Deliverable 2.2
Semantic annotation tool

Author(s):
Konstantinos Tarabanis (UOM)
Efthimios Tambouris (UOM)
Maria Zotou (UOM)

Editor(s):
Konstantinos Tarabanis (UOM)
Efthimios Tambouris (UOM)

Responsible Organisation: UOM

Version-Status: V1 Final

Submission date: 31/12/2016

Dissemination level: PU
### Deliverable factsheet

<table>
<thead>
<tr>
<th>Project Number:</th>
<th>562236-EPP-1-2015-1-EL-EPPKA3-PI-FORWARD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Acronym:</td>
<td>PBL3.0</td>
</tr>
<tr>
<td>Project Title:</td>
<td>Integrating Learning Analytics and Semantics in Problem Based Learning</td>
</tr>
<tr>
<td>Title of Deliverable:</td>
<td>D2.2 – Semantic Annotation Tool</td>
</tr>
<tr>
<td>Work package:</td>
<td>WP2 – PBL_LA semantic model design and annotation tool</td>
</tr>
<tr>
<td>Due date according to contract:</td>
<td>31/12/2016</td>
</tr>
<tr>
<td>Editor(s):</td>
<td>UOM</td>
</tr>
<tr>
<td>Contributor(s):</td>
<td>UOM</td>
</tr>
<tr>
<td>Reviewer(s):</td>
<td>AAU, OUNL, UAH, BOC</td>
</tr>
<tr>
<td>Approved by:</td>
<td>All Partners</td>
</tr>
</tbody>
</table>

**Abstract:**

This report aims to present the semantic annotation environment that was developed in order to foster mainstreamed content discoverability and linkage of learning resources with the Web of Data. The environment can be used by all learning content providers such as MOOC providers, OER providers, educational institutions, national organizations etc., independent of domain and sector.

The report provides the review carried out on the functionalities supported by existing open-source software tools that are available for annotating educational resources. The review results led to the decision to develop a user-friendly web-based
interface because no existing tool can be used as-is for the project semantic annotation environment’s needs (e.g. automatic annotation, structured input of data based on the SALO model, usage of educational RDF vocabularies, export of RDF data for further re-use, linking of data with other resources etc.). The semantic annotation environment will allow the visualization of the SALO model’s concepts in a structured form that can be used by any stakeholder independent of the domain and technical background.

The report also presents a review carried out on existing RDF vocabularies that focus on or include classes and properties of elements within the educational domain. A sub-set of RDF classes and properties was selected to describe all elements of the SALO model and learning resources and the relations between them so that they are understandable by humans and computer machines.

The semantic annotation environment that was developed includes four main tools. The semantic annotator supports the description of learning objects and resources in accordance with the concepts of the SALO model and automatic semantic annotation of each concept. This data is exported in Turtle format for future re-use and also stored to an RDF store for future discoverability. The semantic publisher performs automatic linkage of the annotated learning objects with relevant resources such as DBpedia, ESCO etc. when such links are identified, and it automatically publishes the semantically annotated data as linked data for further linkage with the Web of data. The semantic browser allows users to search existing learning objects and their visualization as published linked data. Finally, the semantic enricher supports automatic provision of published linked data from the PBL_LA framework designed in D1.3 based on the PBL step entered for the learning object during annotation.

**Keyword List:**
Learning semantics, semantic annotation, RDF vocabularies, linked data
## Consortium

<table>
<thead>
<tr>
<th>Role</th>
<th>Name</th>
<th>Short Name</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Coordinator, academic partner</td>
<td>University of Macedonia</td>
<td>UOM</td>
<td>Greece</td>
</tr>
<tr>
<td>2. Technology enhanced learning expert</td>
<td>Open University of the Netherlands</td>
<td>OUNL</td>
<td>Netherlands</td>
</tr>
<tr>
<td>3. PBL expert</td>
<td>Aalborg University</td>
<td>AAU</td>
<td>Denmark</td>
</tr>
<tr>
<td>4. Semantic and Learning analytics expert</td>
<td>University of Alcala</td>
<td>UAH</td>
<td>Spain</td>
</tr>
<tr>
<td>5. IT partner</td>
<td>BOC</td>
<td>BOC</td>
<td>Austria</td>
</tr>
</tbody>
</table>
Revision History

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Revised by</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>V01</td>
<td>15/11/2016</td>
<td>UOM</td>
<td>First version of semantic annotation environment complete</td>
</tr>
<tr>
<td>V02</td>
<td>25/11/2016</td>
<td>UOM</td>
<td>Contributions added</td>
</tr>
<tr>
<td>V03</td>
<td>22/12/2016</td>
<td>UOM</td>
<td>Contributions added</td>
</tr>
<tr>
<td>V04</td>
<td>27/12/2016</td>
<td>UOM</td>
<td>Contributions from partners added</td>
</tr>
<tr>
<td>V1</td>
<td>28/12/2016</td>
<td>UOM</td>
<td>Final version reviewed and ready for submission</td>
</tr>
</tbody>
</table>

Statement of originality:
This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Disclaimer:
This project has been funded with support from the European Commission. This publication reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.
Table of Contents

DELIVERABLE FACTSHEET ........................................................................................................... 2
CONSORTIUM .............................................................................................................................. 4
REVISION HISTORY .................................................................................................................... 5
LIST OF FIGURES ......................................................................................................................... 8
LIST OF TABLES ........................................................................................................................... 9
LIST OF ABBREVIATIONS ........................................................................................................... 10
EXECUTIVE SUMMARY ............................................................................................................. 11

1 INTRODUCTION ......................................................................................................................... 13
   1.1 SCOPE ................................................................................................................................ 13
   1.2 AUDIENCE ......................................................................................................................... 13
   1.3 STRUCTURE ......................................................................................................................... 13

2 METHODOLOGY ......................................................................................................................... 14

3 SEMANTIC ANNOTATION ........................................................................................................... 15
   3.1 REVIEW OF SEMANTIC ANNOTATION TOOLS .................................................................. 16
      3.1.1 LOCO-Analyst .............................................................................................................. 17
      3.1.2 DBpedia Spotlight ..................................................................................................... 17
      3.1.3 Pundit Annotator ....................................................................................................... 19
      3.1.4 AIDA .......................................................................................................................... 20
      3.1.5 Apache Stanbol .......................................................................................................... 21
      3.1.6 Open Calais ................................................................................................................ 22
      3.1.7 CWRC-Writer ............................................................................................................ 23
      3.1.8 Marvin text annotator ............................................................................................... 24
      3.1.9 Federated knOwledge eXtraction Framework (FOX) .................................................. 25
      3.1.10 RDFazer .................................................................................................................... 27
   3.2 SEMANTIC ANNOTATION ENVIRONMENT REQUIREMENTS SPECIFICATION ............. 27

4 SEMANTIC VOCABULARIES FOR SEMANTIC ANNOTATION IN EDUCATION .................... 32
   4.1 SALO MODEL ....................................................................................................................... 32
   4.2 REVIEW OF RDF VOCABULARIES IN EDUCATION ......................................................... 33
      4.2.1 LOM vocabulary .......................................................................................................... 33
      4.2.2 LRMI vocabulary ......................................................................................................... 35
      4.2.3 AIISO vocabulary ........................................................................................................ 35
      4.2.4 ESCO vocabulary ........................................................................................................ 36
      4.2.5 IMS – LD vocabulary .................................................................................................. 37
      4.2.6 Teaching Core vocabulary ......................................................................................... 37
      4.2.7 ReSIST Courseware vocabulary .................................................................................. 38
4.2.8 Open Vocabulary........................................................................................................................................... 38
4.2.9 LUDO-xAPI vocabulary ................................................................................................................................... 38
4.2.10 FOAF vocabulary............................................................................................................................................. 39
4.2.11 Dublin Core vocabulary................................................................................................................................. 39
4.3 PBL3.0 RDF VOCABULARY .................................................................................................................................. 40

5 SEMANTIC ANNOTATION ENVIRONMENT ........................................................................................................... 45
5.1 SEMANTIC ANNOTATION ENVIRONMENT ARCHITECTURE .................................................................................. 45
5.2 SEMANTIC ANNOTATION ENVIRONMENT OVERVIEW .......................................................................................... 46
5.2.1 Scenario 1: Semantic annotator - Annotation of a learning object ................................................................. 49
5.2.2 Scenario 2: Semantic browser - Discovery of existing learning objects ......................................................... 53
5.2.3 Scenario 3: Semantic enricher - Receiving enriched information per PBL step ............................................. 55

6 CONCLUSION .......................................................................................................................................................... 59

REFERENCES ............................................................................................................................................................ 61
List of Figures

FIGURE 1 SEMANTIC REPRESENTATION OF WEB INFORMATION (BARKER & CAMPBELL, 2015, PP. 11-13) ............ 16
FIGURE 2 SEMANTIC ANNOTATION OF LEARNING ARTIFACTS IN LOCO-ANALYST (DICHEVA ET AL., 2009) .......... 17
FIGURE 3 DBPEDIA SPOTLIGHT DEMO PAGE ...................................................................................... 18
FIGURE 4 SELECTION OF ONTOLOGIES FOR ANNOTATION WITH DBPEDIA SPOTLIGHT ......................... 19
FIGURE 5 DEMO USAGE OF PUNDIT ANNOTATOR .................................................................................. 20
FIGURE 6 DEMO USAGE OF AIDA ........................................................................................................ 21
FIGURE 7 DEMO USAGE OF APACHE STANBOL .................................................................................. 22
FIGURE 8 DEMO USAGE OF OPEN CALAIS .......................................................................................... 23
FIGURE 9 DEMO USAGE OF CWRC-Writer ......................................................................................... 24
FIGURE 10 WORKFLOW OF MARVIN ANNOTATOR (MILOSEVIC, 2016) .................................................. 25
FIGURE 11 DEMO USAGE OF FOX TOOL (SPECK & NGOMO, 2014, P. 2) .................................................. 26
FIGURE 12 DEMO USAGE OF RDFAZER ................................................................................................ 27
FIGURE 13 SALO CONCEPTUAL MODEL (PBL3.0 D2.1 “SEMANTIC MODEL DESIGN”) ......................... 32
FIGURE 14 LOV DATABASE .................................................................................................................. 33
FIGURE 15 LOM ONTOLOGY (KOUTSOMITRPOULOS ET AL., 2010) ......................................................... 34
FIGURE 16 IMS-LD SPECIFICATION .................................................................................................. 37
FIGURE 17 ReSIST COURSEWARE VOCABULARY VISUALIZATION WITH PROTEGÉ .................................. 38
FIGURE 18 DUBLIN CORE VOCABULARY ............................................................................................ 39
FIGURE 19 SEMANTIC ANNOTATION ENVIRONMENT ARCHITECTURE ................................................... 45
FIGURE 20 PBL3.0 SEMANTIC ANNOTATION TOOL INTERFACE ........................................................... 48
FIGURE 21 SEMANTIC ANNOTATION OF LEARNING OBJECT .................................................................. 50
FIGURE 22 SEMANTICALLY ANNOTATED LEARNING OBJECT .................................................................. 50
FIGURE 23 PART OF SEMANTICALLY ANNOTATED LEARNING OBJECT .................................................. 52
FIGURE 24 SEARCH EXISTING LEARNING OBJECTS ............................................................................. 53
FIGURE 25 LIST OF EXISTING LINKED LEARNING OBJECTS ................................................................ 54
FIGURE 26 BROWSING PUBLISHED LINKED DATA WITH PUBBY ............................................................. 54
FIGURE 27 SEARCHING COURSES ..................................................................................................... 55
FIGURE 28 BROWSING PUBLISHED LINKED COURSE .......................................................................... 55
FIGURE 29 ANNOTATION OF LEARNING OBJECT ................................................................................ 56
FIGURE 30 ENRICHED PUBLISHED LINKED DATA ............................................................................... 57
FIGURE 31 BROWSING AN ENRICHED ELEMENT OF PBL_LA FRAMEWORK (1) ............................... 57
FIGURE 32 BROWSING AN ENRICHED ELEMENT OF PBL_LA FRAMEWORK (2) .................................. 58
List of Tables

TABLE 1 SEMANTIC ANNOTATION TOOLS OVERVIEW ................................................................. 29
TABLE 2 SEMANTIC ANNOTATION TOOLS FUNCTIONALITIES MATRIX ........................................ 30
TABLE 3 PBL3.0 RDF VOCABULARY: CLASSES ........................................................................ 40
TABLE 4 PBL3.0 RDF VOCABULARY: PROPERTIES ...................................................................... 42
**List of Abbreviations**

The following table presents the acronyms used in the deliverable in alphabetical order.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIISO</td>
<td>Academic Institution Internal Structure Ontology</td>
</tr>
<tr>
<td>CMS</td>
<td>Content Management System</td>
</tr>
<tr>
<td>ESCO</td>
<td>European Skills/Competences, qualifications and Occupations</td>
</tr>
<tr>
<td>FOAF</td>
<td>Friend Of A Friend</td>
</tr>
<tr>
<td>FOX</td>
<td>Federated knOwledge eXtraction Framework</td>
</tr>
<tr>
<td>JSON</td>
<td>JavaScript Object Notation</td>
</tr>
<tr>
<td>LA</td>
<td>Learning Analytics</td>
</tr>
<tr>
<td>LD</td>
<td>Learning Design</td>
</tr>
<tr>
<td>LS</td>
<td>Learning Semantics</td>
</tr>
<tr>
<td>LOM</td>
<td>Learning Object Metadata</td>
</tr>
<tr>
<td>LOV</td>
<td>Linked Open Vocabularies</td>
</tr>
<tr>
<td>PBL</td>
<td>Problem-Based Learning</td>
</tr>
<tr>
<td>RDF</td>
<td>Resource Description Framework</td>
</tr>
<tr>
<td>SALO</td>
<td>Semantically Annotated Learning Object</td>
</tr>
<tr>
<td>SKOS</td>
<td>Simple Knowledge Organization System</td>
</tr>
<tr>
<td>SOLO</td>
<td>Structure of Observed Learning Outcomes</td>
</tr>
<tr>
<td>XML</td>
<td>eXtensible Markup Language</td>
</tr>
</tbody>
</table>
Executive Summary

The overall aim of the PBL3.0 project is to enhance Problem Based Learning (PBL) with Learning Analytics (LA) and Learning Semantics (LS). WP2 aims to provide a semantic model and a semantic annotation environment that will allow and guide educators and other stakeholders to semantically describe learning objects and resources in order to increase their discoverability, underpin their proper utilization in the learning process and assessment as well as enrich the information about them.

This deliverable is the last deliverable of WP2, and aims to present the semantic annotation environment that was developed taking into consideration the SALO model designed in D2.1. The tool can be used by all learning content providers such as MOOC providers, OER providers, educational institutions, national organizations etc., independently of domain and sector. This will foster mainstreamed content discoverability and linkage of learning resources with the Web of Data.

To this end, the report initially presents a review carried out on 10 existing open-source software tools that are available for annotating educational resources. The report then lists the different functionalities that are currently supported (e.g. manual-automatic annotation, plain unstructured text processing, limited RDF vocabularies, limited linking of data with other resources, no ability to import RDF vocabularies or export RDF data etc.) and assesses which of these are in accordance with the functionalities that should be supported by the semantic annotation environment specified in the project (namely, automatic annotation, structured input of data based on the SALO model, usage of educational RDF vocabularies, export of RDF data for further re-use, linking of data with other resources etc.).

Based on this requirements specification, we conclude that no existing tool can be used as-is for the project tool’s needs, and the configuration of any of the tools would require complex and time-consuming effort with no guarantee for success. This leads to the decision of developing a new environment with a user-friendly interface that allows the visualization of the SALO model’s concepts and performs most of the technical effort on the back-end, so that it can be used by any stakeholder independent of the domain of interest and technical background.

This deliverable also presents the review carried out on existing RDF vocabularies that focus on or include classes and properties of elements within the educational domain. This review studies 11 vocabularies that provide RDF terms for annotating academic institutions, courses, learning objects, learning resources, learning designs, people etc. We derive a sub-set of classes found in the studied vocabularies which correspond and can be used to describe all elements of the SALO model and learning resources. We also derive a sub-set of properties found in the vocabularies that can describe the relations formed amongst the SALO concepts so that they are understandable by humans and computer machines. These two lists comprise the PBL3.0 vocabulary that is used to annotate learning objects and resources with semantic meaning.
The semantic annotation environment that is developed comprises of four main toolkits, each supporting specific functionalities as follows:

- **Semantic annotator.** Description of learning objects and resources in accordance with the concepts of the SALO model and automatic semantic annotation of each concept based on the relevant PBL3.0 vocabulary classes and properties. Additionally, the annotator supports export of the semantically annotated data in Turtle format and automatic saving to an RDF store for further re-use.

- **Semantic publisher.** Automatic linkage of the annotated learning objects with relevant resources such as DBpedia, ESCO etc., and automatic publishing of the semantically annotated data as linked data for further linkage with external resources.

- **Semantic browser.** Search of existing learning objects and their visualization as published linked data.

- **Semantic enricher.** Automatic provision of published linked data from the PBL_LA framework designed in D1.3 (e.g. activities, LA methods, LA technologies, data generated) based on the PBL step entered for the learning object during annotation.

The effort required from the user is minimum, as the majority of the workload is carried out in the back-end so that the tool can be mainstreamed to any professional domain and sector and used by any stakeholder.

The tool will be tested for validation by stakeholders from both academia and business and from different domains. Any possible improvements and corrections required based on the testing will be carried out so that the final version of the tool by the end of the project will be optimized. Future works also includes the identification of ways the environment and its results can be incorporated within e-learning platforms such as LMSs, MOOCs etc.
1 Introduction

The aim of this section is to introduce the background of the work pursued with Task 2.3 “Semantic annotation tool”. The scope and the objective that the current document has set out to achieve are presented in sub-section 1.1. The intended audience for this document is described in sub-section 1.2 while sub-section 1.3 outlines the structure of the rest of the document.

1.1 Scope

The present document is the Deliverable 2.2 “Semantic annotation tool” (henceforth referred to as D2.2) of the PBL3.0 project. The main objective of D2.2 is to present all preparatory work carried out in order to construct an RDF vocabulary out of existing educational semantic vocabularies that are used for describing and annotating educational elements and to develop a semantic annotation tool that allows users to semantically describe learning objects of any educational or training domain.

1.2 Audience

The intended audience for this document is the PBL3.0 consortium, the European Commission, and the public interested in semantically annotating learning objects and discovering existing semantic learning objects to improve their learning and training settings.

1.3 Structure

The structure of the document is as follows:

- Section 2 describes the methodology followed in this deliverable.
- Section 3 presents the review carried out on existing semantic annotation tools and the requirements specification that was derived from the review.
- Section 4 presents the review carried out on existing RDF vocabularies that aim to describe educational concepts such as courses, institutions, learning objects, learning materials, learning designs etc., and describes the sub-set of RDF vocabularies, classes and properties that were chosen as most relevant based on the SALO model of D2.1 in order to be employed in the annotation of learning objects.
- Section 5 presents the semantic annotation environment that was developed, the architecture of the tool and the functionalities it supports.
- Section 6 concludes the document.
2 Methodology

The main goal of this deliverable is to present a tool that will be available to all stakeholders that want to utilize semantics and annotate resources in order to improve teaching and learning and that will be accessible for usage by any individual, for any field in the educational domain, with no or limited technical background requirements.

To this end, we initially performed a review of existing software tools available for semantic annotation and examined their functionalities. We focused on studying open-source and free software tools that could be also configured or extended, in case one or multiple tools were suitable to be used as the project’s semantic annotation environment.

Furthermore, we performed a review of the existing semantic vocabularies that are used to describe educational elements in order to examine the types of information that can be semantically depicted about a learning object. Our research focused on vocabularies that provide RDF elements (e.g. classes and properties) that are in line with the elements of the SALO model designed in D2.1. To this end, we searched the Linked Open Vocabularies (LOV)1 database by using SALO’s concepts as keywords, e.g. course, programme, activity, student, learning object etc. The aforementioned research resulted in the design of a combined PBL3.0 vocabulary that consists of the semantic classes and properties that can be used to describe a learning object in accordance to the SALO model and that will allow the object’s discoverability by other stakeholders.

Finally, we developed an open-source, web-based semantic annotation environment that allows any stakeholder to describe a learning object based on the concepts of the SALO model, with no need for technical or semantic-based knowledge. The tool was developed using the Eclipse IDE, MySQL database, the Java and AngularJS programming languages and the Bootstrap framework for the user interface. Furthermore, the tool utilizes the Fuseki and Virtuoso RDF stores and the SPARQL language for RDF annotation and the Pubby tool for Linked Data publishing. The annotation is done on the back-end, where we annotate each concept with the classes of the ontological model based on what each description is about, and we connect the concepts with RDF properties based on their relations.

1 https://lov.okfn.org/dataset/lov/
3 Semantic annotation

The rapidly growth of today’s markets and professional sectors require workforce with lifelong learning skills that can be constantly up-to-date with ongoing developments and able to be competitive in any professional field. This calls for a new paradigm in learning that will be mostly driven by business and societal demands and relevant to the context of each field (Mahmoud et al., 2013).

Furthermore, ongoing research in education and e-learning shows current trends and challenges in the field that should be addressed and solved (Howell et al., 2003; Merrill, 2003; Koper, 2003):

- Knowledge and data are growing rapidly, increasing the need for lifelong learning skills and e-learning technologies that can support distance learning.
- Learning is becoming more learner-centred and demands activities where learners solve real-world problems.
- Learning environments should support problem-centred learning and flexible, adaptive pathways based on learners’ progress.
- Effort of educators should be focused on supporting and scaffolding learners’ active engagement in learning, and thus automatic functions that can alleviate effort of teaching should be encouraged.

The increasing usage of the semantic web in education could support the above issues, as semantics allow more automatic representation and discovery of information. The inventor of the World Wide Web also states that the semantic web can support the communication and shared understanding between humans and computers on a semantic basis (Berners – Lee, 2001). This understanding is underpinned through the usage of domain ontologies and formal specifications / vocabularies, where learning objects are annotated with pieces of information and organized into blocks of descriptive statements (Stojanovic et al., 2001).

Figure 1 also shows how the semantic representation of knowledge can help computers organize the information in a structured and meaningful manner.
This semantic structure can improve learning implementation and assessment by allowing educators to perform their tasks more effectively through automated recommendations based on specific problem-based tasks or learners’ profiles. This can be supported when courses and their relevant concepts are semantically annotated, where software agents are aware of these concepts’ meanings and can thus support filtering of appropriate resources, managing and recommending workflow of activities based on learners’ profiles etc. (Bergin et al., 2000).

The semantic representation of courses also supports the exchange and re-use of learning objects across different educational fields and sectors and thus contribute to the courses’ improvement and efficiency (Littlejohn, 2003), as they are more easily discoverable when the computers are aware of what types of elements they are and what information they hold. Furthermore, educators can enrich their knowledge on their educational field when their courses and learning objects are linked with external semantic information (e.g. DBpedia concepts) and are therefore constantly up-to-date with the ongoing growth of information.

### 3.1 Review of semantic annotation tools

This section presents the research carried out on existing technologies that allow annotation of concepts with semantic information from domain ontologies. These tools are available for usage, either in open-source format for download and installation, or through a provided API. Additionally, the tools reviewed are not dependent on a specific domain but can be used to annotate different types of concepts.
3.1.1 LOCO-Analyst

LOCO-Analyst is an open-source tool that allows semantic annotation in order to connect different learning artefacts (Jovanović, 2008). It is a desktop application available for download\(^2\), and a part of the Reload Content Packaging Editor, an open-source tool for creating courses compliant with the IMS Content specification. The semantic annotation functionality is carried out via the Knowledge and Information Management (KIM) platform, which is, however, neither open-source nor free for commercial usage.

During the annotation process, each learning artefact, e.g. lesson, quiz question, chat message, forum post etc., is assigned none or multiple annotations from domain ontologies (Dicheva et al., 2009).

As shown in the example in Figure 2, let us assume that we have a learning artefact e.g. a specific pseudocode, which we want to annotate and give meaning to it so that it is understandable by computers as well as humans. The tool allows us to assign meaning to the pseudocode with a specific identifier (e.g. http://www.iHelpCourses.org/CMPT100/ProgProcess#Intro) and specify that it belongs to a domain topic (by using the RDF property lc:hasDomainTopic) with the identifier kim-wkb:pseudocode and that it has a humanly recognizable label “pseudocode”. The elements used to describe the pseudocode are instances of domain ontologies’ classes and properties such as alocom-core, lc, proton and rdfs.

3.1.2 DBpedia Spotlight

DBpedia Spotlight is an open-source web-based open-source application that allows automatic annotations of content based on resources stored in DBpedia (Daiber et al., 2013).

\(^2\) http://semwebcentral.org/scm/viewvc.php/?root=locoanalyst
The tool allows users to copy paste a text in the form and prepare the annotation process by setting different parameters. These parameters include:

- **Language of text.** DBpedia stores content in different languages; thus, the tool allows annotation of content in 10 different languages, including English, French, German, Italian, Dutch etc.

- **Confidence percentage.** Users can determine how confident the annotation process should be, e.g. on what level to annotate concepts that are very or less similar to content stored in DBpedia.

- **Selection of specific ontologies and types.** Users can select which classes to use in the annotation for a more domain-specific annotation. The tool allows selection of classes from the DBpedia ontology, the Freebase ontology, the Schema ontology as well as based on a custom SPARQL query, as shown in Figure 4.

---

The tool has a user-friendly interface and is web-based; however, its usage has specific limitations, since it only allows semantic annotation of text and not of structured information like we aim for our tool. Additionally, it requires that the users have knowledge on semantics and ontological structures when choosing classes and ontologies.

The tool is available for download\(^4\) and also provides a demo\(^5\) for experimentation.

### 3.1.3 Pundit Annotator

Pundit is an open-source web application that allows users to annotate web pages using semantic elements (Grassi et al., 2013; Grassi et al., 2012). The tool can be embedded as a Chrome plugin for annotating web pages that users visit or it can be embedded into a custom project for annotating locally saved web pages.

Users can highlight specific text within a web page and annotate it with an RDF triple element (i.e. subject, property, object) as well as with entities from existing ontologies, as shown in Figure 5. This way, RDF triples are created which describe the subject concept using elements from the RDF/RDFS specifications, the Dublin Core ontology and a custom Pundit ontology. Pundit also allows the connection of the concepts with linked data, focusing on the DBpedia database.

---

\(^4\) [https://github.com/dbpedia-spotlight/dbpedia-spotlight](https://github.com/dbpedia-spotlight/dbpedia-spotlight)

The Chrome plugin is easy to use for non-technical experts; however, it focuses on the annotation of web pages and allows manual annotation, which would require significant effort from educators.

The tool is available for download[^6] and also provides a demo for experimentation[^7].

### 3.1.4 AIDA

AIDA is an open-source tool developed by the Databases and Information Systems Department at the Max Planck Institute for Informatics in Saarbrücken, Germany, and it identifies mentions of named entities (e.g., persons, organizations, locations, songs, products, etc.) in English language text and links them to a unique identifier (Gangemi, 2013).

The tool also maps the identified concepts with entities that are registered in the Wikipedia-driven YAGO2 knowledge base and thus provides additional information, enriching the concepts.

Figure 6 shows AIDA’s interface, where users can insert a text in the top right part of the screen and decide on different parameters for the disambiguation process, e.g., robustness levels, ambiguity degrees, etc.

[^6]: [https://github.com/net7/pundit2](https://github.com/net7/pundit2)
[^7]: [http://thepund.it/annotator-pro-web-annotation/](http://thepund.it/annotator-pro-web-annotation/)
As shown in the example, AIDA identifies the different locations and people present in the text, and attempts to assign meaning to each concept by matching it with similar entities found in the YAGO2 knowledge base.

The tool is available for download\(^9\) and also provides a demo\(^{10}\) for experimentation.

### 3.1.5 Apache Stanbol

Apache Stanbol is an open-source tool, part of the Apache Foundation Software, which aims to extend existing Content Management Systems (CMS) with semantic functionalities. These functionalities aim to semantically describe data and release the results as RDF and JSON, as shown in Figure 7.

---

\(^8\) [https://gate.d5.mpi-inf.mpg.de/webaida/](https://gate.d5.mpi-inf.mpg.de/webaida/)

\(^9\) [https://github.com/yago-naga/aida](https://github.com/yago-naga/aida)

\(^{10}\) [https://gate.d5.mpi-inf.mpg.de/webaida/](https://gate.d5.mpi-inf.mpg.de/webaida/)
The data for annotation is initially converted into plain text for analysis; then, the tool detects the language of the text and proceeds to extract different entities (e.g., persons, organizations, locations etc.). Finally, Apache Stanbol allows the linkage of the identified entities with DBpedia elements.

The tool is available for download\footnote{https://stanbol.apache.org/overview.html} and can be installed as a standalone desktop or web application.

3.1.6 Open Calais

Open Calais Thomson Reuters is a web service developed to allow tagging of people, places, companies, facts, and events in a text. The tool utilizes machine learning and statistical methods in order to assign meaning to concepts of the text and also maps the tags to Thomson Reuters unique identifiers that will allow linkage of the data with related information from Thomson Reuters datasets.

\footnote{https://svn.apache.org/repos/asf/stanbol/trunk/}
Figure 8 shows an example of using Open Calais to tag a text by identifying entities and assigning each entity into a category based on the data processing carried out.

The tool is available for use through the provided API.

### 3.1.7 CWRC-Writer

CWRC-Writer is an open-source in-browser text markup editor developed by the Canadian Writing Research Collaboratory (CWRC). The tool allows standard XML WYSIWYG editing, where the text of the editors can be annotated with specific entities (e.g. people, locations, organizations), as shown in Figure 9.

---

D2.2 Semantic Annotation Tool

The tool supports annotation by using multiple supporting services, such as:

- a document store that processes the documents and texts inserted in the tool
- an XML schema or multiple schemas that provide structure for the annotation process
- an XML validation service that validates the semantically annotated data
- an annotation store that manages the annotations produced and
- a service that searches existing URIs of linked data entities or creates new URIs.

The tool is available for download\(^\text{16}\) and a demo version\(^\text{17}\) is available for experimentation.

### 3.1.8 Marvin text annotator

Marvin is an open-source semantic annotation tool that uses various sources to annotate text (Milosevic, 2016). The sources used include the Wordnet library for recognizing words in English and their synonyms, the DBpedia database for enriching data with similar information and the MetaMap program for recognizing biomedical concepts in the text, as shown in Figure 10.

---

\(^{15}\) [https://github.com/cwrc/CWRC-Writer-Reference-Implementation](https://github.com/cwrc/CWRC-Writer-Reference-Implementation)

\(^{16}\) [https://github.com/cwrc/CWRC-Writer](https://github.com/cwrc/CWRC-Writer)

\(^{17}\) [https://sites.google.com/site/cwrcwriterhelp/](https://sites.google.com/site/cwrcwriterhelp/)
The workflow presented shows that the inputted document is initially tokenized using the Apache OpenNLP machine learning toolkit (Baldridge, 2005) in order to be annotated with concepts from the external sources. The annotation is extracted in the SKOS (Simple Knowledge Organization System) format. The tool is available for download\(^\text{18}\) and can be used as a standalone program or as Java library for an external project.

### 3.1.9 Federated knOwledge eXtraction Framework (FOX)

FOX (Federated knOwledge eXtraction Framework) is an open-source framework and toolkit developed to receive text input and provide semantic output in RDF Turtle format (Speck & Ngomo, 2014). The FOX interface allows the input of data which users can indicate if it concerns text or a URL. Additionally, users indicate the output format they prefer (e.g. RDF, JSON-LD, Turtle etc.) and the type of entity recognition method to be used.

\(^{18}\) https://github.com/nikolamilosevic86/Marvin
Furthermore, FOX allows linkage of the identified entities against Linked Data databases such as DBpedia, using the AGDISTIS Named Entity Disambiguation framework and focusing on identifying entities of the “Person”, “Organization” and “Location” types (Ngomo et al., 2011).

The tool is available for download\textsuperscript{19} and a demo\textsuperscript{20} version is available for experimentation.

\textsuperscript{19} https://github.com/AKSW/fox
\textsuperscript{20} http://fox-demo.aksw.org/index.html#!\#demo
3.1.10 RDFazer

RDFazer is a plugin for the Chrome browser and it allows the annotation of websites with RDFa concepts. The tool allows users to manually tag text within a static HTML page with concepts from an existing ontology and outputs the results in RDFa format, as shown in Figure 12.

![Demo usage of RDFazer](https://github.com/Rahien/RDFazer)

RDFazer also supports annotation of text using formal specifications (e.g. EU’s ESCO ontology) through the SPARQL endpoint search functionality (Zotou et al., 2014), which requires knowledge on SPARQL queries.

The tool is available for download and can be installed in the Chrome browser for usage.

### 3.2 Semantic annotation environment requirements specification

This section provides an overview of the project’s objectives on the functionalities to be supported by the semantic annotation environment, and the functionalities currently supported by the existing software tools. The comparison of these two will facilitate us to determine whether one or multiple tools and/or functionalities can be utilized for the PBL3.0 semantic annotation environment.

---

21 [https://github.com/Rahien/RDFazer](https://github.com/Rahien/RDFazer)

22 [https://github.com/Rahien/RDFazer](https://github.com/Rahien/RDFazer)
The functionalities that should be supported by the project’s environment based on current challenges and needs include:

- Automatic annotation of structured data. Support automatic annotation of the learning resources, as the main end-users will be educators of any domain. Thus, the required technical skills for using the tool should be limited, in order to facilitate mainstreaming of the tool’s usage. Additionally, educators and any other stakeholders should be able to enter information about learning objects, which will represent instances of the SALO’s model concepts. Thus, it is important that the developed tool’s user interface allows input of data in a structured way and not solely plain text.

- Annotation using educational vocabularies. Map the SALO’s concepts to specific RDF educational vocabularies that combine multiples specifications and are presented in Section 4.

- Creation of linked published data. Support the automatic linkage of the data with external sources for enrichment of knowledge and publishing of the data for easier browsing and future connection with other concepts within the Web of data.

- Support the export of the RDF data, as it is essential to allow reuse and support the discoverability of the semantic learning information.

The study of the functionalities supported by the existing tools generated a set of criteria for assessment, as follows:

- Type of annotation. It is important to note if the annotation allowed is manual or automatic. The review shows that 6 of the tools support automatic annotation, while the rest allow users to manually tag the resources.

- Type of content. We document what kind of content each tool supports for annotation, e.g. text, documents, web pages etc. Based on the overview, the vast majority of the existing tools support annotation of plain, unstructured text.

-Usage of RDF vocabularies in annotation. We document whether each tool supports the annotation of data based on RDF vocabularies, such as FOAF, Dublin Core or custom-developed ontologies. Based on the overview, the majority of the tools support annotation with vocabularies such as FOAF, Dublin Core and DBpedia. None of the tools support customized vocabulary import.

- Creation of Linked data. We document whether each tool supports the linkage of the annotated concepts to external sources of information, such as DBpedia or other RDF repositories. Based on the overview, it seems that 7 of the 10 tools support linkage of data, mostly through the DBpedia database.
- Export of RDF for reuse. We document whether each tool allows the reuse of the annotated data. Based on the overview, it seems that the majority of the tools do not support export of the semantic data for future exploitation.

Table 1 shows the values of each criterion per tool.

<table>
<thead>
<tr>
<th>Tool</th>
<th>Type of annotation</th>
<th>Type of content</th>
<th>Usage of RDF vocabularies in annotation</th>
<th>Creation of Linked data</th>
<th>Export of RDF for reuse</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOCO-Analyst</td>
<td>Manual</td>
<td>Unstructured plain text</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>DBpedia Spotlight</td>
<td>Automatic</td>
<td>Unstructured plain text</td>
<td>Yes (DBpedia)</td>
<td>Yes (DBpedia)</td>
<td>No</td>
</tr>
<tr>
<td>Pundit</td>
<td>Manual</td>
<td>Web pages</td>
<td>Yes (Dublin Core, CiTO etc.)</td>
<td>Yes (DBpedia)</td>
<td>No</td>
</tr>
<tr>
<td>AIDA</td>
<td>Automatic</td>
<td>Unstructured plain text</td>
<td>No</td>
<td>Yes (YAGO2)</td>
<td>No</td>
</tr>
<tr>
<td>Apache Stanbol</td>
<td>Automatic</td>
<td>Unstructured plain text</td>
<td>Yes (DBpedia, Dublin Core, FOAF etc.)</td>
<td>Yes (DBpedia)</td>
<td>Yes</td>
</tr>
<tr>
<td>Open Calais</td>
<td>Automatic</td>
<td>Unstructured plain text</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>CWRC-Writer</td>
<td>Manual</td>
<td>Unstructured plain text</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>RDFazer</td>
<td>Manual</td>
<td>HTML pages</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Marvin</td>
<td>Automatic</td>
<td>Unstructured plain text</td>
<td>No</td>
<td>Yes (DBpedia)</td>
<td>Yes</td>
</tr>
<tr>
<td>FOX</td>
<td>Automatic</td>
<td>Unstructured plain text</td>
<td>Yes</td>
<td>Yes (limited DBpedia)</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Based on the above table, the review results show:

- Type of annotation. The review shows that 6 of the tools support automatic annotation, while the rest allow users to manually tag the resources.
Type of content. Based on the overview, the vast majority of the existing tools support annotation of plain, unstructured text.

Usage of RDF vocabularies in annotation. Based on the overview, the majority of the tools support annotation with vocabularies such as FOAF, Dublin Core and DBpedia. None of the tools support customized vocabulary import.

Creation of Linked data. Based on the overview, it seems that 7 of the 10 tools support linkage of data, mostly through the DBpedia database.

Export of RDF for reuse. Based on the overview, it seems that the majority of the tools do not support export of the semantic data for future exploitation.

We proceed to assess if one or multiple tools could be used as-is or configured in order to support all functionalities aforementioned, as shown in Table 2.

<table>
<thead>
<tr>
<th>Tool</th>
<th>Automatic annotation</th>
<th>Structured data input</th>
<th>Usage of educational vocabularies</th>
<th>Creation of Linked published data</th>
<th>Export linked data</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOCO-Analyst</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>DBpedia</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Spotlight</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pundit</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIDA</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apache</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Stanbol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open Calais</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CWRC-Writer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>RDFazer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Marvin</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>FOX</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Based on the above, we cannot identify an existing tool that can cover all target functionalities and features for the PBL3.0 semantic annotation environment. Moreover, we note that the configuration of any of the aforementioned technologies in order to support structured data input, automatic annotation with educational technologies and export of the data as RDF output would require many
changes, as all of them focus on text processing. We do not proceed to choose an existing tool for our project’s needs due to the in-depth study of each tool’s source code, the comprehension of the structure of the workflow and the programming of large number of changes.

Thus, we decided to develop a new tool that will support the majority of features provided by the existing semantic annotation tools, but with enriched functionalities. We also aim to connect the learning resources described with the tool with the concepts of the PBL_LA educational framework developed in D1.3. Elaboration on the tool and its features will be provided in Section 5.
4 Semantic vocabularies for semantic annotation in education

This section provides an overview of the RDF vocabularies that have been studied in order to decide which classes and properties will be used to describe the SALO model’s concepts (D2.1 “Semantic model design”).

4.1 SALO model

The SALO model proposes that its concepts should describe the following categories:

- Programs and course structure. This category aims to provide information of a learning object regarding the related cohort (e.g. academic year), programme, semester, course, project work etc. This information adds context to the learning object and should therefore be annotated with corresponding semantic concepts.

- Activity design. This category aims to provide terms that describe the design of activities for the learning object.

- Group works and group assessments. This category aims to provide information on how groups work is evaluated when students interact with the learning object.

- Learning objectives. This category aims to provide information on the different skills and knowledge that should be gained through interaction with the learning object.

- Activity structure. This category aims to provide information on how activities are structured, what are their types and what they accomplish.

The following Figure shows the conceptual SALO model as designed in D2.1.

![SALO conceptual model](image)

Figure 13 SALO conceptual model (PBL3.0 D2.1 “Semantic model design”)
The semantic annotation environment endeavours to allow stakeholders to describe their learning objects by providing information on all SALO concepts and thus expose this information to the Web of data for future discoverability and improving teaching and learning.

4.2 Review of RDF vocabularies in education

The annotation of learning objects and resources requires the identification of classes and objects that allow their appropriate semantic description. To this end, we carried out a research on which RDF vocabularies are available that focus on semantically describing educational elements. In our research, we focused on searching the LOV database, which contains 587 vocabularies that can be used for semantically describing resources from multiple domains, including education.

![LOV database](https://lov.okfn.org/dataset/lov/)

In total, we found eleven vocabularies that are well-defined and include classes and properties for concepts that can be used to describe educational elements. We proceed to decide which of these classes and properties better describe the concepts of the SALO model and we also take into consideration the concepts included in the PBL_LA educational framework of D1.3.

4.2.1 LOM vocabulary

The LOM (Learning Object Metadata) ontology exposes the IEEE LOM standard used for educational content as linked data (Li, 2004; Koutsomitropoulos et al., 2010). Thus, the ontology provides a

---

23 [https://lov.okfn.org/dataset/lov/](https://lov.okfn.org/dataset/lov/)
vocabulary with classes and properties that will provide metadata information about any learning object. The model and concepts supported by the vocabulary are shown in Figure 15.

In summary, some of the classes supported include:

- **Learning object.** The specific learning object that we are describing.
- **Identifier.** A unique identifier assigned to the learning object described.
- **Annotation:** Comments on the educational use of a learning object, including information on when and by whom the comments were created.
- **Contribution:** The entities (i.e., people, organizations) that have contributed to the state of this learning object during its life cycle (e.g., creation, edits, publication).
- **Relation:** A concept which defines the relationship between a learning object and other learning objects, if any.

Additionally, some of the properties included in the vocabulary include:

- **hasResourceIdentifier:** a globally unique label that identifies the learning object.
- **description:** a text description of the learning object.

---

24 [http://slor.sourceforge.net/ontology/lom.owl](http://slor.sourceforge.net/ontology/lom.owl)
• educational: the key pedagogical characteristics of the learning object.
• classificationTaxon: a term within a taxonomy that aims to classify a learning object.
• duration: duration of the learning object.

4.2.2 LRMI vocabulary

The Learning Resource Metadata Initiative (LRMI) has developed a specification for semantically describing learning resources, and also provides an RDF version of its terms. The specification aims to facilitate educators and other stakeholders in more easily searching, identifying and sharing educational content across different store engines (Barker & Campbell, 2015). The LRMI vocabulary builds on the extensive vocabulary provided by Schema.org and other standards. The classes it supports are:

• AlignmentObject: An intangible item that describes an alignment between a learning resource and a node in an educational framework, e.g. specific lesson in a curriculum.
• EducationalAudience: An individual or group for whom the learning resource was created or useful

The vocabulary also includes several properties, such as:

• alignmentType: a category of alignment between the learning resource and the node of an educational framework. Recommended values for the property include: “assesses”, “teaches”, “requires”, “educationalSubject”, and “educationLevel”.
• educationalRole: the role that describes the target audience of the learning resource.
• educationalUse: the purpose of a learning resource in the context of a course. Recommended values for the property include “group work”, “assignment” etc.
• interactivityType: the prominent mode of learning supported by the learning resource. Recommended values include “active”, “mixed”, “expositive”.
• learningResourceType: the prominent type characterizing the learning resource. Recommended values for the property include “presentation”, “exercise”, “handout” etc.

4.2.3 AIISO vocabulary

The AIISO (Academic Institution Internal Structure Ontology) ontology provides a vocabulary with classes and properties that describe the organizational structure of an academic institution (Styles &
Shabir, 2008; Mouromvtsev et al., 2013). This includes information about classes, departments, subjects etc. For example, some of the classes available are as follows\textsuperscript{27}:

- Center: a group of people recognised by an organization as a cohesive group
- Course: a cohesive collection of educational material
- KnowledgeGrouping: a collection of resources, learning objectives, timetables, and other materials
- Programme: a KnowledgeGrouping that represents a cohesive collection of educational material referred to by an organization as a programme
- Subject: a KnowledgeGrouping that represents a cohesive collection of educational material referred to by the organization as a subject

Additionally, AIISO provides a list of properties that can be used to further describe educational elements, such as:

- name: the name of the described element.
- responsibilityOf: when a concept is the responsibility of another concept.
- responsibleFor: when a concept is responsible for another concept.

4.2.4 ESCO vocabulary

The ESCO (European Skills/Competences, qualifications and Occupations) ontology provides a vocabulary with classes and properties for describing skills, occupations, qualifications and organisations (Zotou et al., 2014). Examples of such classes include\textsuperscript{28}:

- Skill: when a concept represents a skill.
- Group: when a concept represents a group of members.
- Occupation: when a concept represents an occupation.
- Qualification: when a concept represents a qualification.

Additionally, some of the properties available in ESCO are:

- name: the name of the described concept.
- isOptionalSkillFor: when a skill is optional for another skill or for an occupation
- subject: the topic of the concept.

\textsuperscript{27} http://vocab.org/aiiso/
\textsuperscript{28} https://ec.europa.eu/esco/resources/data/static/model/html/model.xhtml
4.2.5 IMS – LD vocabulary

The IMS-LD conceptual model is very commonly used to describe learning elements in order to guide educators when designing their courses and enable learning resources’ interoperability (Derntl et al., 2012). To this end, an RDF version of the ontology has been created which transforms all its concepts into RDF classes and properties for easier discovery and linkage with the Web of data. Figure 16 shows a segment of the IMS-LD vocabulary that is available for re-use.

![IMS-LD specification](http://jelenajovanovic.net/ontologies/loco/LearningDesignOnt.rdf)

Some of the classes available in the vocabulary include ActivityStructure, Activity, Environment, Role, LearningObjective etc., and some of the properties include terms such as hasPrerequisite, hasLearningObjective, hasLearner etc.

4.2.6 Teaching Core vocabulary

The Teaching Core Vocabulary provides classes and properties for teachers to use when describing their courses. Some of the classes available are Course, Assignment, Material, Student, Teacher, Lecture etc.

Additionally, properties of the vocabulary include terms such as courseTitle, teacherOf, hasAssignment etc.

---

29 [http://jelenajovanovic.net/ontologies/loco/LearningDesignOnt.rdf](http://jelenajovanovic.net/ontologies/loco/LearningDesignOnt.rdf)

4.2.7 ReSIST Courseware vocabulary

The ReSIST Courseware vocabulary provides classes and properties for describing courses and resources. As shown in Figure 17, the ReSIST Courseware classes can be used to describe concepts such as assessment methods, Courses, student interactivities etc, while some of the properties are: has-title, has-courseware, has-assessment-method etc.

![Figure 17 ReSIST Courseware vocabulary visualization with Protégé](image)

4.2.8 Open Vocabulary

This vocabulary includes terms in RDF format for multiple domains, including education. Some of the classes included are Course, CourseTeacher, Database etc., and some of the relevant properties are semester, teaches, deliveredby etc.

4.2.9 LUDO-xAPI vocabulary

The LUDO-xAPI vocabulary is an OWL representation of the statements defined in the Experience API (xAPI) specification. This vocabulary is focused on describing interactions within Serious Games,

31 [http://lov.okfn.org/dataset/lov/vocabs/crsw](http://lov.okfn.org/dataset/lov/vocabs/crsw)
aiming to provide personalized recommendations on helpful learning resources that could improve students’ learning progress. Even though the focus is Serious Games, there are multiple classes and properties that focus on generic activities based on the xAPI concepts and thus can be adopted for other learning processes (e.g. Activity, InteractionActivity, Group, hasInteractionType etc.).

4.2.10 FOAF vocabulary

The FOAF (Friend Of A Friend) vocabulary is one of the most commonly used when describing people and organizations. The classes and properties of FOAF can be used for annotating specific people, an organization (e.g. a University or a training company), a project, a tool, a document etc.

4.2.11 Dublin Core vocabulary

The Dublin Core vocabulary includes terms to semantically describe web resources such as web pages, videos, images etc., as well as physical resources such as books, documents, as shown in Figure 18.

![Diagram of Dublin Core vocabulary](http://ns.inria.fr/ludo/v1/docs/xapi.html)

![Diagram of Dublin Core vocabulary](http://dublincore.org/documents/dcmi-terms/)

---

33 [http://ns.inria.fr/ludo/v1/docs/xapi.html](http://ns.inria.fr/ludo/v1/docs/xapi.html)

4.3 PBL3.0 RDF vocabulary

All aforementioned RDF vocabularies were studied and evaluated in regards to their relevance to the SALO concepts that will be included in the semantic annotation environment. In the RDF framework, every statement about a resource is represented as a triple with the structure: Subject – Property– Object. Table 3 shows the vocabularies’ classes that are used within the tool for describing the “Subject” or the “Object” segment of each resource description, where we annotate the type of the resource.

<table>
<thead>
<tr>
<th>SALO term</th>
<th>Vocabulary</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Object</td>
<td>LOM</td>
<td><a href="http://data.opendiscoveryspace.eu/lom_ontology_ods.owl#Learning">http://data.opendiscoveryspace.eu/lom_ontology_ods.owl#Learning</a> Object</td>
</tr>
<tr>
<td>COHORT</td>
<td>AIISO</td>
<td><a href="http://purl.org/vocab/aiiso/schema#KnowledgeGrouping">http://purl.org/vocab/aiiso/schema#KnowledgeGrouping</a></td>
</tr>
<tr>
<td>PROGRAM</td>
<td>AIISO</td>
<td><a href="http://purl.org/vocab/aiiso/schema#Programme">http://purl.org/vocab/aiiso/schema#Programme</a></td>
</tr>
<tr>
<td></td>
<td>Teaching Core</td>
<td><a href="http://linkedscience.org/teach/ns#StudyProgram">http://linkedscience.org/teach/ns#StudyProgram</a></td>
</tr>
<tr>
<td>SEMESTER</td>
<td>AIISO</td>
<td><a href="http://purl.org/vocab/aiiso/schema#KnowledgeGrouping">http://purl.org/vocab/aiiso/schema#KnowledgeGrouping</a></td>
</tr>
<tr>
<td></td>
<td>Dublin Core</td>
<td><a href="http://purl.org/dc/terms/extent">http://purl.org/dc/terms/extent</a></td>
</tr>
<tr>
<td>COURSE</td>
<td>Teaching Core</td>
<td><a href="http://linkedscience.org/teach/ns#Course">http://linkedscience.org/teach/ns#Course</a></td>
</tr>
<tr>
<td></td>
<td>AIISO</td>
<td><a href="http://purl.org/vocab/aiiso/schema#Course">http://purl.org/vocab/aiiso/schema#Course</a></td>
</tr>
<tr>
<td></td>
<td>ReSIST</td>
<td><a href="http://courseware.rkbexplorer.com/ontologies/courseware#Course">http://courseware.rkbexplorer.com/ontologies/courseware#Course</a></td>
</tr>
<tr>
<td>COURSE WORK</td>
<td>AIISO</td>
<td><a href="http://purl.org/vocab/aiiso/schema#Module">http://purl.org/vocab/aiiso/schema#Module</a></td>
</tr>
<tr>
<td>PROJECT</td>
<td>IMS-LD</td>
<td><a href="http://www.lornet.org/LOCO#Method">http://www.lornet.org/LOCO#Method</a></td>
</tr>
<tr>
<td></td>
<td>FOAF</td>
<td><a href="http://xmlns.com/foaf/spec/Project">http://xmlns.com/foaf/spec/Project</a></td>
</tr>
<tr>
<td>PROJECT WORK</td>
<td>IMS-LD</td>
<td><a href="http://www.lornet.org/LOCO#Play">http://www.lornet.org/LOCO#Play</a></td>
</tr>
<tr>
<td>SKILL</td>
<td>ESCO</td>
<td><a href="http://data.europa.eu/esco/model#Skill">http://data.europa.eu/esco/model#Skill</a></td>
</tr>
</tbody>
</table>

| SOLO TAXONOMY | LOM | http://data.opendiscoveryspace.eu/lom_ontology_ods.owl#Taxon |
| PHASE | LOM | http://data.opendiscoveryspace.eu/lom_ontology_ods.owl#Classification |
| ACTIVITY STRUCTURE TYPE | IMS-LD | http://www.lornet.org/LOCO#ActivityStructure |
| | Dublin Core | http://purl.org/dc/terms/MethodOfInstruction |
| ACTIVITY TYPE | IMS-LD | http://www.lornet.org/LOCO#LearningActivity |
| | LOM | http://slor.sourceforge.net/ontology/lom.owl#InteractivityTypeVocabulary |
| LEARNING ACTIVITY | IMS-LD | http://www.lornet.org/LOCO#Activity |
| | LUDO-xAPI | http://ns.inria.fr/ludo/v1/docs/xapi.html#Activity |
| STUDENT | IMS-LD | http://www.lornet.org/LOCO#Learner |
| | Teaching Core | http://linkedscience.org/teach/ns/#Student |
| ROLE | IMS-LD | http://www.lornet.org/LOCO#Role |
| STUDENT GROUP | Teaching Core | http://linkedscience.org/teach/ns/#StudentGroup |
| | LRMI | http://purl.org/dcx/lrmi-terms/EducationalAudience |
| PROJECT SUPERVISOR | IMS-LD | http://www.lornet.org/LOCO#Staff |
| | Teaching Core | http://linkedscience.org/teach/ns/#Teacher |
| CENSOR | IMS-LD | http://www.lornet.org/LOCO#Staff |
| | Teaching Core | http://linkedscience.org/teach/ns/#Teacher |
| TOOL | IMS-LD | http://www.lornet.org/LOCO#Tool |
| | Dublin Core | http://purl.org/dc/dcmitype/Software |
| LEARNING CONTENT | Teaching Core | http://linkedscience.org/teach/ns/#Material |
| | IMS-LD | http://www.lornet.org/LOCO#Resource |
| COMPETENCY | IMS-LD | http://www.lornet.org/LOCO#LearningObjective |
Table 4 shows the properties that can be used to describe the relations between the annotated concepts within SALO’s PBL3.0 RDF Vocabulary.

<table>
<thead>
<tr>
<th>Property / Vocabulary</th>
<th>Subject</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://purl.org/vocab/aiiso/schema#name">http://purl.org/vocab/aiiso/schema#name</a></td>
<td>COHORT</td>
<td>String value</td>
</tr>
<tr>
<td></td>
<td>PROGRAMME</td>
<td></td>
</tr>
<tr>
<td></td>
<td>COURSEWORK</td>
<td></td>
</tr>
<tr>
<td><a href="http://data.opendiscoveryspace.eu/lom_ontology_ods.owl#educationalInteractivityType">http://data.opendiscoveryspace.eu/lom_ontology_ods.owl#educationalInteractivityType</a></td>
<td>LEARNING</td>
<td>ACTIVITY</td>
</tr>
<tr>
<td></td>
<td>OBJECT</td>
<td>STRUCTURE TYPE</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ACTIVITY</td>
<td>STRUCTURE TYPE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>String value</td>
</tr>
<tr>
<td><a href="http://linkedscience.org/teach/ns#courseTitle">http://linkedscience.org/teach/ns#courseTitle</a></td>
<td>COURSE</td>
<td>String value</td>
</tr>
<tr>
<td><a href="http://courseware.rkbexplorer.com/ontologies/courseware#has-title">http://courseware.rkbexplorer.com/ontologies/courseware#has-title</a></td>
<td>COURSE</td>
<td>String value</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="http://linkedscience.org/teach/ns#academicTerm">http://linkedscience.org/teach/ns#academicTerm</a></td>
<td>COURSE</td>
<td>SEMESTER</td>
</tr>
<tr>
<td><a href="http://data.opendiscoveryspace.eu/lom_ontology_ods.owl#taxonEntry">http://data.opendiscoveryspace.eu/lom_ontology_ods.owl#taxonEntry</a></td>
<td>SOLO</td>
<td>String value</td>
</tr>
<tr>
<td><a href="http://courseware.rkbexplorer.com/ontologies/courseware#course-objectives">http://courseware.rkbexplorer.com/ontologies/courseware#course-objectives</a></td>
<td>COURSE</td>
<td>SKILL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="http://purl.org/dc/terms/instructionalMethod">http://purl.org/dc/terms/instructionalMethod</a></td>
<td>ACTIVITY</td>
<td>String value</td>
</tr>
</tbody>
</table>

Table 4 PBL3.0 RDF vocabulary: Properties
<table>
<thead>
<tr>
<th>URI</th>
<th>STRUCTURE TYPE</th>
<th>VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://purl.org/dc/elements/1.1/title">http://purl.org/dc/elements/1.1/title</a></td>
<td>TOOL</td>
<td>String value</td>
</tr>
<tr>
<td><a href="http://purl.org/vocab/aiiso/schema#part_of">http://purl.org/vocab/aiiso/schema#part_of</a></td>
<td>LEARNING OBJECT</td>
<td>COURSE PROGRAMME</td>
</tr>
<tr>
<td><a href="http://www.lornet.org/LOCO#isOfConferenceType">http://www.lornet.org/LOCO#isOfConferenceType</a></td>
<td>MEETING</td>
<td>String value</td>
</tr>
<tr>
<td><a href="http://purl.org/dc/terms/description">http://purl.org/dc/terms/description</a></td>
<td>MEETING</td>
<td>String value</td>
</tr>
<tr>
<td><a href="http://data.opendiscoveryspace.eu/lom_ontology_ods.owl#classification">http://data.opendiscoveryspace.eu/lom_ontology_ods.owl#classification</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="http://www.lornet.org/LOCO#hasPlay">http://www.lornet.org/LOCO#hasPlay</a></td>
<td>PROJECT</td>
<td>PROJECT WORK</td>
</tr>
<tr>
<td><a href="http://www.lornet.org/LOCO#hasAssignedRole">http://www.lornet.org/LOCO#hasAssignedRole</a></td>
<td>STUDENT</td>
<td>ROLE</td>
</tr>
<tr>
<td><a href="http://www.lornet.org/LOCO#hasAssignedActivity">http://www.lornet.org/LOCO#hasAssignedActivity</a></td>
<td>STUDENT</td>
<td>LEARNING ACTIVITY</td>
</tr>
<tr>
<td><a href="http://www.lornet.org/LOCO#title">http://www.lornet.org/LOCO#title</a></td>
<td>LEARNING</td>
<td>PROJECT WORK</td>
</tr>
<tr>
<td><a href="http://www.lornet.org/LOCO#hasLearningObjective">http://www.lornet.org/LOCO#hasLearningObjective</a></td>
<td>ACTIVITY TYPE</td>
<td>COMPETENCY</td>
</tr>
<tr>
<td><a href="http://www.lornet.org/LOCO#hasActivityDescription">http://www.lornet.org/LOCO#hasActivityDescription</a></td>
<td>ACTIVITY</td>
<td>String value</td>
</tr>
<tr>
<td><a href="http://purl.org/vocab/aiiso/schema#teaches">http://purl.org/vocab/aiiso/schema#teaches</a></td>
<td>SUPERVISOR,</td>
<td>COURSE</td>
</tr>
<tr>
<td><a href="http://www.lornet.org/LOCO#hasResource">http://www.lornet.org/LOCO#hasResource</a></td>
<td>ACTIVITY</td>
<td>LEARNING CONTENT</td>
</tr>
<tr>
<td><a href="http://linkedscience.org/teach/ns/#studentGroup">http://linkedscience.org/teach/ns/#studentGroup</a></td>
<td>COURSE</td>
<td>STUDENT GROUP</td>
</tr>
<tr>
<td><a href="http://xmlns.com/foaf/spec/#name">http://xmlns.com/foaf/spec/#name</a></td>
<td>PROJECT</td>
<td>String value</td>
</tr>
<tr>
<td><a href="http://www.w3.org/2004/02/skos/core#prefLabel">http://www.w3.org/2004/02/skos/core#prefLabel</a></td>
<td>SKILL</td>
<td>String value</td>
</tr>
<tr>
<td><a href="http://courseware.rkbexplorer.com/ontologies/courseware#has-assessment-method">http://courseware.rkbexplorer.com/ontologies/courseware#has-assessment-method</a></td>
<td>COURSE</td>
<td>COURSE EVALUATION</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EXTENDED EXAMINATIO</td>
</tr>
<tr>
<td>URL</td>
<td>Category</td>
<td>Type</td>
</tr>
<tr>
<td>--------------------------------------------------------------------</td>
<td>-----------------------</td>
<td>------------</td>
</tr>
<tr>
<td><a href="http://dublincore.org/lrmi-terms/#learningResourceType">http://dublincore.org/lrmi-terms/#learningResourceType</a></td>
<td>LEARNING CONTENT</td>
<td>String value</td>
</tr>
<tr>
<td><a href="http://dublincore.org/lrmi-terms/#isBasedOnUrl">http://dublincore.org/lrmi-terms/#isBasedOnUrl</a></td>
<td>LEARNING CONTENT</td>
<td>String value</td>
</tr>
<tr>
<td><a href="http://dublincore.org/lrmi-terms/#educationalUse">http://dublincore.org/lrmi-terms/#educationalUse</a></td>
<td>LEARNING CONTENT</td>
<td>String value</td>
</tr>
<tr>
<td><a href="http://dublincore.org/lrmi-terms/#interactivityType">http://dublincore.org/lrmi-terms/#interactivityType</a></td>
<td>LEARNING CONTENT</td>
<td>String value</td>
</tr>
<tr>
<td><a href="http://dublincore.org/lrmi-terms/#timeRequired">http://dublincore.org/lrmi-terms/#timeRequired</a></td>
<td>LEARNING CONTENT</td>
<td>String value</td>
</tr>
<tr>
<td><a href="http://dublincore.org/lrmi-terms/#useRightsUrl">http://dublincore.org/lrmi-terms/#useRightsUrl</a></td>
<td>LEARNING CONTENT</td>
<td>String value</td>
</tr>
</tbody>
</table>
5 Semantic annotation environment

This section presents the semantic annotation environment that was developed for the purpose of the project. The following sub-sections elaborate on the architecture of the environment as well as the tools and functionalities it supports, as also mentioned in Section 3.2. Finally, the section provides detailed description of the tool’s usage through user scenarios.

5.1 Semantic annotation environment architecture

The semantic annotation environment was developed using the Eclipse IDE, the Java programming language, the AngularJS JavaScript library and the Bootstrap framework. Furthermore, the development stores the resulted annotated data in a RDF store called Fuseki, and utilizes the Pubby tool for publishing the annotated learning objects as Linked Data, as shown in Figure 19.

![Figure 19 Semantic annotation environment architecture](image)

The environment includes four main tools that comprise the semantic annotation tool, as follows:

- Semantic annotator. The user interface provides two columns with a structured visualization of the SALO model’s concepts that the stakeholders should populate with data relevant to the learning object they want to annotate. One column includes the SALO concepts for a learning object represented as HTML elements and the other column includes more detailed terms for describing in more detail the learning resource relevant to the object.
The data that is input by the users are then sent to the back-end via web services, where they get annotated using the PBL3.0 custom vocabulary described in Section 4.3. This annotation is then stored in an RDF store for future discovery through SPARQL queries using the SGVizler JS library and is also exposed in Turtle format to the users for future exploitation.

- Semantic publisher. Once the data is annotated, they are automatically linked to similar entities from external resources such as DBpedia and the ESCO ontology, and are published as linked data using the Pubby linked data tool.

- Semantic browser. Apart from adding annotations, the users can also search existing annotations that have been saved in the RDF store. This way, they can discover similar learning objects that have already been annotated and enrich their knowledge. The search results provide links of each published identified object, which users can visit and browse the object through any web browser.

- Semantic enricher. Based on the PBL step that the annotated learning object belongs to, the semantic annotation tool exposes RDF and published linked data that are related to this step, based on the concepts and instances that are included in the PBL_LA framework of D1.3. These instances were RDFized using the OpenRefine tool, linked with corresponding DBpedia entities where applicable, and published as linked data with the Pubby tool. For example, if a teacher annotates a learning object which belongs in the PBL step “Group creation”, the enriched data that will show up will suggest him with activities that he could use in this step, data that is usually generated during this step, LA methods he could employ and the corresponding tool he can use to perform this LA method.

### 5.2 Semantic annotation environment overview

The semantic annotation environment’s interface is shown in Figure 20. The interface is divided into two main panels, two visible from the start, and two visible only when relevant information is entered by the user. These panels represent the environment’s tools and they involve:

- Searching of existing learning objects
- Discovery of existing learning objects as published linked data
- Annotation of new learning object
- Enrichment of knowledge on the PBL_LA paradigm based on each PBL step.

The following sub-sections will describe each of the above panels and the functionalities they support.

As shown in Figure 20, the first panel / tool (Semantic browser) regards the search of existing learning objects. This includes a text field where users can provide a keyword for searching learning
objects that have been previously semantically annotated and saved in an RDF store. The second panel / tool (Semantic publisher) generates the search results as links, where each link is a published linked version of the data as performed by the Semantic publisher.

The third panel / tool (Semantic annotator) regards the annotation of new learning objects and is divided into two columns. The first column includes text fields, dropdowns and lists for each SALO important concept. Some concepts (e.g. student names) will be added in the next version of the environment. Each text field accepts free text from the user, and we provide representative examples under each label to guide the users on how to annotate their learning objects. Apart from free text, the following concepts allow a specific range of choices:

- SOLO taxonomy. A dropdown list is provided where the user must choose one specific layer of the SOLO taxonomy based on what is more appropriate for the learning object. We also provide a link to the SOLO taxonomy terms as a guiding reference.

- Phase. A list is provided with the three main pedagogical phases as set in the SALO model in D2.1., i.e. Problem analysis, Problem design, Project report. The user can choose one or multiple phases in order to indicate where the learning object belongs to.

- Activity structure type. A list is provided with the 9 main steps of the Aalborg PBL model as stated in D1.1. The user can choose one or multiple PBL steps that the annotated learning object is related to.

- Activity structure. A list is provided with the different structures that represent activities as stated in D1.1 and D1.3. The user can choose one or multiple activity structure that the annotated learning object is related to.
This column allows the semantic annotation of the learning resource / content that corresponds to the learning object. The concepts chosen are specified in the LRMI vocabulary which specializes in learning resources' semantic description. The majority of the concepts can be annotated through

http://195.251.218.39:8080/pbl3_semanot/
free text, apart from the Interactivity type, which is a list because, based on the vocabulary, it can include only the values active, expositive, or mixed.

Finally, the forth panel / tool (Semantic enricher) appears once a user has annotated a learning object, where the tool dynamically and automatically presents a table with knowledge from the PBL_LA framework that is related to the PBL step annotated for and related to the user’s learning object.

The following sub-sections elaborate in more detail the tools within the environment as case scenarios.

5.2.1 Scenario 1: Semantic annotator - Annotation of a learning object

Let us assume that we want to annotate the learning object that has the information shown in Figure 21, which is the same object as the case study presented in D2.1.
Once we fill in all the information we want to annotate, we press the button “Annotate learning object”. The tool receives the inputted data and proceeds to create RDF statements for each concept based on what the concept represents, as shown in Figure 22. This data is automatically saved in a Fuseki RDF store in the back-end in order to support future discoverability for the semantic objects.

Figure 21 Semantic annotation of learning object

Figure 22 Semantically annotated learning object
Figure 23 presents a closer look at the RDF data. The main subject of our annotation is the learning object, which is assigned a dynamic identifier (URI). In this example, the object’s URI is `<http://pbl3_semannot/28Cohort2015ComputerGr28>`, which is a combination of the following elements:

- A base URI used in all identifiers for all concepts for a common root: http://pbl3_semannot
- A number derived from a random generator function to ensure that there will be no duplicate URIs
- A combination of substrings from the SALO concepts “Cohort” and “Course” to provide contextual understanding for the user
- A number derived from a random generator function to ensure that there will be no duplicate URIs

The learning object is initially annotated with a Class that describes what type of information it is, as follows:

**Subject**: `<http://pbl3_semannot/28Cohort2015ComputerGr28>`

**Property**: `<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>`

**Object**: `<http://data.opendiscoveryspace.eu/lom_ontology_ods.owl#LearningObject>`

We proceed to assign relations to the object using statements that have the same subject `<http://pbl3_semannot/28Cohort2015ComputerGr28>` and property – object groups. Some representative examples are as follows:

- `<http://purl.org/vocab/aiiso/schema#part_of> <http://pbl3_semannot/28Cohor28>` which shows that the learning object is a part of an element with the URI `<http://pbl3_semannot/28Cohor28>`. This element is the Cohort that was also annotated.
- `<http://purl.org/vocab/aiiso/schema#part_of> <http://pbl3_semannot/28Media28>` which shows that the learning object is a part of an element with the URI `<http://pbl3_semannot/28Media28>`. This element is the Programme that the learning object is related to.
- `<http://data.opendiscoveryspace.eu/lom_ontology_ods.owl#classification> <http://pbl3_semannot/28Prob28>` which shows that the learning object belongs to a classification that is represented by an element with the URI `<http://pbl3_semannot/28Prob28>`. This element is the Phase that this object belongs to.
- `<http://data.opendiscoveryspace.eu/lom_ontology_ods.owl#educationalInteractivityType> <http://pbl3_semannot/28Anal28>` which shows that the learning object is related to an educational interactivity type that is represented by an element with the URI `<http://pbl3_semannot/28Anal28>`. This element is the Activity structure type (PBL step) that the learning object is related to.
- <http://data.opendiscoveryspace.eu/lom_ontology_ods.owl#educationalInteractivityType>
  <http://pbl3_semanannot/28Desi28> which shows that the learning object is related to an educational interactivity type that is represented by an element with the URI <http://pbl3_semanannot/28Desi28>. This element is the second Activity structure type (PBL step) that the user added.

Figure 23 Part of semantically annotated learning object

Each of the concepts available in the interface is annotated on its own and is assigned values and relations with other concepts based on the SALO model. For example, the learning resource annotated has the URI: <http://pbl3_semanannot/28LO12328> and is described using the following statements:

- <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
  <http://linkedscience.org/teach/ns/#Material>, which shows that the learning resource belongs to the RDF Class “Material” from the Teaching core vocabulary and therefore now computers can understand the meaning behind this element.

- <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
  <http://www.lornet.org/LOCO#Resource>, which shows that the learning resource can also be assigned the RDF class “Resource” from the IMS-LD vocabulary. This way, we increase the possibility that the URI <http://pbl3_semanannot/28LO12328> will be discovered, since it can be searched with multiple types.

- <http://www.lornet.org/LOCO#title> "LO123-HowtomakegoodUserInterfaces", which shows that the learning resource has a specific title with the value "LO123-HowtomakegoodUserInterfaces".

- <http://dublincore.org/dcx/lrmi-terms/#learningResourceType> "Presentation", which shows that the learning resource has a specific type with the aforementioned string value.

- <http://dublincore.org/dcx/lrmi-terms/#isBasedOnUrl>
  "http://usabilitypost.com/2009/04/15/8-characteristics-of-successful-user-interfaces/", which shows that the learning resource is based on another resource’s URL that has the aforementioned string value.
• <http://dublincore.org/dcx/lrmi-terms/#educationalUse> "Theoryreading", which shows that the learning resource has a specific educational use which is depicted by the aforementioned string value.

• <http://dublincore.org/dcx/lrmi-terms/#interactivityType> “Active”, which shows that the learning resource supports a specific predominant mode of learning that is depicted by the aforementioned string value.

• <http://dublincore.org/dcx/lrmi-terms/#timeRequired> "45minutes", which shows that the learning resource requires 45 minutes to work through.

• <http://dublincore.org/dcx/lrmi-terms/#useRightsUrl> "http://creativecommons.org/licenses/by/3.0/", which shows that the learning resource follows the permission guidelines depicted in the aforementioned string value.

All the information that has been semantically annotated can now be accessed and understood by both humans and computer machines. This can allow meaningful design of courses, monitoring of learners’ progress, assessment of performance and utilization of resources.

5.2.2 Scenario 2: Semantic browser - Discovery of existing learning objects

Users can use the semantic annotation tool to search for any existing learning objects that have already been annotated, increasing discoverability of information in the education and training sector for any domain. Figure 24 shows the panel that supports this functionality, where the user can submit a keyword for what they want to discover.

![Figure 24 Search existing learning objects](image)

The tool searches based on the keyword set all concepts of the learning objects that have been saved on the RDF store. For example, if we search with the keyword “Media”, the tool generates a list of learning objects that include one or multiple concepts with the value “Media”, as shown in Figure 25.
The links shown have been generated on the back-end using the Pubby linked data tool, and allow the publishing of each learning object as linked data and the web browsing of all elements of an RDF resource. Figure 26 shows that the link in the search result depicts a web representation of the RDF statements describing the learning object that includes the value “Media”.

More specifically, based on the above information, users and the computer can understand that the above concept is a programme (through the rdf:type property and the corresponding objects teach:StudyProgram, aiiso: Programme). Furthermore, we understand that the concept’s name is “Medialogy Bachelor”. We can access this information in different formats as well, as shown in the links available at the bottom of the page. Thus, if we want, we can download this data as RDF or Turtle, and save it to another RDF store or upload them to another repository, increasing its mainstreaming and contributing to its discoverability. Users can also upload the XML/RDF version of the data into LMS or other e-learning systems that allow XML / RDF import, and inspect and consult the learning objects during learning.

Additionally, users can use the published data’s link and link it to other RDF learning objects, enriching existing information for any domain.

Figure 27 shows another example of searching existing learning objects. This time, we use the “Comp” keyword in order to find anything that includes this string value. We get the link shown in the Figure as a result.
Vising the link, we are redirected to the published linked data version of the learning object, as shown in Figure 28.

The information shows that the concept found is a course with the title “Computer Graphics Programming”. Additionally, we can see that this course has as course objective an element that is further described by a URI, and belongs to an academic term that is also further described by the URI <http://egov.dai.uom.gr:8085/semlinked/page/28Semester528>. The user can visit these additional URIs and get further information about these semantically annotated elements.

5.2.3 Scenario 3: Semantic enricher - Receiving enriched information per PBL step

The semantic annotation tool aims to guide educators and other interested users in discovering information and gaining knowledge that will help them improve learning and teaching experiences for all stakeholders independent of the domain. To this end, we exploit the concepts (PBL step, activities, LA methods, LA technologies, generated data) and instances that were populated within D1.3 and the PBL_LA framework and visualize the results that are relevant to the user’s input, taking into consideration the PBL step (Activity structure type) chosen.

More specifically, we proceed to transfer the instances into a relational database, we create RDF versions of this data using the OpenRefine tool, we create links with the DBpedia database whenever we encounter similarities and we publish these data using the Pubby linked data tool.

For example, we assume that we annotate the learning object shown in Figure 29, where we state that the object is related to the PBL step “Design”. The semantic annotation tool performs searches on the back end in order to find all the different concepts and corresponding instances of the PBL_LA framework that are related to the PBL step “Design”.
Figure 29 Annotation of learning object

Figure 30 shows all the PBL_LA framework’s elements that have been linked with this specific PBL step. Each row shows a different set of knowledge for the user, e.g. in the 5th row we see that within the “Design” PBL step, the educator can apply the activity “Design strategy”, the LA method “Personalization”, the technology “LeMo” and the data generated is “Content of design requirements”.

---

<table>
<thead>
<tr>
<th>Context of the Learning Object</th>
<th>Annotation of Learning Resource</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title of learning resource:</strong></td>
<td>Creating card charts</td>
</tr>
<tr>
<td><strong>Learning resource is based on:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Educational use:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Interactivity type:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Typical time required:</strong></td>
<td></td>
</tr>
</tbody>
</table>

---

[Context of the Learning Object continues here]
The users can thus become aware of what activities are more relevant for each PBL step, which methods they can apply for successfully implementing LA, which LA tools to use in order to retrieve the target visualizations and what data is generated that can be analysed for insights on the learning process.

As shown in the above Figure, each instance is shown as a link. Each link redirects to the published linked data version of the instance created on the back-end with the Pubby linked data tool. For example, Figure 31 presents the LOCO-Analyst LA tool and all of the related RDF concepts that describe it.

We can see that this concept is a “Tool” or “Software”, has the name “LOCO-Analyst” and belongs to the type “Learning Analytics”.

Figure 32 shows another example of browsing a PBL_LA framework’s concept, i.e. the “Assessment” LA method.
We can also see that this concept is linked with a similar entity from the DBpedia database using the property “owl:sameAs”. This allows users to visit the linked entity through its URI

<http://dbpedia.org/resource/Category:Educational_assessment_and_evaluation_theorists> and gain more insights and information about it.
6 Conclusion

This deliverable aimed to present the semantic annotation environment that was developed taking into consideration the SALO model designed in D2.1. The environment can be used by all learning content providers such as MOOC providers, OER providers, educational institutions, national organizations etc., independent of domain and sector. This will foster mainstreamed content discoverability and linkage of learning resources with the Web of Data.

To this end, the report initially presented a review carried out on 10 existing open-source software tools that are available for annotating educational resources. The report listed the different functionalities that are currently supported (e.g. manual-automatic annotation, plain unstructured text processing, limited RDF vocabularies, limited linking of data with other resources, no ability to import RDF vocabularies or export RDF data etc.) and assessed which of these are in accordance with the functionalities that should be supported by the semantic annotation tool (automatic annotation, structured input of data based on the SALO model, usage of educational RDF vocabularies, export of RDF data for further re-use, linking of data with other resources etc.).

Based on this requirements specification, we concluded that no existing tool can be used as-is for the project tool’s needs, and the configuration of any of the tools would require complex and time-consuming effort with no guarantee for a successful re-design of a new version of an existing tool. This led to the decision of developing a new tool with a user-friendly interface that allows the visualization of the SALO model’s concepts and performs most of the technical effort on the back-end, so that it can be used by any stakeholder independent of the domain of interest and technical background.

The report also presented the review carried out on existing RDF vocabularies that focus on or include classes and properties of elements within the educational domain. This review studied 11 vocabularies that provide RDF terms for annotating academic institutions, courses, learning objects, learning resources, learning designs, people etc. We derived a sub-set of classes found in the studied vocabularies which correspond and can be used to describe all elements of the SALO model and learning resources. We also derived a sub-set of properties found in the vocabularies that can describe the relations formed amongst the SALO concepts so that they are understandable by humans and computer machines. These two lists comprised the PBL3.0 vocabulary that is used to annotate learning objects and resources with semantic meaning.

The semantic annotation tool that was developed supported the following main functionalities:

- Description of learning objects and resources in accordance with the concepts of the SALO model and automatic semantic annotation of each concept based on the relevant PBL3.0 vocabulary classes and properties.

- Automatic linkage of the annotated learning objects with relevant resources such as DBpedia, ESCO etc.
• Export of the semantically annotated data in Turtle format and automatic saving to an RDF store for further re-use.

• Automatic publishing of the semantically annotated data as linked data for further linkage with external resources.

• Search of existing learning objects and their visualization as published linked data.

• Automatic provision of published linked data from the PBL_LA framework designed in D1.3 (e.g. activities, LA methods, LA technologies, data generated) based on the PBL step entered for the learning object during annotation.

The effort required from the user when annotating with the project’s tool is minimum, as the majority of the workload is carried out in the back-end so that the tool can be mainstreamed to any professional domain and sector and used by any stakeholder.

The tool will be tested for validation by stakeholders from both academia and business and from different domains. Any possible improvements and corrections required based on the testing will be carried out so that the final version of the tool by the end of the project will be optimized. Future works also includes the identification of ways the environment and its results can be incorporated within e-learning platforms such as LMSs, MOOCs etc.
D2.2 Semantic Annotation Tool

References


