wrenger, B. (eds.) International Working Conference on Artificial and Natural Neural Networks: Engineering Applications of Bio-Inspired Artificial Neural Networks, IWANN 1999 (1999)
2. Davies, J., Studer, R., Warren, P.: Semantic Web technologies - trends and research in ontology-based systems. John Wiley & Sons, Chichester (2006)

3. Wiederhold, G.: Mediators in the architecture of future information systems. IEEE Computer 25, 38–49 (1992)
4. Heflin, J., Hendler, J.: Searching the web with SHOE. In: Artificial Intelligence for Web Search, pp. 36–40. AAAI Press, Menlo Park (2000)
5. Ciuksys, D., Caplinskas, A.: Reusing ontological knowledge about business process in IS engineering: process configuration problem. Informatica 18, 585–602 (2007)
6. Schreiber, G., Akkermans, H., Anjewierden, A., de Hoog, R., Shadbolt, N., van de Velde, W., Wielinga, B.: Knowledge engineering and management - The CommonKADS methodology. The MIT Press, Cambridge (1999)
7. Uschold, M., King, M.: Towards a methodology for building ontologies. In: Workshop on Basic Ontological Issues in Knowledge Sharing (IJCAI 1995), Montreal, Canada (1995)
8. Fernandez-Lopez, M., Gomez-Perez, A., Sierra, J.P., Sierra, A.P.: Building a chemical ontology using methontology and the ontology design environment. Intelligent Systems 14 (1999)
9. Uschold, M., Grueninger, M.: Ontologies: principles, methods and applications. Knowledge Sharing and Review 11 (1996)
10. Kotis, K., Vouros, G.: Human centered ontology management with HCONE. In: IJCAI 2003 Workshop on Ontologies and Distributed Systems (2003)
11. Sure, Y.: Methodology, Tools & Case Studies for Ontology based Knowledge Management. In: Institute AIFB, vol. 332, PhD. University of Karlsruhe (2003)
12. Nicola, A.D., Navigli, R., Missikoff, M.: Building an eProcurement ontology with UPON methodology. In: 15th e-Challenges Conference, Ljubljana, Slovenia (2005)
13. Vanderlinden, E.: Finance ontology (2011)
14. Farrar, S., Lewis, W., Langendoen, T.: A Common ontology for linguistic concepts. In: Knowledge Technologies Conference, USA, Seattle (2002)
15. Zhang, Z., Zhang, C., Ong, S.S.: Building an Ontology for Financial Investment. In: Leung, K.-S., Chan, L., Meng, H. (eds.) IDEAL 2000. LNCS, vol. 1983, pp. 308–313. Springer, Heidelberg (2000)
16. Mellouli, S., Bouslama, F., Akande, A.: An ontology for representing financial headline news. Journal of Web Semantics 8, 203–208 (2010)
17. Qin, H., Taffet, M.D.: Vocabulary use in XML standards in the financial market domain. Knowledge and Information Systems 6, 269–289 (2004)
18. Chen, Y., Zhou, L., Zhang, D.S.: Ontology-supported web service composition: An approach to service-oriented knowledge management in corporate financial services. Journal of Database Management 17, 67–84 (2006)
19. Cheng, H., Lu, Y.C., Sheu, C.: An ontology-based business intelligence application in a financial knowledge management system. Expert Systems with Applications 36, 3614–3622 (2009)
20. Castells, P., Foncillas, B., Lara, R., Rico, M., Alonso, J.L.: Semantic web technologies for economic and financial information management. In: Bussler, C.J., Davies, J., Fensel, D., Studer, R. (eds.) ESWS 2004. LNCS, vol. 3053, pp. 473–487. Springer, Heidelberg (2004)
21. Ying, W., Sujanani, A., Ray, P., Paramesh, N., Lee, D., Bhar, R.: Design and Development of Financial applications using ontology-based Multi-Agent Systems. Computing and Informatics 28, 635–654 (2009)
22. Lavbič, D., Krisper, M.: Facilitating Ontology development with continuous evaluation. Informatica 21, 533–552 (2010)
23. Horrocks, I., Patel-Schneider, P.F., Bechhofer, S., Tsarkov, D.: OWL rules: A proposal and prototype implementation. Journal of Web Semantics 3, 23–40 (2005)