OBAMA, an Ontology-Based Software Tool for Agile Method Adoption

Soreangsey Kiv\textsuperscript{a}, Yves Wautelet\textsuperscript{c}, Samedi Heng\textsuperscript{b},
Manuel Kolp\textsuperscript{a}

\textsuperscript{a}LouRIM, UCLouvain, Louvain-La-Neuve, Belgium
{soeangsey.kiv, manuel.kolp}@uclouvain.be
\textsuperscript{b}HEC Liège, Université de Liège, Liège, Belgium
samedi.heng@ulg.be
\textsuperscript{c}KULeuven, Leuven, Belgium
{yves.wautelet}@kuleuven.be

Abstract

Tools like Protégé support the creation and edition of one or more ontologies in a single workspace. They nevertheless require a user to be familiar with this kind of abstractions and their supporting techniques such as a reasoner and SPARQL queries. This paper presents a step-by-step implementation of a prototype-tool that allows retrieving and displaying easily the information about agile practices contained in an ontology using Python programming language. Future development includes the flexible insertion, modification, and removal of knowledge by the user.

Keywords: Knowledge Representation, Ontology, Python Programming Language, Prototype-Tool.
1. Introduction

A lot of different practices are grouped under the umbrella of agile software development. Being agile does not necessarily mean adopting all of these practices (Eilers et al., 2020; Rahman et al., 2018). Organizations or individual software development teams often select a method (like Scrum, XP, SAFe, etc.) and try to adapt it in a specific business context. Then, all the practices prescribed by the chosen method do then not need to be implemented in the specific business environment and a customized on-demand implementation is generally preferable (Abbas et al., 2010; Kiv et al., 2021a).

The Ontology-Based for Agile Method Adoption (OBAMA) tool[^1] is a software system that aggregates a lot of knowledge on the adoption of agile practices.

[^1]: A demo in the form of a video of the tool can be found at [https://bit.ly/3B5dcJN](https://bit.ly/3B5dcJN)
found within the scientific literature. First, it fully implements an ontology depicting the relevant concepts in the field of systematic (agile) practices adoption built out of a systematic literature review. Second, it hosts a lot of knowledge gathered from literature reporting on the experience in adopting specific agile practices. Finally, through SPARQL queries, practitioners can easily access the most relevant information when adopting specific agile practices. The process to retrieve relevant information from the ontology using available tools such as Protégé tool requires users to follow some steps including (1) creating a new instance to represent a development team, (2) describing the team’s properties (3) executing the reasoner to automatically link the team instance to the other instances based on inference rules, (4) and using SPARQL to retrieve the desired concerns. Since this information retrieval process is complicated and requires a user to have some preliminary knowledge, a user-friendly tool to simplify it is deemed important to build.

The tool presented in this paper is part of a larger research effort to build and validate an ontology for agile practice adoption. A primary ontology for practices adoption was presented in Kiv et al. (2017a,b, 2019) while the entire research to build the ontology validating and populating it has been depicted in Kiv et al. (2022). Ontology creation has been used previously in the domain of agile methods like in Wautelet et al. (2014) with the aim of creating requirements (conceptual) models and assist in software engineering (see Wautelet et al. (2017)); it is here used as a database structure to support the knowledge retrieval of agile practice adoption in a flexible and custom manner. This paper serves as a reference for the OBAMA tool, it describes the technology supporting the software as well as its architecture and impact.

In Section 2.1, we provide a basic explanation about each of the important components needed to develop the tool. Section 2.2 explains the technical architecture of our tool, with a set of steps from loading until displaying the information to the users. Section 2.3 describes the functionality of the tool. Finally, we conclude with the result, the impact and future research directions in Section 3.
2. Software Description

2.1. Main Technologies Supporting the OBAMA-Tool

There are four main technologies used for the development of our supporting tool are:

- An ontology file written in Web Ontology Language (OWL) \cite{McGuinness:2004}. This file constitutes the database for storing the collected knowledge about agile practice adoption;

- The Python programming language was chosen because it provides an extensive Application Programming Interface (API) to easily work with the ontology;

- Owlready2, an ontology-oriented programming package in Python that allows loading, modifying, saving, and executing a reasoner on the ontology \cite{Lamy:2017}. Since the HermiT OWL reasoner included in Owlready2 is written in Java programming language, we need to install Java.exe to execute it;

- wxPython, a cross-platform GUI toolkit that we use to create the interface \cite{Precord:2015}.

These technologies are described in this section.

2.1.1. The Ontology File Written in Web Ontology Language

All the descriptions of concepts, relationships, and inserted knowledge related to agile practice adoption are written in the OWL format. The ontology file can be seen as the database file of the system.

OWL is an ontology language recommended by the World Wide Web Consortium (W3C) and it provides three sub-languages: OWL Lite, OWL Description Logic (DL), OWL Full \cite{McGuinness:2004}. OWL Lite is the core language for creating an ontology with a minimum number of restrictions and constraints. OWL DL is a restricted version of OWL Full. It has some restrictions on the context in which language construction can be used to ensure
decidability. OWL Full uses the same vocabulary as the above presented OWL DL version. However, it is much more expressive and has fewer restrictions than other versions. We use OWL Full for our ontology creation as it is more descriptive and less restrictive, which makes it the most convenient choice for usage. OWL can be seen as an extension of the Resource Description Framework (RDF) vocabulary because each OWL document is an RDF. Like RDF, OWL vocabulary includes a set of XML syntaxes that are used to describe knowledge in triples, “Subject-Predicate-Object”, just like a statement in a natural language. For example, within the RDF format statement “practice is composed of an activity”: practice is the subject, is composed of is the predicate, and activity is the object.

Following the procedure of ontology creation (Noy & McGuinness, 2001), we started by defining the key concepts (classes) and then hierarchically refined the concepts into sub-classes. In the concept refinement process, one of the decisions to take was whether we should introduce a new sub-class or simply distinguish through instances with different property values (Noy & McGuinness, 2001). For instance, there are seventeen different types of situation that can affect the agile adoption result, including team size, project size, type of communication, team distribution, etc. In this case, we need to decide whether we should represent each type of situation by a sub-class or an instance. According to (Noy & McGuinness, 2001), we can identify a class when each type has a different relationship with other class(es). Based on this principle, as each type of situation has a different effect (hurt or harm) on the requisite, we represent each type of situation by a sub-class in our ontology model. An example of OWL syntax to create class and sub-class in our ontology is illustrated in Figure 1. In the example, Role and Situation are the main classes where Communication is the sub-class of Situation.

Once all classes were created, we continued with the definition of their properties that include two types: data and object. (1) The data property is used to link the individuals (i.e., instances) and data value. Every class in our model has only two data properties (i.e., Name and Description). Based on the literature,
these two data properties are the only common elements used to describe each class by agile practitioners. These properties are String types. (2) Object property is used to link individuals (i.e., instances) and individuals (i.e., instances). Both links are built in the form of “Domain - data/object property - Range”. For instance, the relationship “Practice - Achieve - Goal” has Achieve as an object property where Practice is its domain and Goal is its the range. Figure 2 is an example of OWL syntax to define an object property. In the example, the object property Achieve is the reverse of another object property named Achieved by. In addition, either a team or practice (domain) can achieve either goal or principle (range).

Finally, we created the individuals (i.e., instances) based on the knowledge we had extracted from the literature. We started by creating an individual of a class Team for each case study. We then continued to create other individuals for the other information related to that team, for instance, the practice the team adopted and the activity it performed as part of the practice, etc. When
all the individuals were created, we built their relationships by defining their object properties. Figure 3 is an example of OWL syntax to create an individual Team42: Sprint review and define its relationships with other individuals.

2.1.2. SPARQL

SPARQL is a query language that was standardized in 2008 by the W3C that is used for querying RDF data (Prud’hommeaux & Seaborne, 2005). Most forms of SPARQL queries contain a set of basic triple graph patterns (subject, predicate, object) where each pattern matches a sub-graph of RDF data and the terms from that sub-graph may be substituted for the variables (Schmidt et al., 2010). To query for the information in our ontology, each SPARQL query comprises 4 parts of information in the following order. (1) A prefix that is the abbreviation of the Uniform Resource Identifiers (URIs), (2) the information to return from the query, (3) the triple patterns, and (4) the query modifiers to filter and order the query results. Figure 4 is an example of how to write a SPARQL query to select all the solutions to solve the problems that are encountered by Daily meetings.

2.1.3. OWLReady2

Owlready2 is a module for ontology-oriented programming in Python that we use to access the ontology, execute the reasoner, and query the information. The general structure of the Owlready2 is made of five main components as
They include (1) an optimized RDF quadstore implemented with an SQL database (i.e., SQLite3) and stored either in memory or on disk in a file, (2) meta-classes for OWL classes and constructs, (3) optional ontology-specific Python source files defining methods to insert into OWL classes, (4) the HermiT OWL reasoner for performing automatic classification and (5) the SPARQL engine from the RDFlib Python module that is used for RDF query [Lamy 2017].

The general algorithm that we use to query the ontology is OWLReady2 as shown in Figure 6. First, the ontology is imported into RDF quadstore and their concepts and/or entities are saved in Python objects. Then, the HermiT reasoner is executed. After that, the updated entities of the ontology are automatically stored in a memory in RDF format. We then perform SPARQL queries using the RDFlib graph. Finally, a Python function captures and converts the results into Strings and stores them in a list-type variable.

2.1.4. wxPython

wxPython is a cross-platform toolkit that is used to create desktop Graphical User Interface (GUI) applications. wxPython provides a set of methods and
Python runtime memory

Metaclasses
- Classes
- Methods
- Properties

Ontology-specific Python source files

Methods

Ontologies files
- NTriples
- RDF/XML
- OWL/XML

RDF quadstore

SQL database (SQLite3)

Reasoner (HermiT)

Serializer RDFS-triples

Automatic updates

Lazy parser

Import

update with inferences

Figure 2: General architecture of Owlready. It includes 5 components: (1) an optimized RDF quadstore, implemented with an SQL database in SQLite3 and stored either in memory or on disk in a file, (2) metaclasses for OWL classes that have been redefined in order to adapt OWL ontologies to the Python object model, (3) optional ontology-specific Python source files, (4) the HermiT reasoner [11], for performing automatic classification, (5) defining methods to insert into OWL classes, (6) importing Python source files associated with the ontologies, and wrapping their entities and creating Python objects, implementing in Python an unified interface, mixing OWL statements (properties, class hierarchy, individuals) with Python statements (method definitions).

5.1. General architecture

5.2. Mapping OWL to the Python object model

Figure 5: OWLReady2 Architecture [Lamy, 2017]

Figure 6: Ontology query in Python programming language.
objects that programmers can use to simply and easily create programs with a robust and highly functional GUI (Precord, 2015). Technically, wxPython is a wrapper over a C++ GUI API called wxWidgets which are essential building blocks of a GUI application. wxPython provides a lot of widgets that can be divided into six groups. (1) Base Widgets provide basic functionality for derived widgets, (2) top-level widgets provide different functions independently of each other, (3) containers that contain other widgets, (4) dynamic widgets that can be edited by users, (5) static widgets that are used to only display information but cannot be edited by the users, and (6) random widgets such as status-bar, toolbar, and menu-bar, etc. There is a specific relation among widgets in wxPython that is developed by inheritance. Widgets form a hierarchy; they can inherit functionality from other widgets. The advantages of wxPython include being open-source, cross-platform, easy to write, and easy to understand.

2.2. Tool Architecture

The architecture of the tool is illustrated in Figures 7 and 8 with two different scenarios. One is to retrieve all the information while the other is to retrieve the information based on the user’s inputs.

2.2.1. Technical architecture for all the information related to a practice

In the scenario where all the information in ontology is retrieved and listed, the retrieval process occurs only the first time the program is executed, and the returned information is then stored in the permanent memory for future usage. As there is no any change to the information, loading the information from the memory saves much more time than re-loading from the ontology file. The program is executed in 9 steps, from loading data from the ontology to displaying the results for the users.

1. The program imports the ontology from the source file (OWL), loads their entities recursively, and wraps them in a Python object named OwlReady Object;
2. The HermiT reasoner is then executed to auto-generate the classification of the classes, properties, and individuals based on the inference rules. After that, it updates the Owlready Object;

3. The program loads all the pre-written queries to retrieve the information related to all the concerns (one query for one concern) and stores them in a list-type variable;

4. RDFLib SPARQL engine then iteratively runs all the queries from the list;

5. The results of each query retrieved by the RDFLib SPARQL engine are then converted and stored in another list-type variable;

6. Once all the queries are run, all the query results are stored in permanent memory for future usages;

7. From the GUI, users select a concern from the list;

8. The program loads the information related to the selected concern from the memory;

9. Finally, the information is displayed in table format.

2.2.2. Technical architecture for the information related to practice based on inputs

In the scenario where the information is retrieved based on the users’ inputs, the retrieval process occurs each time the users click on the button “Calculate
result” in the “Input page 1” or the “Input page 2”. There are 11 steps in total from loading data from the ontology to displaying results for the user.

1. From the GUI, users provide the input values (goals or situations);
2. The program imports the ontology source file, loads their entities recursively, and wraps them in a Python object named *OwlReady Object*. This step is done in the previous scenario;
3. The program then creates a temporary ontology object of the class *Team* where the input values are its properties;
4. The *HermiT* reasoner is then executed to auto-generate the classification of the classes, properties, and individuals based on the inference rules. By doing so, the temporary ontology object are now linked to the other individuals based on pre-defined inference rules. After that, it updates the *Owlready Object*;
5. The program loads all the pre-written queries to retrieve the information related to all the concerns (one query for one concern) and stores them in a list-type variable;
6. The *RDFLib SPARQL* engine then iteratively runs all the queries from the list;
7. The results of each query retrieved by RDFLib SPARQL engine are then converted into Strings and stored in another list-type variable;
8. All the query results are stored in a temporary memory;
9. From the GUI, users select a concern from the list;
10. The program loads the information related to the selected concern from the memory;
11. Finally, the information is displayed in the table format.

2.3. Functionality of the Tool

In this paper, we thus focus on the prototype-tool that allows retrieving and displaying information. This tool is meant to help practitioners see how useful the ontology could be when the knowledge is available. The part where a user
can insert, modify, and delete knowledge will be developed at a later stage. It was created in a notebook-style where each functionality is put on a different page as shown in Figure 9.

- The first page, called “Welcome page”, serves to introduce to users the purpose of the tool and some usage guidelines;

- The second page “All the information related to practice” allows users to access all the information related to agile practice adoption that we have inserted. All concerns are listed in a combo-box. By choosing a concern from the list, the information will be displayed underneath in a table format;

- If users want to filter the information based on goals, they need to provide input values in the “Input page 1”. Since most goals in adopting agile methods can be mapped to the Agile Manifesto (Kiv et al., 2018), we provide on this page a list of agile values and principles from which users can choose;

- If users want to filter the information based on their situations, they need to provide their input values in the “Input page 2”. On this page, we provide 13 situational factors which have been collected through an SLR (Campanelli & Parreiras, 2015). Each of the factors has a list of different
values from which users can choose. In both “Input page 1” or “Input page 2”, there is a button “Calculate result” that allows starting the data retrieval process;

- Our tool retrieves the information based on users’ inputs before it displays the results in the page “Information related to practice based on input”. Users can find the desired information by choosing a concern from the list.

Our ontology file and source code in Python programming language can be found at [Kiv et al., 2021b].

![Figure 9: Screenshot of OBAMA tool.](image)

3. Impact and Conclusion

Ontologies have been widely used in many fields, but not many research articles address the difficulties of working with ontologies for an average user and how to overcome this problem. The purpose of this paper is to elaborate on how to create a user-friendly tool in Python to help users retrieve information without any required preliminary knowledge. As said, the software is aimed to be used by practitioners, especially novice ones since their potentially limited
experience makes the software tool especially relevant for them. By using our tool, practitioners can easily get general information on agile practice adoption. They can also filter for only relevant information based on their goals and context. The tool has been evaluated by agile practitioners on a set of criteria related to efficiency in Kiv et al. (2022). Results allow us to declare that all the features in our tool are efficient enough for users. Most experts “Somewhat agree” or “Agree” that they can access, understand and input the information easily. Even though we cannot say it is already satisfying enough for massive adoption by users yet, but this result is overall still reassuring. It is also important to notice that, most experts “Agree” that our tool helps them decide whether or not an agile practice is suitable for their team.

The tool has been developed in a laboratory setting and validated by practitioners on the mentioned elements. More work is currently made with novice agilists in real-life settings to further evaluate the tool but also improve it to align at best to their requirements.

References

Abbas, N., Gravell, A. M., & Wills, G. B. (2010). Using factor analysis to generate clusters of agile practices (a guide for agile process improvement). In AGILE Conference (pp. 11–20). doi:https://doi.org/10.1109/AGILE.2010.15

Campanelli, A. S., & Parreiras, F. S. (2015). Agile methods tailoring—a systematic literature review. Journal of Systems and Software, 110, 85–100. doi:https://doi.org/10.1016/j.jss.2015.08.035

Eilers, K., Simmert, B., & Peters, C. (2020). Doing agile vs. being agile—understanding their effects to improve agile work. In International Conference on Information Systems (ICIS).

Kiv, S., Heng, S., Kolp, M., & Wautelet, Y. (2017a). An intentional perspective on partial agile adoption. In J. Cardoso, L. A. Maciaszek, M. van Sin-
Kiv, S., Heng, S., Kolp, M., & Wautelet, Y. (2018). Agile manifesto and practices selection for tailoring software development: A systematic literature review. In International Conference on Product-Focused Software Process Improvement (pp. 12-30). doi:https://doi.org/10.1007/978-3-030-03673-7_2

Kiv, S., Heng, S., Kolp, M., & Wautelet, Y. (2019). Agile methods knowledge representation for systematic practices adoption. In P. Kruchten, S. Fraser, & F. Coallier (Eds.), Agile Processes in Software Engineering and Extreme Programming - 20th International Conference, XP 2019, Montréal, QC, Canada, May 21-25, 2019, Proceedings (pp. 19-34). Springer volume 355 of Lecture Notes in Business Information Processing. URL: https://doi.org/10.1007/978-3-030-19034-7_2

Kiv, S., Heng, S., Wautelet, Y., & Kolp, M. (2017b). Towards a goal-oriented framework for partial agile adoption. In E. Cabello, J. Cardoso, L. A. Maciaszek, & M. van Sinderen (Eds.), Software Technologies - 12th International Joint Conference, ICSOFT 2017, Madrid, Spain, July 24-26, 2017, Revised Selected Papers (pp. 69-90). Springer volume 868 of Communications in Computer and Information Science. URL: https://doi.org/10.1007/978-3-319-93641-3_4

Kiv, S., Heng, S., Wautelet, Y., & Kolp, M. (2021a). Towards a systematic socio-intentional framework for agile methods tailoring. In 2021 IEEE 23rd Conference on Business Informatics (CBI) (pp. 143–152). IEEE volume 2.

Kiv, S., Heng, S., Wautelet, Y., Poelmans, S., & Kolp, M. (2021b). Obama - ontology-based for agile method adoption [source code]. doi:https://doi.org/10.24433/C0.8008250.v1
Kiv, S., Heng, S., Wautelet, Y., Poelmans, S., & Kolp, M. (2022). Using an ontology for systematic practice adoption in agile methods: Expert system and practitioners-based validation. *Expert Systems with Applications,*(p. 116520).

Lamy, J.-B. (2017). Owlready: Ontology-oriented programming in python with automatic classification and high level constructs for biomedical ontologies. *Artificial intelligence in medicine, 80,* 11–28. doi:[https://doi.org/10.1016/j.artmed.2017.07.002](https://doi.org/10.1016/j.artmed.2017.07.002).

McGuinness, D. L., Van Harmelen, F. et al. (2004). Owl web ontology language overview. *W3C recommendation, 10,* 2004.

Noy, N. F., & McGuinness, D. L. (2001). Ontology development 101: A guide to creating your first ontology. [https://protege.stanford.edu/publications/ontology_development/ontology101.pdf](https://protege.stanford.edu/publications/ontology_development/ontology101.pdf).

Precord, C. (2015). *WxPython Application Development Cookbook.* Packt Publishing Ltd.

Prud’hommeaux, E., & Seaborne, A. (2005). *SPARQL Query Language for RDF.* Technical Report W3C. [https://www.w3.org/TR/rdf-sparql-query/](https://www.w3.org/TR/rdf-sparql-query/).

Rahman, A., Agrawal, A., Krishna, R., Sobran, A., & Menzies, T. (2018). “doing” agile versus “being” agile, .

Schmidt, M., Meier, M., & Lausen, G. (2010). Foundations of sparql query optimization. In *International Conference on Database Theory* (pp. 4–33).

Wautelet, Y., Heng, S., Kiv, S., & Kolp, M. (2017). User-story driven development of multi-agent systems: A process fragment for agile methods. *Comput. Lang. Syst. Struct., 50,* 159–176.

Wautelet, Y., Heng, S., Kolp, M., & Mirbel, I. (2014). Unifying and extending user story models. In M. Jarke, J. Mylopoulos, C. Quix, C. Rolland, Y. Manolopoulos, H. Mouratidis, & J. Horkoff (Eds.), *Advanced Information*
Systems Engineering - 26th International Conference, CAiSE 2014, Thessaloniki, Greece, June 16-20, 2014. Proceedings (pp. 211–225). Springer volume 8484 of Lecture Notes in Computer Science.