

# OntoMaven - Maven-based Ontology Development and Management of Distributed Ontology Repositories

Adrian Paschke

Corporate Semantic Web, Institute of Computer Science,  
Koenigin-Luise-Str. 24, 14195 Berlin, Germany

[adrian.paschke@inf.fu.berlin.de](mailto:adrian.paschke@inf.fu.berlin.de)

<http://www.corporate-semantic-web.de>

**Abstract.** In collaborative agile ontology development projects support for modular reuse of ontologies from large existing remote repositories, ontology project life cycle management, and transitive dependency management are important needs. The Apache Maven approach has proven its success in distributed collaborative Software Engineering by its widespread adoption. The contribution of this paper is a new design artifact called OntoMaven. OntoMaven adopts the Maven-based development methodology and adapts its concepts to knowledge engineering for Maven-based ontology development and management of ontology artifacts in distributed ontology repositories.

**Keywords:** Semantic Web, Ontology Repositories, Ontology Development, Ontology Engineering, Ontology Modularization, Ontology Management, Ontology Life Cycle Management, Corporate Semantic Web, Agile Knowledge Engineering

## 1 Introduction

Sharing and reusing knowledge in ontology-based applications is one of the main aims in the Semantic Web as well as the Pragmatic Web<sup>1</sup> [17,11,12], which requires the support of distributed ontology management, documentation, validation and testing. As pointed out in [5,13] such ontology development life cycles have a similar structural and logic complexity as distributed software development projects. Agile ontology life cycle management methodologies, such as COLM<sup>2</sup> [10], require the collaboration of knowledge engineers and domain experts. Ontologies are developed and maintained in an iterative and distributed way, which requires the support of versioning [9,14] and modularization [2,3]. Moreover, new aspect-oriented ontology development approaches [16] enable weaving of cross-cutting knowledge concerns into the main ontology model, which requires meta-level descriptions of ontology aspects and management of distributed knowledge models.

<sup>1</sup> <http://www.pragmaticweb.info>

<sup>2</sup> <http://www.corporate-semantic-web.de/colm.html>

In this work we adapt a highly successful method and tool in distributed software engineering project management, namely Apache Maven<sup>3</sup>, for the Maven-based management of distributed ontology repositories. Maven is tool for the (distributed) project management and quality assurance in software engineering projects. The goal of Maven is to automate recurring tasks in software development projects such as management of software artifact in distributed remote and local repositories, versioning, dependency management, documentation, testing, packaging and deployment.

We follow a Design Science research methodology and develop a novel design artifact, called OntoMaven<sup>4</sup>, as a relevant and original contribution in distributed ontology engineering. The approach is grounded in the rigorous methods of Maven's engineering approach, which has already proven its value and success in software engineering. The OntoMaven approach supports ontology engineering in the following ways:

- OntoMaven remote repositories enable distributed publication of ontologies as **ontology development artifacts** on the Web, including their metadata information about life cycle management, versioning, authorship, provenance, licensing, knowledge aspects, dependencies, etc.
- OntoMaven local repositories enable the reuse of existing ontology artifacts in the users' local ontology development projects.
- OntoMaven's support for the different development phases from the design, development to testing, deployment and maintenance provides a flexible life cycle management enabling iterative agile ontology development methods, such as COLM [10], with support for collaborative development by, e.g., OntoMaven's dependency management, version management, documentation and testing functionalities, etc.
- OntoMaven plug-ins provide a flexible and light-weight way to extend the OntoMaven tool with existing functionalities and tools, such as semantic version management (e.g., SVont - Subversion for Ontologies [9,14]), semantic documentation (e.g., SpecGen Concept Grouping [3]), dependency management of aspect-oriented ontology artifacts (e.g. [16]), automated testing (e.g., with the W3C OWL test cases and external reasoners such as Pellet), etc.
- Maven's API allows easy integration of OntoMaven into other ontology engineering tools and their integrated development environments (IDE).

The further paper is structured as follows: Section 2 describes related work. Section 3 describes the design of OntoMaven based on Maven's remote and local repositories, the Project Object Model (POM), plug-ins, and important functional concepts of the solution approach for OntoMaven ontology development. Section 4 proves the feasibility of the proposed concepts with a proof-of-concept implementation of the OntoMaven design artifact. Section 5 compares the OntoMaven functionalities to the tool support of the major existing ontology engineering tools, namely Protege, Swoop and Top Braid Composer. This highlights

---

<sup>3</sup> <http://maven.apache.org/>

<sup>4</sup> <http://www.corporate-semantic-web.de/ontomaven.html>

the strengths of OntoMaven with the open approach to model, manage, and reuse ontology (modules) as **ontology development artifacts** including their metadata descriptions in the POM, dependency management, versioning, documentation, etc. Finally, section 6 summarizes the current OntoMaven work and discusses future research.

## 2 Related Work

There are many existing ontology engineering methodologies and ontology editors available. With its Maven-based approach for structuring the development phases into different goals providing different functionalities during the development project's life cycle, OntoMaven supports in particular agile ontology development methods, such as RapidOWL [1] and COLM [10], as well as development methods which are inherently based on modularization such as aspect-oriented ontology development [16].

According to [6] the most popular ontology editors supporting the Semantic Web ontology languages (RDF, RDFS, OWL, SWRL) are Protege<sup>5</sup>, Swoop<sup>6</sup>, and Top Braid Composer<sup>7</sup>. Other editors support, e.g., visual ontology modeling such as Thematrix Visual Ontology Modeler (VOM)<sup>8</sup>, which enables UML-based ontology modeling based on the OMG Ontology Definition Metamodel (OMG ODM<sup>9</sup>), or lightweight domain specific vocabulary development, such as Leone<sup>10</sup> [14]. While the focus of OntoMaven is on supporting the backend-functionalities in ontology development projects, the focus of these editors is on the support of ontology modeling / representation with a user interface. Non of them is based directly on the Apache Maven concepts and its methods. While all of the editors also provide support for ontology repositories and reuse of existing ontologies by imports, OntoMaven mainly differs in the approach how it manages and declaratively describes the development phases, goals, and artifacts in a Maven Project Object Model (POM). Further implementation-specific and plugin-specific differences are in the underlying details of the provided functionalities of OntoMaven such as POM-based dependency management, semantic versioning, semantic documentation etc. For a comparison see the evaluation section 5.

The W3C Wiki lists several existing ontology repositories<sup>11</sup>. Further ontology repositories are, e.g., COLORE<sup>12</sup> for ontologies written in the ISO Common Logic (CL) ontology languages and Ontohub<sup>13</sup> which maintains a set of heterogeneous ontologies. The current focus of these projects is on collection and listing of

<sup>5</sup> <http://protege.stanford.edu/>

<sup>6</sup> <http://www.mindswap.org/2004/SWOOP/>

<sup>7</sup> [http://www.topquadrant.com/products/TB\\_Composer.html](http://www.topquadrant.com/products/TB_Composer.html)

<sup>8</sup> <http://thematrix.com/tools/vom/>

<sup>9</sup> <http://www.omg.org/spec/ODM/>

<sup>10</sup> leone - <http://www.corporate-semantic-web.de/leone.html>

<sup>11</sup> [http://www.w3.org/wiki/Ontology\\_repositories](http://www.w3.org/wiki/Ontology_repositories)

<sup>12</sup> <http://stl.mie.utoronto.ca/colore/>

<sup>13</sup> <http://ontohub.org/>

existing ontologies. Apart from simple search functionalities there is no support for repository-based ontology development which is the focus of OntoMaven and OntoMaven repositories.

New standardization efforts such as OMG Application Programming Interfaces for Knowledge Bases (OMG API4KB)<sup>14</sup> and OMG OntoIOP aim at the accessibility and interoperability of heterogeneous ontologies via standardized interfaces and semantic transformations defined on the meta-level of the ontology models, e.g., by the Distributed Ontology Language (DOL) [8]. These approaches do not address ontology engineering directly, but can provide a standardized repository back-end for OntoMaven ontology development projects.

### 3 OntoMaven's Design and Concept

This section describes the approach and the concepts of the new design artifact, called OntoMaven, which adapts Apache Maven for the Maven-based ontology development and management of distributed OntoMaven ontology repositories. Maven is not just an automated build tool but also supports software artifact management and quality assurance in software projects. The main required functionalities provided by Maven are:

- source code compilation
- dependency management
- testing with test suites
- automated documentation and reporting
- installation and deployment of generated code

The main design concepts of Maven are:

- The Project Object Model (POM) is the main declarative XML description for managing a project and its development artifacts. Based on the instructions in a POM file Maven automates the different project goals in the life cycle of a software development project.
- Maven plug-ins implement the functionality of the different Maven goals and lead to a modular and extensible architecture of Maven. The plug-ins are executed by Maven using the descriptions in the POM file. Maven has three predefined life cycles, namely the *Clean* life cycle, which cleans the project, the *Default* life cycle, which processes, builds, tests and installs locally or deploys remotely, and the *Site* life cycle, which reports, documents and deploys the created HTML documentation, e.g. on a Web server.
- Maven local and remote repositories manage the used plug-ins and artifacts including support for versioning and dependency management. The general approach is that libraries of existing software frameworks and Maven plug-ins which are required in a software development project are downloaded from distributed remote repositories to the local Maven repository so that

---

<sup>14</sup> [www.omgwiki.org/API4KB/](http://www.omgwiki.org/API4KB/)

Maven can work with these artifacts locally during the development. This distributed approach supports sharing and reuse of existing software artifacts. The information about the used artifacts and their remote addresses (typically a URL) as well as dependency information are described in the POM file of a project. The downloaded artifacts have their own POM files in order to support e.g. transitive dependencies.

In the following subsections we adapt the main concepts of Maven, so that they can be used in ontology development and ontology life cycle management. In particular, we focus on the (distributed) management of knowledge artifacts (ontologies / ontology modules) and their versioning, import and dependency management, documentation, and testing.

### 3.1 Management and Versioning of Ontology Artifacts

One of the design patterns in ontology engineering is the reuse of existing ontologies and ontology modules. Finding ontologies on the Web is supported, e.g., by specialized ontology search engines such as Swoogle<sup>15</sup> and Watson<sup>16</sup>. Since such found ontologies typically cannot be used directly, but need to be modularized, refactored and enhanced, before they can be reused in an ontology development project, there is need for versioning and life cycle management of such ontologies. Furthermore, combinations with other existing ontologies (by ontology match-making and alignment) might lead to transitive dependencies which need to be described and managed. OntoMaven therefore adopts Maven's artifact concept. It describes and manages ontologies as ontology artifacts in a Maven Project Object Model (POM). The typical steps to add an ontology (module) as an OntoMaven artifact to a POM are:

1. Find ontology module(s)
2. Select the right module and version
3. Analyse and resolve dependencies of the modules
4. Declaratively describe the ontology artifact in a POM

Many ontology languages support imports or integration of distributed ontologies. For instance, the W3C Web Ontology Language (OWL) therefore has a specialized `owl:import` statement and the ISO Common Logic standard supports modularized imports by a segregation semantics which distinguishes the universe of discourse of the main ontology from the segregated universe of discourse of the imported ontology.

Typical recurring tasks which are automated by OntoMaven are in such modular import and reuse scenarios are,

- check the existence of the imported ontology (module) referenced by the defined URI in the import statement (and find alternative URLs from pre-configured repositories if the ontology does is not found at the import URI).

<sup>15</sup> <http://swoogle.umbc.edu/>

<sup>16</sup> <http://kmi-web05.open.ac.uk/WatsonWUI/>

- management of ontologies / ontology modules as ontology artifacts in Maven repositories including their metadata descriptions such as versioning information.
- download of ontology artifacts from remote repositories (including transitive imports) to a local development repository in order to support offline development of ontologies

Another important aspect in the agile and collaborative development of ontologies is the support for version management. Typical requirements are maintaining consistency and integrity, as well as provenance data management (e.g. authorship information) throughout the version history. Compared to version management solutions in software engineering which just maintain syntactic versions on the level of code line differences, the difficulty in knowledge engineering is that the versions and their differences need to be managed on a semantic level. Different syntactic versions of an ontology knowledge model might still have the same semantic interpretation. That is, a semantic version management system needs to compute the semantic difference, e.g., to detect and resolve version conflicts. The approach in OntoMaven is based on the ontology versioning tool *SVont*<sup>17</sup>, which is an extensions of the version management tool Subversion. [9,14]

### 3.2 Import and Dependency Management

OntoMaven adopts the dependency management of Maven by describing the dependencies in an ontology development project from existing ontology artifacts in a POM. This is illustrated in the following listing from a POM example:

```
<project>
...
  <dependencies>
    <dependency>
      <groupId>de.onto.maven</groupId>
      <artifactId>TimeOntologie</artifactId>
      <version>1.0</version>
    </dependency>
  ...
</dependencies>
</project>
```

The listing describes a dependency of an ontology artifact identified by the ID `TimeOntologie` version 1.0 which belongs to the group `de.onto.maven`.

It is possible to define multiple repositories in which OntoMaven will look for dependent ontology artifacts. If the defined ontology artifact is not found in the repository, users will be informed. They can start a search with the ontology name and URI defined in the artifact description in configured ontology search engines. OntoMaven supports Swoogle and Watson. The found ontologies will can be added as new ontology artifacts to an OntoMaven repository.

<sup>17</sup> <http://www.corporate-semantic-web.de/svont.html>

### 3.3 Documentation

Despite first automated ontology matching and alignment approaches developing high quality ontologies still remains a manual knowledge modeling effort, where domain experts work together with knowledge engineers. Documentation is an important step which facilitates this work and in particular makes maintenance and reuse easier. The typical distinction is into *user documentation* and *technical documentation*. While the former supports the users of an ontology, e.g. in their task to populate the ontology with instance data, the latter, technical documentation, supports the ontology developer.

There exist several ontology documentation tools such as OWLDoc<sup>18</sup>, VocDoc<sup>19</sup>, or SpecGen<sup>20</sup>, which document ontologies on a technical level (much like tools such as JavaDoc in Java programming). Unfortunately, the lack of good documentation in the published ontologies on the Web makes reuse difficult, because the analysis and decision process on the applicability of a candidate ontology becomes very time-consuming. Therefore, additional automated support needs to be provided, e.g., for analysing larger ontologies on an abstract level by creating concept groupings which reduce the complexity of the ontology model. This groupings and summarizations of concepts provides the reader with an easier way to understand the ontology vocabulary. Typically such concept groups are additionally presented in easy to understand visualization formats, e.g. by tools such as OWLViz<sup>21</sup>, OntoGraf<sup>22</sup>, Sonivis OWL plugin<sup>23</sup>, SOVA<sup>24</sup>, TGVizTab<sup>25</sup>, or as UML models, e.g. by VOM<sup>26</sup>. Such concept grouping and visualisations can be used in the user documentation of an ontology. [3]

Maven supports the documentation phase and provides goals for creating and publishing automated reports. In OntoMaven we make use of SpecGen and the SpecGen extension<sup>27</sup> for automated concept grouping, in order to create the technical and user documentation in an OntoMaven plugin which is executed by the `mvn site` command.

### 3.4 Testing

Testing is an important phase in the ontology life cycle. In particular, in agile iterative development processes testing allows detecting inconsistencies, anomalies, improper design, as well as validation against, e.g., the intended results of domain experts' competency questions which are represented as ontology test

<sup>18</sup> <http://docpp.sourceforge.net/>

<sup>19</sup> <http://kantenwerk.org/vocdoc/>

<sup>20</sup> <https://github.com/specgen/specgen>

<sup>21</sup> <http://www.co-ode.org/downloads/owlviz/>

<sup>22</sup> <http://protegewiki.stanford.edu/wiki/OntoGraf>

<sup>23</sup> <http://www.corporate-semantic-web.de/ontology-modularization-framework.html>

<sup>24</sup> <http://protegewiki.stanford.edu/wiki/SOVA>

<sup>25</sup> <http://users.ecs.soton.ac.uk/ha/TGVizTab/>

<sup>26</sup> <http://thematix.com/tools/vom/>

<sup>27</sup> <http://www.corporate-semantic-web.de/concept-grouping.html>

cases. Maven supports a testing phase in which automated tests are executed and the results are reported by the Maven command `mvn test`.

The W3C OWL recommendation<sup>28</sup> defines a collection of test cases with different test types and tests. As standard test types OntoMaven by default supports the W3C OWL test cases *syntax checker*, *consistency checker*, and *entailment test*. The produced test results are compliant to the W3C recommendation and the created test reports show if the ontology model is **consistent**, **inconsistent** or if the result is **unknown**. Further test types can be implemented as Maven plug-ins and added to the OntoMaven projects test suites by the user.

## 4 Proof-of-Concept Implementation - OntoMaven PlugIns

The implementation of OntoMaven extends and adapts Maven [7], so that it supports the management of ontology modules in Maven repositories. This section describes how the OntoMaven approach and the above described concepts are implemented using *Maven repositories* and the *Maven plug-in* extension mechanism. A Maven plug-in is a collection of one or more goals. For instance, the plug-in `archetype` implements the goals `create` and `generate` which create a Maven project. The implementation of a Maven plug-in is done in an *Maven Plane Old Java Object (MOJO)*. Maven supports the automated generation of a Mojo project with the goal `generate` in the Maven plugin `Archetype`:

```
Mvn archetype:generate -DinteractiveMode=false
-DarchetypeArtifactId=maven-archetype-mojo -DgroupId=[] -DartifactId=[]
```

By defining the `groupId` and the `artifactID` the command creates a Mojo project with a Mojo class which is used for the plug-in implementation.

In the OntoMaven approach, the phases and goals, which the plug-in implements, are defined by JavaDoc annotations in the source code of the Mojo class. For instance, the following annotations define that the implemented plug-in is used in the phase `test` and that it has a goal called `test-syntax`:

```
@phase test //plug-in used in test phase
@goal test-syntax // goal with the name "test-syntax"
```

Parameters are used to configure the plug-in execution. For instance, the following code snippet defines a required parameter `complianceMode` with the default value `strict`:

```
*@parameter expression = complianceMode
*default-value="strict"
*@required
```

Such plug-in parameters can be configured in a POM.xml file or directly when calling a goal, e.g. `mvn ... -DcomplianceMode=strict`. An implemented plug-in can be installed using Maven `mvn install` and the plug-in goals can be integrated into the POM.xml of an OntoMaven project, as the following example listing shows for the plug-in `SVontPlugin` and the goal `semantic-diff`:

<sup>28</sup> <http://www.w3.org/TR/owl-test/>

```

<build>
  <plugins>
    <plugin>
      <groupId>de.csw.ontomaven</groupId>
      <artifactId>SVontPlugin</artifactId>
      <version>1.0-SNAPSHOT </version>
      <executions>
        <execution>
          <goals>
            <goal>semantic-diff</goal>
          </goals>
        </execution>
      </executions>
    </plugin>
  </plugins>
</build>

```

The following subsections provide further details about the proof-of-concept implementations of the main plug-ins in OntoMaven. We first describe the OntoMaven repositories which are the persistence and back-end layer for storing and managing ontologies.

#### 4.1 OntoMaven Repositories

OntoMaven can use all Maven compliant repositories. One of the strengths of Maven is that it uses a folder structure following a standard folder layout for its repositories; sources are in `$basedir/src/main/java`, resources in `$basedir/src/main/resource`, tests in `$basedir/src/test`, classes in `$basedir/target/classes`, and packaged libraries in `$basedir/target/`.

For the OntoMaven proof-of-concept implementation we adapted the Apache Archiva Build Artifact Repository Manager<sup>29</sup> as a managing tool providing a user interface for the OntoMaven repositories. It supports finding and managing OntoMaven artifacts. Figure 1 shows the upload user interface.

**Upload Artifact**

Group Id\*:

Artifact Id\*:

Version\*:

Packaging\*:

Classifier:

Generate Maven 2 POM

Artifact File\*:

POM File:

Repository Id:

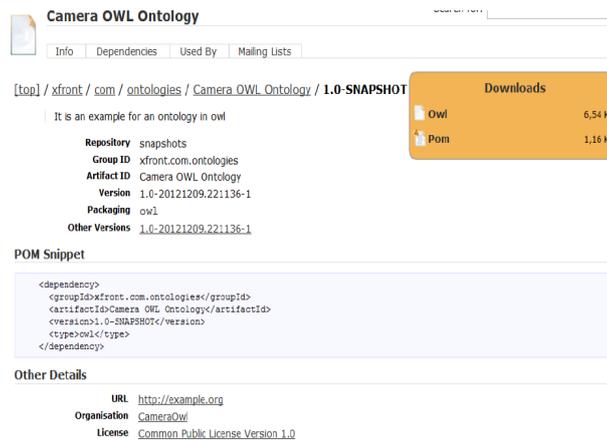
**Fig. 1.** Archiva User Interface - Ontology Artifact Upload

Via this form an ontology can be uploaded to an OntoMaven repository together with its POM file. The artifact's metadata contains information about

<sup>29</sup> <http://archiva.apache.org/>

the group id, artifact id, version, packaging and optional additional classifier information. The POM provides all necessary information about the artifact and its dependencies. In OntoMaven these dependencies are used to describe (transitive) imports from an ontology, which are resolved by the OntoMvnImport plug-in (see subsection 4.2) and are defined in the ontology's POM file.

Figure 2 gives an example of the Archiva user interface showing the management information of an ontology artifact called **Camera OWL Ontology**. Under the interface menu link **Dependencies** the dependencies of this ontology can be found.



**Camera OWL Ontology**

Info Dependencies Used By Mailing Lists

[top] / xfront / com / ontologies / Camera OWL Ontology / 1.0-SNAPSHOT

It is an example for an ontology in owl

Repository	Group ID	Artifact ID	Version	Packaging	Other Versions
snapshots	xfront.com.ontologies	Camera OWL Ontology	1.0-20121209.22:1136-1	owl	1.0-20121209.22:1136-1

Downloads	Size
Owl	6.54 kB
Pom	1.16 kB

**POM Snippet**

```
<dependency>
  <groupId>xfront.com.ontologies</groupId>
  <artifactId>Camera OWL Ontology</artifactId>
  <version>1.0-SNAPSHOT</version>
  <scope>owl</scope>
</dependency>
```

**Other Details**

URL <http://example.org>  
 Organisation [cameraOwl](#)  
 License [Common Public License Version 1.0](#)

**Fig. 2.** Archiva User Interface - Management of Ontology Artifact

Once managed in an online OntoMaven repository, an ontology artifact can be used in any OntoMaven ontology development project. The following listing gives an example how a remote repository can be configured and a dependency to an ontology artifact (here the Camera-OWL-Ontology) can be defined in the POM.xml document of a project.

```
<profiles>
  <profile>
    <id>2</id>
    <activation>
      <activeByDefault>true</activeByDefault>
    </activation>
    <repositories>
      <repository>
        <snapshots>
          <enabled>true</enabled>
        </snapshots>
        <id>snapshots</id>
        <name>OntoMaven Snapshot Repository</name>
        <url>www.corporate-semantic-web.de/repository/snapshots/</url>
      </repository>
    </repositories>
  </profile>
</profiles>
```

```

<dependencies>
  <dependency>
    <groupId>xfront.com.owl.ontologies</groupId>
    <artifactId>Camera-OWL-Ontology</artifactId>
    <version>1.0-SNAPSHOT</version>
    <type>owl</type>
  </dependency>
</dependencies>

```

## 4.2 OntoMvnImport

This plug-in implements the imports of Ontologies into the Maven repositories. It also checks if the import statements in the ontology including transitive imports can be resolved. Therefore, it maintains an updated list of reference URIs to the ontology resources loaded to the repository. This list follows the OASIS XMLCatalog standard which also specifies a technique for the automated replacement of external references in XML documents. A XML parser validates if defined replacement rules in an XMLCatalog apply to the references in the validated XML document and in case they apply, it replaces these references with the references defined in the XMLCatalog. In OntoMaven we use this automated replacement technology to replace the URI references to imported **external ontologies** with references to the **internal ontology artifacts**, which are locally managed in an OntoMaven repository after they have been loaded by the plug-in. This replacement approach avoids the continuous import and use of external ontologies during an OntoMaven development project. As an example the following catalog entry defines that the original reference to the imported ontology `example.owl` can be replaced by the URI referencing the stored ontology artifact in the local repository.

```

<?xml version="1.0" encoding="UTF-8"?>
<catalog xmlns="urn:oasis:names:tc:entity:xmlns:xml:catalog">
  <system systemId="www.example.com/example.owl" uri="src/resource/owl"/>
</catalog>

```

After the first loading of an ontology as repository artifact, including all transitive imports which are resolved and stored as dependent ontology artifacts, by the OntoMvnImport plug-in, the plug-in always checks if there is an ontology artifact listed in the XMLCatalog. If there is an existing reference to an ontology artifact, it will use it instead of any externally referenced ontology. A special situation is, if the import statement cannot be resolved, e.g. because the ontology is no longer existent under the given reference. In this case the plug-in notifies the user.

The following listing shows how the plug-in can be used in a OntoMaven POM. In the `configuration` it defines the input ontology and sets the `local` parameter to `true`, indicating that the ontology should be loaded to the local repository and that the local version of the ontology should be used.

```

<build> <plugins> <plugin>
  <groupId>de.csw.ontomaven</groupId>
  <artifactId>OntoMvnImport</artifactId>

```

```

<version>1.0-SNAPSHOT</version>
<configuration>
  <owlfile>src/resource/reputation.owl</owlfile>
  <local>true</local>
</configuration>
<executions>
  <execution>
    <goals>
      <goal>owlimport</goal>
    </goals>
  </execution>
</executions>
</plugin> </plugins> </build>

```

### 4.3 OntoMvnSvn

The `OntoMvnSvn` plug-in provides ontology versioning support for OntoMaven. As discussed in section 3.1 standard (code) versioning tools such as Concurrent Version System (CVS) and Subversion cannot be directly used, because version differences are only computed syntactically, but not semantically, as it is required for versioning interpreted knowledge models such as ontologies.

We therefore implemented an extension to Subversion called `SVont`<sup>30</sup> [9] which can compute semantic differences<sup>31</sup> and which can version ontologies. `SVont` supports typical Subversion commands such as `checkout`, `status`, `diff`, `commit`, and `info`. In the following the implementation of the commands `status` and `diff` are described in more detail.

For the `status` command the `OntoMvnSvn` plug-in first does a repository checkout of the ontology to a temporal folder and then compares the repository version with the currently developed working version. The possible status results are *identical* and *changed*. This status information is used by the plug-in to either report *ontology changed* or *ontology is up-to-date*.

The `diff` command first does a checkout of the committed repository version. It then computes the semantic difference to the working version. Therefore, the OntoMaven implementation is using the semantic difference computation implemented by `SVont`. The `OntoMvnSvn` plug-in goal `diff` additionally performs a dependency analysis. It lists all dependencies, e.g. dependencies by domain and range properties and subclass relations as the following example shows:

```

----- DIFF INFORMATION -----
Ontology File : ..\..\..\..\..\camera.owl
===== ACTUAL CHANGES =====
Axioms were added to the repository, or deleted from the working
copy.
SubClassOf(<www.xfront.com/owl/ontologies/camera/#Money>
owl:Thing)
Declaration(Class(<www.xfront.com/owl/ontologies/camera/#Money>))
=====
----- MORE INFO -----
The above changes of the OWL classes are dependent on the following
axiom.
currency <----- DataProperty (Domain)
cost <----- ObjectProperty (Range)
-----

```

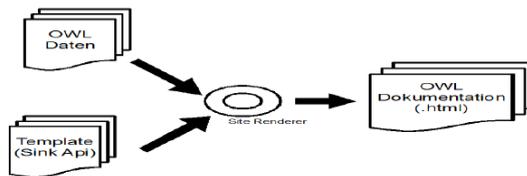
<sup>30</sup> <http://www.corporate-semantic-web.de/svont.html>

<sup>31</sup> for the description logic  $\mathcal{EL}$

#### 4.4 OntoMvnReport

The plug-in is implemented as Maven report plug-in<sup>32</sup>. The goal `site` of this plug-in creates four different documentations about the ontology - a general *project documentation*, an *ontology report summary*, a *technical report*, and an *ontology visualization*.

For rendering the ontology information into an HTML documentation it uses the Sink API<sup>33</sup> as illustrated in Figure 3.



**Fig. 3.** Simplified Diagram for the Report Rendering in the OntoMvnReport Plug-In

With the Maven Site Descriptor<sup>34</sup> the layout and content menus in OntoMaven reports can be adapted.

The general *ontology project documentation* is created from the description of the ontology artifact in the POM file, which includes project metadata about, e.g. the project, project team, dependencies, the plug-ins, issues, source repositories, licenses, ontology developers and their roles, etc. The following listing gives an example of typical project's metadata in a POM which is used in the project documentation which can be selected from the report menu (see figure 4):

```

<description>here's the descripton of an ontology </description>
<organization>
  <name>Corporate Semantic Web, Freie Universitt Berlin</name>
  <url>www.corporate-semantic-web.de</url>
</organization>
<inceptionYear>2013</inceptionYear>
<licenses>
  <license>
    <name>LGPL-3.0</name>
    <url>www.gnu.org/licenses/lgpl.txt</url>
  </license>
</licenses>
<developers>
  <developer>
    <name>Adrian Paschke</name>
    <email>paschke@inf.fu-berlin.d</email>
    <organization>Corporate Semantic Web</organization>
    <organizationUrl>www.corporate-semantic-web.de/
    </organizationUrl>
    <roles>
      <role>developer</role>
    </roles>
  </developer>
</developers>
  
```

<sup>32</sup> <http://docs.codehaus.org/display/MAVENUSER/Write+your+own+report+plugin>

<sup>33</sup> <http://maven.apache.org/doxia/doxia/doxia-sink-api/>

<sup>34</sup> <http://maven.apache.org/plugins/maven-site-plugin/examples/sitedescriptor.html>

```

</developer>
</developers>

```



**Fig. 4.** OntoMaven Project and Ontology Documentation Menu

The *ontology report summary* is created by the goal `ontologyreport`. Figure 5 shows an example ontology summarization which gives an overview about the general description, the format, the semantic profile, imported ontologies and a summary about the ontology's statistics (number of classes, datatype properties, object properties, etc.).

Imports :	
found no import	
	number
Total Number of Classes	13
Total Number of Datatype Properties	6
Total Number of Object Properties	7
Total Number of Annotations Properties	1
Total Number of Individuals	2

**Fig. 5.** OntoMaven Ontology Report Summary

For the documentation of the ontology, the plug-in uses existing automated ontology documentation tools. We have integrated the SpecGen ontology documentation tool which creates a HTML page containing detailed information about the classes and the properties. We further extended SpecGen with various algorithms for creating structure based concept groupings. [3,4] These groupings are used as basis for a visual documentation of the ontology. To support this process of creating such concept groups for the documentation of ontologies we extended the SpecGen tool with an automatic concept grouping functionality<sup>35</sup> and embedded it for the OntoMaven documentation.

<sup>35</sup> <http://www.corporate-semantic-web.de/concept-grouping.html>

A more detailed insight is given by the *technical ontology report*, which is created by the goal `technicalreport`. This goal produces a listing of classes and properties as shown in figure 6. By clicking on a particular class or property the technical details about it are shown.

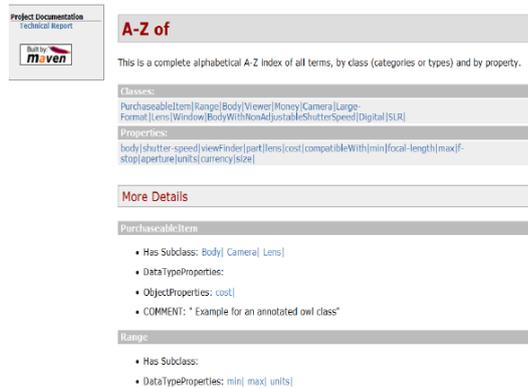


Fig. 6. OntoMaven Technical Report

The goal `visualizer` produces a network visualization of the ontology concepts (classes) and its relations using different graph visualizations <sup>36</sup>. Figure 7 gives a visualization example.

The following listing shows how to use the `OntoMvnReport` plugin in a project POM.xml

```
<reporting><plugins><plugin>
<groupId>de.nbi.MvnOnt</groupId>
<artifactId>MvnOwlReport</artifactId>
<version>1.0-SNAPSHOT</version>
<reportSets>
  <reportSet>
    <configuration></configuration>
    <reports>
      <report>ontologyreport</report>
      <report>technicalreport</report>
      <report>visualizer</report>
    </reports>
  </reportSet>
</reportSets></plugin></plugins></reporting>
```

The produced reports can be found in the Maven folder `target/site`.

#### 4.5 OntoMvnTest

The *OntoMvnTest* plug-in implements functionalities for the test phase. The plug-in executes the configured test using the goal `test`. It is also used internally

<sup>36</sup> <http://www.corporate-semantic-web.de/ontology-modularization-framework.html>

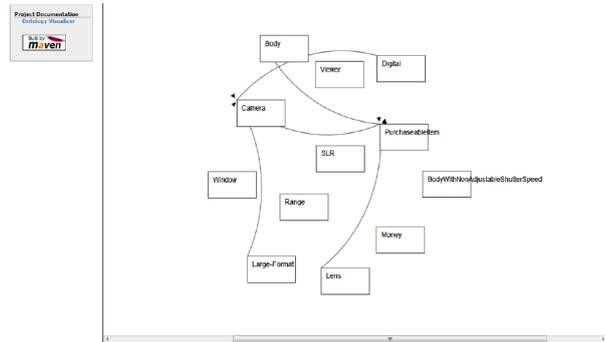


Fig. 7. OntoMaven Ontology Visual Report

in other phases such as the `package` goal. The plug-in implementation uses the Pellet reasoner<sup>37</sup> to execute the ontology test cases.

As default test suites the plug-in supports the W3C OWL Test Cases<sup>38</sup>. This test collection contain different types of test cases, such as a test that determines and returns the OWL sublanguage, tests for inconsistency checks, and entailment tests, which test if the intended conclusions (represented by an output ontology) are entailed in the input ontology model. For instance, the intended entailment test result values are **Entailment** (positive test result) or **NoEntailment** (negative test result). The following listing shows how the plug-in can be used in a POM.xml.

```
<build>
<plugins>
  <plugin>
    <groupId>de.csw.MvnOnt</groupId>
    <artifactId>MvnOwlTest</artifactId>
    <version>1.0-SNAPSHOT</version>
    <configuration>
      <owlfile>owl/1a.owl</owlfile>
    </configuration>
    <executions>
      <execution>
        <goals>
          <goal>owltest</goal>
        </goals>
      </execution>
    </executions>
  </plugin>
  <plugin>
    <groupId>de.csw.MvnOnt</groupId>
    <artifactId>MvnOwlEntailment</artifactId>
    <version>1.0-SNAPSHOT</version>
    <configuration>
      <premise_file>owl/1a.owl</premise_file>
      <conclusion_file>owl/1aconclusion.
        owl</conclusion_file>
    </configuration>
    <executions>
```

<sup>37</sup> <http://clarkparsia.com/pellet/>

<sup>38</sup> <http://www.w3.org/TR/owl-test/>

```

        <execution>
          <goals>
            <goal>owlentailment</goal>
          </goals>
        </execution>
      </executions>
    </plugin></plugins></build>

```

## 5 Evaluation

OntoMaven is not a full ontology development tool as e.g. Protege, Swoop and Top Braid Composer, which provide a development user interface. Instead OntoMaven’s has its strength in the management of distributed ontology modules including support for reuse (transitive imports), dependency management and collaboration (semantic versioning).

Table 1 compares OntoMaven to the most used ontology development tools by its functional support in typical ontology engineering life cycles.

**Table 1.** Functional Comparison of OntoMaven with Ontology Development Tools

	OntoMaven	Protege	Swoop	Top Braid Composer
Repositories	yes (local and remote)	yes (local and remote)	no	yes (by Allegro Graph 4 PlugIn)
Reuse (Import)	yes (dependency management)	yes	yes	yes
Collaboration Support (Versioning)	yes (semantic diff)	no	no	no
Documentation	yes (text and visual)	yes (text and visual)	yes (only text)	yes (text and visual in Maestro version)
Testing	yes	yes	yes	yes
Extensibility	yes	yes (many existing plugins)	yes	yes (commercial)

## 6 Conclusion

Apache Maven is a widespread and highly successful tool in Software Engineering for build automation and development project life cycle management. This paper has adapted the Maven approach and concepts for Knowledge Engineering in (agile) ontology development. The contribution is a new design artifact called OntoMaven which has been implemented as a proof-of-concept implementation.

OntoMaven uses a Project Object Model (POM) XML file to describe the ontology project being developed, its dependencies on other external ontology modules, the development life cycle order, directories, and required plug-ins. It comes with pre-defined targets for performing certain well-defined tasks in typical agile ontology development phases.

OntoMaven dynamically downloads distributed ontologies and Maven plugins from one or more OntoMaven remote repositories and stores them in a local

cache, the local repository. This local cache of downloaded artifacts can be further updated with ontology artifacts created by local projects as well as developed artifacts can be uploaded to remote OntoMaven repositories, so that they can be shared and reused.

OntoMaven is built using Maven’s plugin-based architecture that allows it to make use of any application controllable through standard input. The proof-of-concept of OntoMaven implements several useful plugins which interface with existing ontology development tools and functionalities such as the plug-ins *OntoMvnImport*, *OntoMvnSVN*, *OntoMvnReport*, and *OntoMvnTest*.

In future research we plan to study the benefits of the OntoMaven support for ontology engineers in real agile ontology development projects and compare it with other development approaches. We plan to quantify the results on the basis of ontology development costs (in person months) and compare them against the estimated cost models (based on our work on Agile ONTOCOM 2<sup>39</sup> [15]).

In our aspect-oriented ontology development research [16] we plan to make use of OntoMaven for the distributed management of ontology modules and their description as *aspect-oriented ontology artifacts* in the POM.

## 7 Acknowledgements

This work has been partially supported by the InnoProfile project ”Corporate Semantic Web” funded by the German Federal Ministry of Education and Research (BMBF).

## References

1. Sören Auer. The rapidowl methodology—towards agile knowledge engineering. In *WETICE*, pages 352–357, 2006.
2. G. Coskun, M. Luczak-Rösch, R. Heese, and A. Paschke. Applying ontology modularization for corporate ontology engineering. In *Proceedings of the Intl. Conference on Semantic Systems (I-SEMANTICS 2009)*, pages 669–674, Graz, Austria, Sep 2009.
3. Gökhan Coskun, Mario Rothe, and Adrian Paschke. Ontology content “at a glance”. In M Donnelly and G Guizzardi, editors, *Proceedings of the 7th International Conference on Formal Ontology in Information Systems*, pages 147–159, Graz, Austria, 2012. IOS Press.
4. Gökhan Coskun, Mario Rothe, Kia Teymourian, and Adrian Paschke. Applying community detection algorithms on ontologies for identifying concept groups. In *WoMO*, pages 12–24, 2011.
5. Antonio De Nicola, Michele Missikoff, and Roberto Navigli. A software engineering approach to ontology building. *Inf. Syst.*, 34(2):258–275, April 2009.
6. M. Rahamatullah Khondoker and Paul Mueller. Comparing ontology development tools based on an online survey. In *Proceedings of the World Congress on Engineering 2010 (WCE 2010)*, 2010.

<sup>39</sup> <http://www.corporate-semantic-web.de/corporate-ontology-engineering.html>

7. Onur Kilic. Erweiterung von maven zur toolbasierten verwaltung von ontologiemodulen, 2013.
8. Christoph Lange, Oliver Kutz, Till Mossakowski, and Michael Grüninger. The distributed ontology language (dol): Ontology integration and interoperability applied to mathematical formalization. *CoRR*, abs/1204.5093, 2012.
9. Markus Luczak-Rsch, Gkhan Coskun, Adrian Paschke, Mario Rothe, and Robert Tolksdorf. Svont - version control of owl ontologies on the concept level. In Klaus-Peter Fhnrich and Bogdan Franczyk, editors, *GI Jahrestagung (2)*, volume 176 of *LNI*, pages 79–84. GI, 2010.
10. Markus Luczak-Rsch and Ralf Heese. Managing ontology lifecycles in corporate settings. In Tassilo Pellegrini, Sren Auer, Klaus Tochtermann, and Sebastian Schaffert, editors, *Networked Knowledge - Networked Media*, volume 221 of *Studies in Computational Intelligence*, pages 235–248. Springer Berlin Heidelberg, 2009.
11. Adrian Paschke and Harold Boley. Rule responder: Rule-based agents for the semantic-pragmatic web. *International Journal on Artificial Intelligence Tools*, 20(6):1043–1081, 2011.
12. Adrian Paschke, Harold Boley, Alexander Kozlenkov, and Benjamin Larry Craig. Rule responder: Ruleml-based agents for distributed collaboration on the pragmatic web. In *ICPW*, pages 17–28, 2007.
13. Adrian Paschke, Gkhan Coskun, Ralf Heese, Markus Luczak-Rsch, Radoslaw Oldakowski, Ralph Schfermeier, and Olga Streibel. Corporate semantic web: Towards the deployment of semantic technologies in enterprises. In Weichang Du and Faezeh Ensan, editors, *Canadian Semantic Web*, pages 105–131. Springer US, 2010.
14. Adrian Paschke, Gökhan Coskun, Dennis Hartrampf, Ralf Heese, Markus Luczak-Rösch, Mario Rothe, Radoslaw Oldakowski, Ralph Schäfermeier, and Olga Streibel. Realizing the corporate semantic web: Prototypical implementations. TR-B-10-05:1–49, 02/2010 2010.
15. Adrian Paschke, Gökhan Coskun, Harasic Marko, Ralf Heese, , Radoslaw Oldakowski, Ralph Schäfermeier, Olga Streibel, Kia Teymourian, and Alexandru Todor. Corporate semantic web report vi: Validation and evaluation. TR-B-13-01:1–64, 01/2013 2013.
16. Ralph Schäfermeier and Adrian Paschke. Towards a unified approach to modular ontology development using the aspect-oriented paradigm. In *7th International Workshop on Modular Ontologies (WoMO 2013)*, 2013.
17. Hans Weigand and Adrian Paschke. The pragmatic web: Putting rules in context. In *RuleML*, pages 182–192, 2012.