LEARNING PERSONALISATION
APPROACH BASED ON RESOURCE
DESCRIPTION FRAMEWORK

Andrius Berniukevičius

October 2016

Abstract

This report is aimed to analyse the problem of learning personalisation applying Resource Description Framework (RDF) standard model. Research results are two-fold: first, the results of systematic literature review on RDF application in learning are presented, and, second, RDF-based learning personalisation approach is proposed. First of all, systematic literature review was conducted in Thomson Reuters Web of Science database and using Semantic Scholar search tool. The review has shown that RDF data model is based upon the idea of making statements about web resources in the form of subject–predicate–object expressions. These expressions are known as triples in RDF terminology. The subject denotes the resource, and the predicate denotes traits or aspects of the resource and expresses a relationship between the subject and the object. The review revealed that linked data and triples-based RDF standard model could be successfully used in education. On the other hand, although linked data approach and RDF standard model are already well-known in scientific literature, only few authors have analysed its application to personalise learning process, but many authors agree that linked data and RDF-based learning personalisation trends should be further analysed. Original RDF-based learning personalisation approach is also presented in the report. According to this approach, RDF-based personalisation of learning should be based on applying students’ learning styles and intelligent technologies. The main advantages of this approach are analysis of interconnections between students’ learning styles and suitable learning components (i.e. learning resources, learning methods and activities, learning tools and technologies etc.) based on using pedagogically sound vocabularies of learning components, experts’ collective intelligence, and intelligent technologies (e.g. expert evaluation, ontologies, recommender systems, software agents etc.). This pedagogically sound RDF-based personalisation approach is aimed at improving learning quality and effectiveness.

Keywords: learning personalisation, Resource Description Framework, linked data, learning styles, intelligent technologies
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>Error! Bookmark not defined.</td>
</tr>
<tr>
<td>Systematic review</td>
<td>4</td>
</tr>
<tr>
<td>Findings of the Systematic review</td>
<td>7</td>
</tr>
<tr>
<td>Learning personalisation approach based on Resource Description Framework</td>
<td>Error! Bookmark not defined.</td>
</tr>
<tr>
<td>Conclusion</td>
<td>10</td>
</tr>
<tr>
<td>Literature</td>
<td>11</td>
</tr>
</tbody>
</table>
1 Introduction

The aim of the report is to analyse the problem of learning personalisation applying Resource Description Framework (RDF) standard model. The results of the performed systematic review on RDF and semantic description in learning and personalisation are discussed, and an original learning personalisation framework addressing student’s learning styles and based on RDF and intelligent technologies is presented.

2 Systematic review

The main goal of the systematic review was to find out the state-of-the-art strategies and semantic web approaches to personalise learning by identification suitable learning objects (LOs) for student in conformity with his/her learning styles.

The following research questions have been raised to perform systematic literature review:

1. What Semantic Web strategies and approaches are used to semantically describe and retrieve learning resources?
2. How Semantic Web technologies, like RDF, can be used to support learning personalisation?

Systematic literature review was performed on March 18, and updated on May 12, 2016. In order to get wider view on semantic web technologies that can be used for our goal formulated above, we did not include “learning styles” into search keywords. The search was undertaken in Thomson Reuters Web of Science database and then using Semantic Scholar tool (https://www.semanticscholar.org/).

Over time, several competing metadata standards and educational metadata schemas have been proposed, e.g. the widely adopted IEEE LOM (Learning Object Metadata), ISO/IEC MLR – ISO 197884 (Metadata for Learning Resources – MLR), Dublin Core, IMS etc. However, the adoption of a sole metadata schema is usually not sufficient to efficiently characterize learning resources. A number of taxonomies, vocabularies, and application profiles (AP) are defined to address this problem. Repositories also exploit diverse interface mechanisms such as OAI-PMH or SQI (Dietze et al, 2013). Gabor et al (2013) state that existing metadata standards (LOM and similar) lack educational feature descriptions about educational needs, especially if we deal with multimedia LO. Therefore, metadata are usually augmented with additional information by application of semantic web approaches. Navarro et al (2013) argue that certain complex LOs need to be complemented with different types of domain-dependent information for their pedagogical planning and retrieval. The authors propose a theoretical approach that permits to dynamically change domain-dependent information schemas and use a single LO repository for classifying and enriching learning. The approach is based on a meta-relational model for the dynamic definition of specific domain-dependent relational database schemas used for classifying and enriching LOs.

A number of studies discuss metadata usage in conjunction with ontologies to ensure effectiveness in LO description, search, and retrieval. For instance, Solomou et al (2015) present an ontology model for the developed IEEE LOM educational AP with
enhanced technological and educational fields, aimed to improve the discovery and retrieval of LOs within intelligent e-learning systems. The authors develop a personalised learning system that uses this educational AP and its ontological representation in order to offer advanced services to students. Huang et al (2013) present an approach of the construction of knowledge retrieval and navigation system, based on ontology-based knowledge organization model for learning resources. The knowledge organization model includes the parts of subject domain ontology, metadata extraction, and automatic classification for LOs. In this model, the metadata of LOs are semantically described basing on the domain ontology, and in this way the knowledge organization and navigation for learning resources is implemented.

The design of recommender systems is an ongoing research area where intelligence is incorporated into web content systems to be able to provide recommendations to students on the basis of their learning preferences, i.e. based on their learning profiles (Sunil and Saini, 2013). The authors discuss the design of a recommender system based on ontology, mapped to the learning content, and learner profiles created in the system. In order to provide support for personalised access to the resources that exist in open educational repositories, Almudena et al (2014) propose the recommendation strategy combining a description of the LOs based on metadata standards enriched by ontology-based semantic indexing, and contextual information about the user.

Resource Description Framework (RDF) proves to be a widely used semantic web framework to solve the problems we address in this report. The Semantic Web is a collection of components working together so that a machine is able to process and understand information. In order for this vision to be implemented, formal standards for representing and interpreting data are used, including the RDF and machine processible ontologies (Algosaibi and Melton, 2014). RDF as a recommended format for representing data is one of the most important contributions to the semantic web concept. It brings opportunity to develop new approaches to data analysis. The main idea is to represent each piece of data as a triple: "subject-proposition-object", where the "subject" is an entity being described, "object" is an entity that describes the subject, and the "proposition" is a connection (a relation) between subject and object. A subject of one triple can be an object of another triple, and vice versa. This gives a network of interconnected triples by Chen and Reformat (2014). RDF data can be analysed with various query languages, e.g. SPARQL.

Chen (2015) proposes an approach to transform metadata from equivalent lexical element mapping into semantic mapping with contextual relationships, based on RDF. RDF is used as a crosswalk model to represent the contextual relationships implicitly embedded between described objects and their elements. The semantic, hierarchical, granular, syntactic and multiple object relationships are included to achieve semantic metadata interoperability at the data element level. RDF-based expressions let manifest into a semantic representation the sets of shared terms, contextual relationships between described objects and their metadata elements. The author has developed nine types of mapping rules to achieve a semantic metadata crosswalk.

By combining semantic descriptions already lying or implicit within the descriptive metadata, reasoning-based or semantic searching of these collections can be enabled and produce novel possibilities for content browsing and retrieval (Solomou and Koutsomitropoulos, 2015). The authors employ semantic searching techniques on MII-DS-09P-14-5.
digital repositories and introduce a methodology to pragmatically evaluate and get measurable results of the semantic searching in such scenarios. Chen and Reformat (2014) suggest building categories based on similarity of entities contained in the data to provide more benefits in addition to properties indicating data type and subject, provided in RDF-based data.

There is a wide variety of technologies available to deal with exposing, sharing and integrating educational web data, but according to the number of publications in the recent years, it can be stated that Linked Data based approaches have gained a lot of attention and started realising the vision of highly accessible and web-wide reusable learning resources by providing the standards, tools, and Web infrastructure to expose and interlink educational data at web-scale (Dietze et al, 2013).

Semantic Web technologies and Linked Data are changing the way information is stored, described and exploited (Chicaiza et al, 2014). The “Linked Data” term refers to a set of best practices for publishing and connecting structured data on the Web. Chicaiza et al (2014) deal with improvement of the associations between learning subjects, areas and topics, including semantic relations and recommendations about resources for learners. The advantages of linked data web are used to support semi-automatic classification of educational resources. The relations of the resources are encoded in RDF language and stored in the repository, a query language is used to retrieve data, and the knowledge of organizational systems and linked data is used to classify the web resources according to the domain.

The survey presented in (Dietze et al, 2013) is one of the first comprehensive surveys on the topic of Linked Data for education and provide an extensive overview of the Linked Data approaches for technology-enhanced learning. It aims to provide rich and well-interlinked data for the educational domain, using the existing technology-enhanced learning data on the web by allowing its exposure as linked data, and using automated enrichment and interlinking techniques.

New opportunities for relating learning resources identified by URIs combined with the usage of RDF as a lingua franca for describing them are arising with the emergence of web of data (Rajabi et al, 2015). The authors present an approach for exposing existing IEEE LOM metadata as Linked Data. IEEE LOM elements (simple and structured, as well as with multiplicity) are transformed into XML representation and RDF triples (subject, predicate and object). The metadata are linked to the datasets in LOD (Linking Open Data), e.g. DBPedia. A case study and a reference implementation along with an evaluation have proved the concept of this mapping. Selected queries passed a performance testing on both relational database and triple store.

Vert and Andone (2014) suggest using Linked Data principles to discover, integrate and reuse online learning resources, using standards and principles proven to foster web interoperability, like RDF and SPARQL. The authors concentrate on the solutions for open educational resources. The publishing of resources as Linked Data is done in several steps: selection of data sources, usage of vocabularies and ontologies to model the data, conversion to the RDF data model, including cleaning of the data, publishing the semantic-enriched data to linked learning resources repositories and consuming the data, usually through SPARQL endpoints.
Dessi and Atzori (2016) address the problem of ranking among properties of the entities used in RDF datasets, Linked Data and SPARQL endpoints. The authors provide applications for property tagging and entity visualisation, and propose to apply Machine Learning to Rank techniques to the problem of ranking RDF properties. The major advantages of the approach are: flexibility/personalisation, speed, effectiveness.

Linking Open Data (LOD) cloud is a collection of linked RDF data with over 31 billion RDF triples. Accessing linked data is a challenging task due to ontology schema specifics in each data set (Zhao and Ichise, 2013). To solve this issue, the authors propose an automatic method to integrate different ontology schemas: Mid-Ontology learning approach that can automatically construct an ontology, linking related ontology predicates (class or property) in different data sets. The approach consists of three main phases: data collection, predicate grouping, and Mid-Ontology construction. Experiments show that our Mid-Ontology learning approach successfully integrates diverse ontology schema, and effectively retrieves related information.

Chung and Kim (2015) design an ontological semantic model of achievement standards (the standards, providing guidelines about what has to be taught and assessed by teachers and what has to be studied and achieved by students). Mapping rules are defined to formalize the semantic model to RDF/OWL specification. The approach is based on Linking Open Data. The proposed semantic model is used to create linked data profile searching and browsing, sharing, modification history tracing, learning resource linking.

While personalisation, adaptation and recommendation are central features of web-based educational environments, recommender systems apply information retrieval techniques to filter and deliver learning resources according to user preferences and requirements (Taibi et al, 2013). The authors state that, however, the suitability of possible recommendations is fundamentally dependent on the available data, i.e. metadata about learning resources and data about the users. To solve the limitation in quantity and quality of both types of data, the Linked Data movement has become very active over the recent years. Taibi et al (2013) propose a large-scale educational dataset, generated by exploiting Linked Data methods and applying clustering and interlinking techniques to extract, import and interlink a wide range of educationally relevant data.

Research work, presented in (Morshed et al, 2013) is aimed to develop knowledge recommendation system for the Linking Open Data Cloud using semantic machine learning approach. Knowledge is stored in a triplestore using RDF triples format (subject, predicate, and object) along with the complete metadata. The authors argue that such a RDF representation made the developed intelligent knowledge base very flexible to integrate with the Linking Open Data (LOD) cloud.

3 Findings of the Systematic review
The review has shown that many authors agree that “pure” metadata approaches to describe learning objects lack flexibility to address the issues of personalisation. RDF provides facilities for data merging even if the underlying schemas differ. RDF data model is based upon the idea of making statements about web resources (LOs) in the form of subject–predicate–object expressions. RDF extends the linking structure of the resources to use URIs to name the relationship between “subject” and “object” as well as the two ends of the link (this is referred to as a “triple”). The “subject” denotes the resource, and the “predicate” denotes traits or aspects of the resource and expresses a relationship between the “subject” and the “object”.

The review has also revealed that linked data and triple-based RDF standard model could be successfully used in education. Although Linked Data approach and RDF standard model are already well-known in scientific literature, only few authors have analysed its application to personalise learning process. Thus, Almudena et al (2014) propose the recommendation strategy combining a description of the LOs based on metadata standards enriched by ontology-based semantic indexing, and contextual information about the user. According to Taibi et al (2013), the suitability of possible recommendations is fundamentally dependent on the available metadata about LOs and data about the users. Solomou et al (2015) develop a personalised learning system that uses IEEE LOM educational AP and its ontological representation in order to offer advanced services to students.

On the other hand, many authors agree that linked data and RDF-based learning personalisation trends should be further analysed.

4 Learning Personalisation Approach Based on RDF and Intelligent Technologies

The proposed personalisation approach is based on Kurilova et al (2014) and Kurilovas (2015). According to this approach, RDF-based personalisation of learning should be based on applying students’ learning styles and intelligent technologies. The main advantages of this approach are analyses of interlinks between students’ learning styles and suitable learning components (i.e. LOs, learning methods and activities, learning tools and technologies etc.) based on using pedagogically sound vocabularies of learning components, experts’ collective intelligence, and intelligent technologies (e.g. expert evaluation, ontologies, recommender systems, software agents etc.).

According to this approach, RDF triples should interlink (1) LOs (“subject”) including metadata, (2) contextual information about particular learner (“object”), and (3) suitable learning methods, activities and tools (“predicate”). In this RDF triple, the “subject” denotes the resource, and the “predicate” denotes traits or aspects of the resource and expresses a relationship between the subject and the object.

According to Kurilovas (2015), implementation of this approach consists of the following stages:

- Creating learners’ dynamic profiles/models according to their learning styles and other features.
• Creating interlinks and ontologies to establish suitability of learning components to particular learning styles.
• Creating recommender system to recommend suitable learning components to particular students.

Thus, creating interlinks and ontologies to establish suitability of LOs, learning methods/activities, and learning tools/environments is the following stage after creating students’ profiles.

In (Kurilovas et al, 2014), personalisation is analysed in terms of suitability of LOs, teaching/learning methods and learning activities to particular learning styles according to Honey and Mumford (2000) learning styles model. Analysis of interrelations between learning styles, learning activities, teaching methods, and LOs types is presented in (Kurilovas et al, 2014). After that, an example of interlinks between teaching/learning methods (M) and learning resource types (T) for problem-solving learning activity was presented:

**Figure 1.** Example of interlinks between teaching/learning methods and LOs types for problem-solving learning activity (according to (Kurilovas et al, 2014))

Further on, ontology example is presented in (Kurilovas et al, 2014). This ontology presents a query for finding suitable learning activities by methods (i.e. “Problem Solving” activity could be found using “Blogging” teaching/learning method).
In (Kurilovas et al, 2014), the authors have analysed only “Learning Resource Type” LOs metadata field. These interlinks could be enriched by analysing several additional fields of LOs metadata according to IEEE LOM standard such as “Structure”, “Format”, “Interactivity Type”, and “Interactivity Level” (Dorça et al, 2016).

According to Kurilovas (2015), after interlinking and ontologies creation stage, recommender system should be created to link students’ personal data in their profiles, relevant LOs according to corresponding metadata fields, and learning activities and tools suitable to particular students according to their learning styles and other profiles’ data.

Interlinking and ontologies creation should be based on the expert evaluation results. Experienced experts should evaluate suitability of learning components (LOs, learning activities and tools) to particular students’ needs, e.g. learning styles. The higher suitability ratings the better learning components fit the needs of particular learners. Pedagogically sound vocabularies of learning components should be applied at this stage.

In (Juškevičienė and Kurilovas, 2014), an example of the method to create personalised recommender system was presented. The prototype of recommender system has been developed following the working principles of the knowledge-based recommender system. The domain knowledge was conceptualised in the ontology.
Recommender system should form the preference lists of the learning components according to the expert evaluation results of suitability of learning components and students’ data e.g. learning styles. Probabilistic suitability indexes should be identified for all learning components in terms of their suitability level to particular learners.

Thus, personalised learning packages/scenarios could be created for particular learners. A number of intelligent technologies should be applied to implement this approach, e.g. ontologies, recommender systems, intelligent agents, personal learning environments etc.

5 Conclusion

Systematic review presented in the report has shown that RDF data model is based upon the idea of making statements about web resources in the form of subject–predicate–object expressions (RDF triples). The subject denotes the resource, and the predicate denotes traits or aspects of the resource and expresses a relationship between the subject and the object. The review revealed that linked data and triples-based RDF standard model could be successfully used in education. On the other hand, although linked data approach and RDF standard model are already well-known in scientific literature, only few authors have analysed its application to personalise learning process, but many authors agree that linked data and RDF-based learning personalisation trends should be further analysed.

Original RDF-based learning personalisation approach is presented in the report. According to this approach, RDF-based personalisation of learning should be based on applying students’ learning styles and intelligent technologies. The main advantages of this approach are analysis of interlinks between students’ learning needs e.g. learning styles and suitable learning components (i.e. learning objects, learning methods/activities, learning tools/technologies etc.) based on using pedagogically sound vocabularies of learning components, experts’ collective intelligence to evaluate suitability of learning components to particular learners’ needs, and application of intelligent technologies (e.g. expert evaluation, ontologies, recommender systems, software agents etc.).

This pedagogically sound RDF-based personalisation approach is aimed at improving learning quality and effectiveness. The learning package (scenario, unit) of the highest quality for particular student means a methodological sequence of learning components with the highest Suitability Indexes. The level of students’ competences, i.e. knowledge/understanding, skills and attitudes/values directly depends on the level of application of high-quality learning packages in real pedagogical practice.

6 References


MII-DS-09P-14-5


