

Zorica Bogdanović¹, Marijana Despotović-Zrakić¹,
Miloš Milutinović¹, Mihajlo Anđelić¹, Suzana Milinović²
¹ University of Belgrade, Faculty of Organizational Sciences,
² Falc East d.o.o.

Model for Enhanced Data Management, Visualization, and Adaptation in e-learning

UDC: 37.018.43:004.738.5 ; 371.3:004
DOI: 10.7595/management.fon.2013.0030

In modern educational systems, the final outcome of the learning process heavily depends on providing learners with an access to the right information at the right time. An appropriate presentation and visualization of information can strongly influence the learning experience by allowing learners to explore an educational domain instinctively, gaining a better insight into its structure. The main goal of this paper is to enhance data management, visualization, and adaptation within an e-learning system. A model for semantic annotation of concepts present in learning management systems is presented, including the description of a structure of semantic applications that will utilize annotated resources to enhance data handling, browsing, searching, visualization, and adaptivity. The developed solution is based around the principles of modularity and independency among components.

Keywords: semantic web, ontologies, learning objects, adaptive e-learning.

1. Introduction

Intelligent semantic applications that utilize semantic data to provide new value for their users represent a final objective of explicit semantic declaration and knowledge structuring. Semantic annotation of information resources enables an easier search and customizing of results according to users' needs. Well-structured data can be visualized more easily and concisely, allowing for a better perception of the domain (Dadzie & Rowe, 2011). Visualizations and infographics are used to display data in various ways. Numbers are assigned semantics in the form of images, context, and stories.

E-learning is based on delivering adequate information and educational resources to students via network. However, many complexities arise with educational resources existing in different formats, encompassing different topics and varying in scope and complexity. The semantic web and the semantic annotation potentially offer a solution to these, and can stimulate further advances by simplifying the development of visualization and adaptivity schemes.

This paper describes the model for semantic annotation of educational materials and related concepts, as well as the structure of dependent services that extend the functionalities of current learning management systems (LMS). The goal of the model is to enable simple discovery, while allowing formulation and execution of advanced search queries, augmented with visualizations of result sets and domain structure and adaptation of results according to students' interests. Several components were also implemented and their descriptions and other remarks are given following the description of the model.

2. Literature Review

Learning materials and sources are primarily digital in nature in the case of e-learning. This allows them to be easily modified in order to adapt to the users' needs and a learning context. Student and professionally-created learning materials can be found on the Internet in various formats. The amount and diversity of available resources are becoming an issue for both the learners and the information systems that facilitate learning. Using a descriptive and standardized set of attributes (metadata), the aforementioned problems can be solved.

In order to reduce expenses, educational institutions design reusable learning materials (Downes, 2001), commonly by applying the concept of learning objects, small, compact educational units focused on achieving a specific learning outcome (Cisco Systems, 1999). Some of the characteristics specific to learning objects include their reusability, smaller scope, relative independence and also the possibility of aggregation (McGreal, 2004). The power of learning objects stems mostly from their heavy reliance on metadata to describe their technical and educational aspects. IEEE Learning Object Metadata (IEEE Learning Technology Standards Committee, 2002) is a prominent metamodel for description of learning objects. Another popular set of standards, based on the IEEE LOM, is Sharable Content Object Reference Model (SCORM) (Bohl, Scheuhase, Sengler, & Winand, 2002). SCORM learning objects are packed into content packages of specific structure (Advanced Distributed Learning, 2009a), and are based around the use of web-centric technologies, HTML pages and interactivity with the system (Advanced Distributed Learning, 2009b). Many resources like research papers, student-created materials, links to web pages, etc. do not naturally fit into this preconception. In order to ensure an easier and automated searching and indexing of diverse learning resources, a simpler, implementation-agnostic set of metadata is likely a better match. The primary example is the Dublin Core, a set of metadata attributes for description of information resources (Weibel, 2005a).

When learners are using a search functionality to find adequate resources, their queries might be imprecise, resulting in too large or too small result sets. This indicates that there is a need for a system that would organize the information in some useful, efficient way. When metadata is applied on all of the information resources, the search algorithms can efficiently concentrate on specific metadata fields, although errors in spelling and use of synonyms can still present problems.

Some metadata fields are simple and straightforward, often using values from a strictly defined set, e.g. „date” field in Dublin Core; others, like “subject” can contain freely inputted keywords or phrases that describe the subject or content of the resource (Dublin Core Metadata Initiative, 2012). Users can make mistakes or use different synonyms for the same concept, and some kind of a controlled vocabulary for terms used to describe a subject of a resource is required, which represents one of main concerns of classification based on subjects. Vocabularies, thesauri, ontologies, faceted classifications, as well as taxonomies represent methods of subject-based classification (Garshol, 2004). In general, controlled vocabularies only list allowed terms, while other methods define some relations between them, infusing them with additional semantics. Ontologies are the most powerful since they allow any type of concept and relation to be defined at will.

Ontologies are open vocabularies, models that describe the world using types, attributes, and relations. Any form of relationship between two concepts can be expressed using strongly typed relations. A short, concise definition of ontologies states that they are a “formal, explicit specification of a shared conceptualization” (Gruber, 1993). Ontologies are powerful and extendable and can be used to describe information objects, taking over a role of a meta-model. Resources annotated using well-structured technologies can be easily integrated, retrieved, and compared by both the human users and the machines.

Ontologies are tightly connected with the concept of semantic web and represent one of the layers of Berners’ layered cake of semantic web. Using ontologies on web allows the creation of a semantic web. The definitions of both concepts concentrate on structuring data in a way understandable to machines (Gómez-pérez & Corcho, 2002). However, they are often used without consistency, and encompass an entire array of technologies; a brief summary of related terms and enabling technologies is given in (Gerber, Barnard, & Merwe, 2006).

Ontologies applied in education and, specifically, learning management systems, could provide a basis for application of higher-level logic and better browsing and searching capabilities. By appropriately semantically annotating educational resources, actors (students, teachers), and other concepts in the system, existing processes can be simplified, and innovative methods of presentation can be applied. Applying visual methods of presentation like graphs and trees can abstract extraneous details to allow focusing on the desired topic. Visualization libraries are especially abundant in the domain of web technologies and can be plugged in into the existing web applications. Unannotated or weakly annotated information objects are not suitable for application of visualization techniques, and most attempts might amount to displaying a number of disconnected resource “islands”. If rich semantic attributes and relations are available, resources can be abstracted, elements and relations filtered or highlighted, and different perspectives generated, providing insight into a certain aspect of the analyzed resource set.

Visualizing learning resources in a coherent and legible manner can allow learners to better understand the domain of study. The visualization represents knowledge about the domain of study and can be used by the learners as a map, helping them gain an overview of their current level of knowledge and showing them potential avenues of advance. This helps them compose adequate search queries and discover new information resources (Dadzie & Rowe, 2011). However, appropriate visual encoding (mapping of dataset attributes to visual elements) needs to be selected taking the characteristics of purpose/task of visualization and of the human perception system in consideration (Munzner, 2009). (Shneiderman, 1996) gives a type by task taxonomy of information visualizations, describing types of datasets and corresponding possibilities for visualization. Current approaches for data browsing and visualization are given in (Dadzie & Rowe, 2011) where the importance of both the visual and textual representation is stated - text allows fine grained analysis, while visual browsing provides insight into the structure of a domain and allows for the bigger picture to be seen. Both approaches can be used in combination effectively.

Social media and networks are especially interesting for integration with learning management systems; some research works give values of more than 90% examinees using Facebook, with numbers of users on most such sites showing a general trend of growth (Meeker & Wu, 2013). Social media can be used for distribution of educational materials, receiving feedback from students, and even developing educational games and integrating them with existing learning management systems (Labus, Simi, Bara, Despotovi-Zraki, & Radenkovi, 2012).

3. Model for semantic annotation and semantic services in e-learning

In this paper, we describe a model for improving searchability and browsability of educational resources in a LMS by applying semantic annotation to educational resources, actors, and other concepts, and using the semantics to visualize data and enrich it with information gathered from social networks and semantic databases. The semantic web applications are realized in four steps – creating an ontology, serializing it using appropriate languages, marking up concrete instances of resources, and creating actual services that will use the semantic information. Since the focus is on the model and not on the implementation, this chapter describes the first and the fourth steps. Firstly, a generic, expendable educational ontology is defined and then a model of educational and supporting services based on that ontology is given.

3.1 Simple educational ontology framework (SEOF)

Properly designed ontologies can accurately describe any domain, but caution needs to be exercised with balancing their expressiveness and complexity. An existing framework of an upper, mid, and domain ontologies can be expanded to fit the needs of the domain, especially if a suitable ontology already exists. Another approach is to create a new, simpler ontology by only describing specific concepts. The former approach can introduce complexity in the form of unnecessary, generic concepts, while the latter inherently provides less interoperability between different systems. The interoperability can be improved by using another method, for instance basing the attributes in the ontology on an existing set of metadata like Dublin Core.

When defining a new ontology, the first step is to define its domain and scope (Noy & McGuinness, 2001). The domain analyzed in this chapter is that of a learning management system that deals with different kinds of learning resources and concepts like teachers, students, activities, etc. The developed ontology, SEOF (Simple Educational Ontology Framework), represents a first iteration in the inherently iterative development process and concentrates on these base concepts, leaving room for future expansion and refinement. Instead of using specialized ontology development tools that tend to be complex, cumbersome, and include more implementation-oriented features, in case of SEOF, a more general, abstract, and simple approach was used by utilizing UML (Unified Modeling Language) class diagrams. UML diagrams possess concepts – classes, attributes, and relations, that roughly correspond to those in specialized ontology languages.

Most of the SEOF concepts and their relations are shown in Figure 1.

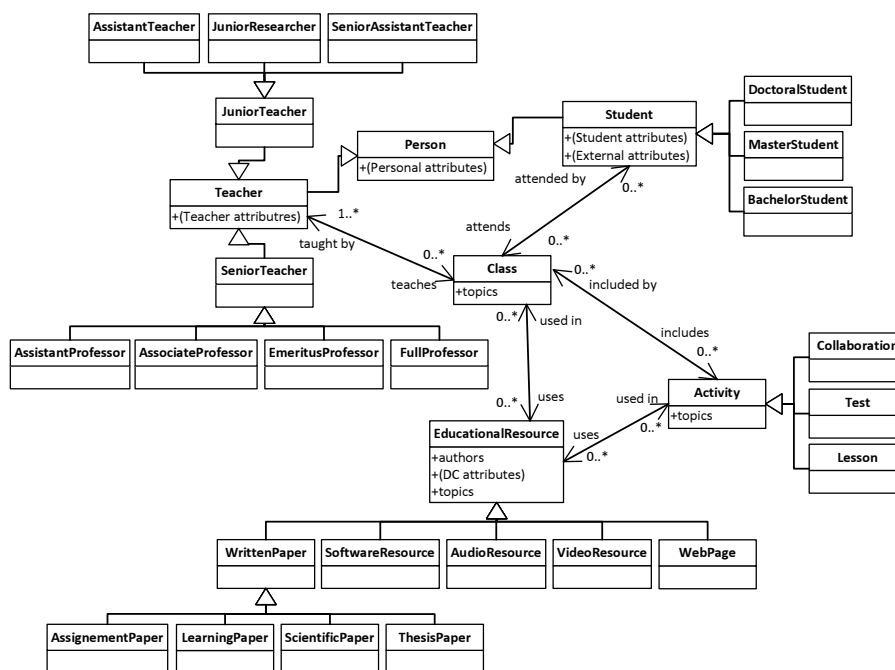


Figure 1: Simple Educational Ontology Framework

Educational Resource is a basic, expandable generalization that represents any type of resource used for learning. In order to ensure some level of interoperability with different systems, Dublin Core set of metadata attributes was added to this class. We suggest that specific subclasses can add other, more specific attributes, but should also define a scheme for translation and aggregation of these attributes into Dublin Core. Even if resources are exported to an external, not directly compatible environment, some of the semantics will still be preserved in the form of digests stored in the Dublin Core attributes. This is a suitable task for Dublin Core since it was designed as an expandable, lowest common denominator for the description of resources (Weibel, 2005b).

Other concepts include classes, activities, teachers and students. Classes can include different activities, while both of them can use different types of educational resources. Some examples of activities are tests, lessons, and collaborations; this set can be expanded according to the functionalities of the learning management system used. Teacher and Student classes inherit attributes from the Person class which defines adequate personal attributes. Students possess additional attributes like grades, attended classes, etc. and any attributes gathered from external sources, the main being a list of all their education-related interests.

An important class that is not shown in Figure 1 is the Topic class. The instances of this class represent concrete educational topics that can be used as allowed values for Dublin Core attribute "Subject" present in Educational Resource class. This concept is in some way related to all of the other classes. Classes and activities can list topics they cover, teachers can have a list of topics they have an expertise in, and students can have a manually inputted or automatically generated list of topics they are interested in. Topics can be related to other topics in many ways - they can be synonyms, antonyms, similar concepts, specializations, generalizations, implementations and so on; a suitable dictionary of relations between topics should be defined, and expanded per need, to properly capture all of the relevant relations that can be useful when browsing the resources in the system.

With all educational resources annotated and other concepts described and stored in some kind of a repository, the learners can perform advanced search and browse operations. If learners want to familiarize with the structure of a domain of study, a visualization of topics and related resources, activities, and classes can be presented to them. Basic searching can be performed on the Dublin Core attributes, while advanced queries can define required values for attributes and relations. For instance, the learner can request all learning papers that cover topics that are a specialization of those topics covered in a specific class.

3.2 Model of semantic services for e-learning

Following the design and serialization of an ontology, instances of defined classes need to be created and stored in some form of a repository. External entities can access, read, and change the repository through the core logic component, shown in Figure 2.

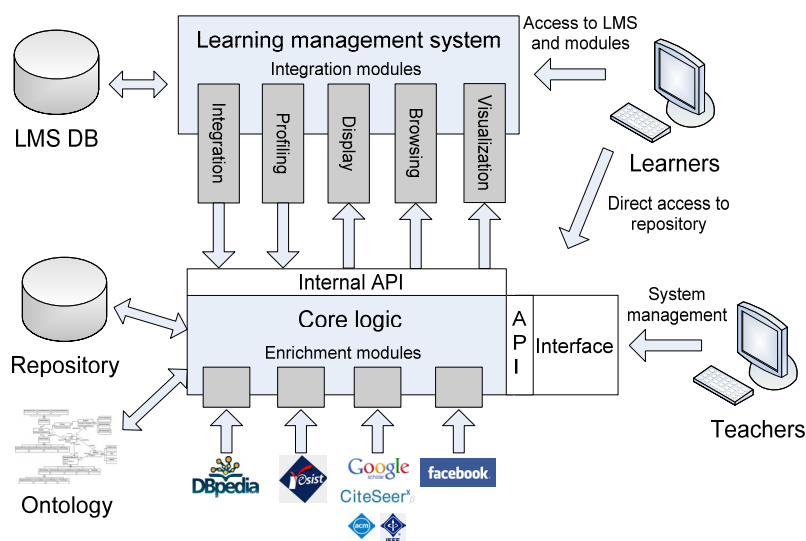


Figure 2: Model of integration of ontologies and semantic services with an existing learning management system

The instances stored in the repository represent metadata about resources and other concepts that exist in a LMS. The actual concepts are generally stored in some kind of a LMS-governed database, and direct access to them may or may not be available. In order to access this information, a modular architecture is proposed. The core should remain independent of any existing LMS and integration should be performed using modules, plugins, or other software entities supported by most such systems. These modules can then use the features and functionalities of the appropriate LMS, lowering the risk of any inconsistencies that may arise from direct, unsanctioned access. The modules suggested in this model are the integration, profiling, display, browsing, and visualization module. These modules are decoupled from the core and should communicate with it through a well-defined API.

The integration module integrates resources and their metadata. It keeps track of any changes made to resources, updating the repository as needed. The profiling module performs a similar task, gathering of information, but specifically targets students and their activities. The display module integrates into the existing interface components of the LMS, if possible, and presents the relevant metadata at all appropriate locations. The browsing module uses the information from the repository to offer an advanced searching and browsing functionality of both internal and external resources directly within the LMS. Finally, the visualization module enables visualizations of the structure of the domain and of instances and their relations.

In order to allow semantic enrichment of internal resources, annotation of external resources, and acquiring personal information about students, the core system should also be capable to integrate with the existing semantic systems, services, and social networks. Different services might possess different interfaces and methods of data access, and modular approach is again appropriate. Services such as DBpedia can be used for semantic enrichment by retrieving relations between concepts that do not exist within the SEOF. Social networks, especially those that are professionally oriented, can provide an invaluable source of personal information about students and their interests.

Finally, the core system should provide an interface for system management by teachers or administrators, allowing them to add, remove, or change metadata entries, configure the modules that integrate with external systems and change other system-wide settings. If the LMS is incapable of supporting some or all of the described modules, the core system also has to allow direct access to students through a separate interface.

4. Implemented components and implementation remarks Model for semantic annotation and semantic services in e-learning

The main advantage of a modular architecture, with sufficiently decoupled components and use of standard interfaces, is the possibility of independent development and modification. This provides a clear separation of concerns between components that gather data, integrate and store it, and generate its representations. By using common software routines and data for the missing parts of the model, several independent components were developed, mostly by adapting earlier research efforts conducted at the Laboratory for electronic business, University of Belgrade. The description of these components and some additional implementation remarks are given here.

The annotation of resources in the LMS and externally can be performed automatically, semi-automatically and manually. A semi-automatic method is recommended, where the first pass is automatic, and the teachers can then perform the second pass gradually, by manually editing and improving the generated metadata. A simple module for automated annotation of papers that contain a “keywords” section was developed and used to generate a test dataset for other developed modules.

The visualizations were developed using the freely available D3 library (Bostock, Ogievetsky, & Heer, 2011). The first visualization (Figure 3) uses some of the specific metadata fields to display relations between the topics. The student can select a starting topic either through a resource tagged with that topic, or by searching through available topics. That topic (PHP in Figure 3) is shown along with its relations to other topics, and their relations to other topics. By clicking on a topic, all available related resources are displayed, along with the option to shift focus to that topic.

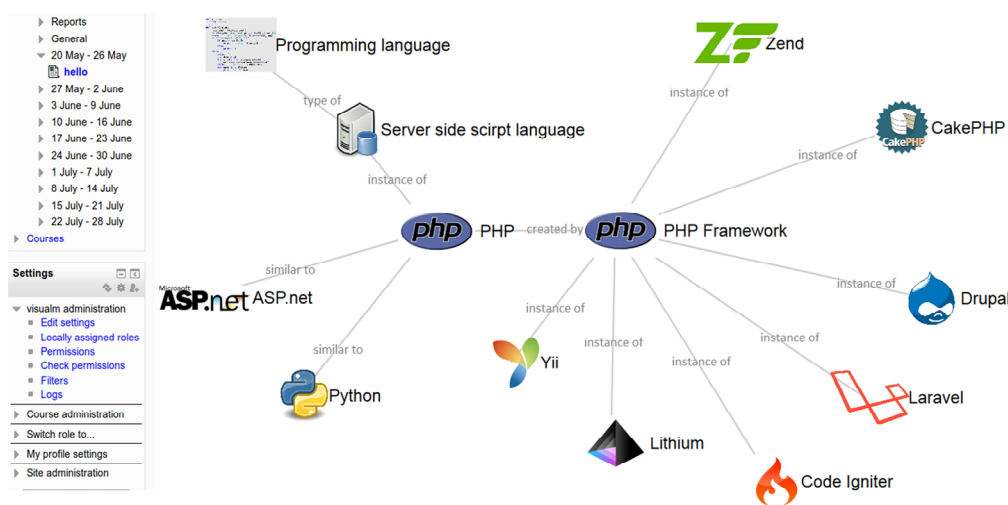


Figure 3: Visualization of topics described in the repository

The second module was developed with the goal of displaying retrieved semantic data from DBPedia. The resulting visualization is similar to the previous one, and is shown in Figure 4. All relations available from DBPedia for one, selected topic, are shown. Similarly, the student can select any topic to get additional info or to move it into focus.

While developing this visualizations, the problem that became evident was that of complexity; even the information shown in Figure 3 would become too complicated in a real system, since many potential relations were omitted from the test dataset. Adding another step away from the central concept would likely make the entire picture incomprehensible. Accordingly, future research should explore the introduction of weighted relationships combined with configurable complexity. Some method of integrating the two visualizations should also be explored, taking into consideration additional visual complexity when both internal and external sources are shown in the same picture.

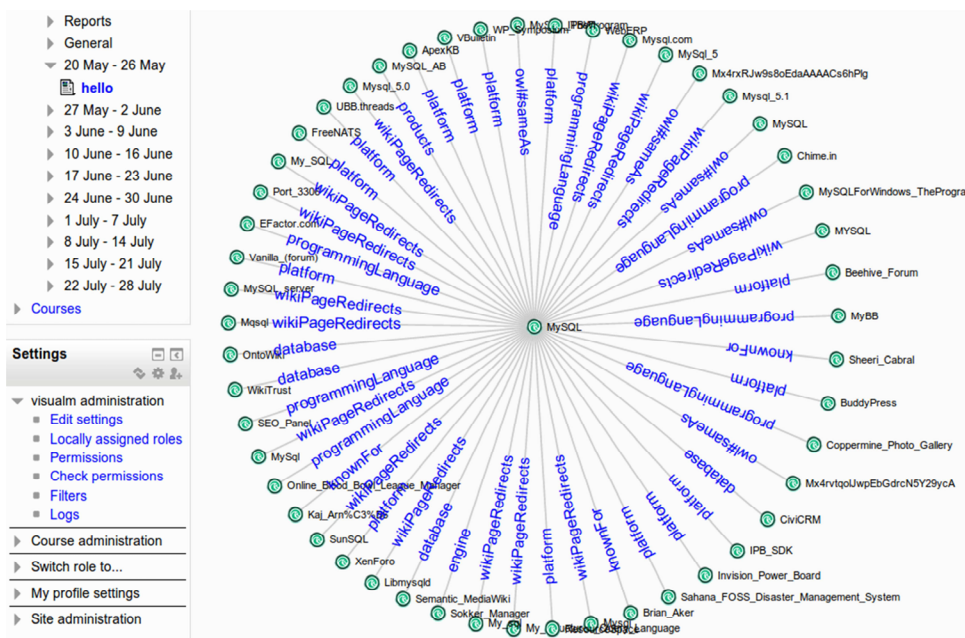


Figure 4: Visualization of related topics retrieved from the DBpedia

In order to gather information about students' interests, two types of sources are available. The first is a compromise of internal, LMS-provided information and heavily depends on the capabilities of the LMS itself. The other sources are external - social networks, and can provide a wider picture of the students' interests. A potentially interesting method of gathering student data is through an edutainment application that would attract learners more easily. This approach was tested at the Laboratory for electronic business by adjusting an existing Facebook edutainment application to gather specific information about its users and deliver it to the interface of the core application. However, gathering students' personal data in order to gain insight into their professional interests can be a sensitive topic and should be optional. It is perhaps preferable to make the gathered information unavailable directly to the administrators/teachers, but only available to the application's internal algorithms. Integration with different social networks can also be a challenge since it might overlap with marketing and educational efforts of the institution.

Conclusion

Semantic web represents a synergy of technologies and any application of semantic web is also likely to be technologically diverse in nature. However, this diversity enables a reuse of a single information resource in multiple different ways, enriching the experience of users. This paper presented a model for the application of core semantic web principles in e-learning and organization of semantic services with the goal of simplifying and enhancing the entire process of e-learning. All envisioned services were based on an ontology of concepts developed to describe the domain of e-learning through a learning management system. The model allowed for both simpler and more advanced searching and browsing of educational resources, as well as application of visualization libraries to represent the structure and content of the domain, with the possibility of adaptation to interests of the students. One of the problems of the suggested method of semi-automatic annotation is the potentially large required investment of effort; future research will explore the best ways to include students in this activity.

The future work will be focused on the implementation of the presented model and exploring the possibilities for integration with different access points of the semantic web and social media sites. The modular architecture will also allow experimentation on different modules and different modalities of integration with learning management systems.

Acknowledgement

The authors are thankful to the Ministry of Education and Science for financial support grant number 174031.

REFERENCES

- [1] Advanced Distributed Learning. (2009a). SCORM 2004 Content Aggregation Model [CAM] (4th ed., pp. 1–234).
- [2] Advanced Distributed Learning. (2009b). SCORM 2004 Run-Time Environment [RTE] (4th ed., pp. 1–202).
- [3] Bohl, O., Scheuhase, J., Sengler, R., & Winand, U. The sharable content object reference model (SCORM) - a critical review. , 1 *International Conference on Computers in Education 2002 Proceedings 950–951 (2002)*. *IEEE Comput. Soc.* doi:10.1109/CIE.2002.1186122
- [4] Bostock, M., Ogievetsky, V., & Heer, J. (2011). D3: Data-Driven Documents. *IEEE transactions on visualization and computer graphics*, 17(12), 2301–9. doi:10.1109/TVCG.2011.185
- [5] Cisco Systems. (1999, June). Cisco Systems Reusable Information Object Strategy: *Definition, creation overview, and guidelines*. San Jose, CA: Cisco Systems, Inc. Retrieved from http://www.cisco.com/warp/public/779/ibs/solutions/learning/whitepapers/el_cisco_rio.pdf
- [6] Dadzie, A.-S., & Rowe, M. (2011). Approaches to Visualising Linked Data: a Survey. *Semantic Web*, 2(2), 89–124. doi:10.3233/SW-2011-0037
- [7] Downes, S. (2001). Learning Objects Resources For Distance Education Worldwide. *The International Review of Research in Open and Distance Learning*, 2(1), 1–35. Retrieved from <http://www.irrodl.org/index.php/irrodl/article/view/32>
- [8] Dublin Core Metadata Initiative. (2012). *Dublin Core Metadata Element Set, Version 1.1*. Retrieved June 22, 2013, from <http://dublincore.org/documents/2012/06/14/dces/>
- [9] Garshol, L. M. (2004). Metadata? Thesauri? Taxonomies? Topic Maps! Making Sense of it all. *Journal of Information Science*, 30(4), 378–391. doi:10.1177/0165551504045856
- [10] Gerber, A. J., Barnard, A., & Merwe, A. J. Van Der. (2006). A Semantic Web Status Model. *Integrated Design and Process Technology*, Special issue: IDPT2006, 1–10.
- [11] Gómez-Pérez, A., & Corcho, O. (2002). Ontology Languages for the Semantic Web. *IEEE Intelligent Systems*, 17(1), 54–60.
- [12] Gruber, T. R. (1993). A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*, 5(2), 199–220.
- [13] IEEE Learning Technology Standards Committee. (2002, July). Draft standard for learning object metadata. *Institute of Electrical and Electronics Engineers, Inc.* Retrieved from http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf
- [14] Labus, A., Simić, K., Barać, D., Despotović-Zrakić, M., & Radenković, B. (2012). Integration of Social Network Services in E-Education Process. *Metalurgia International*, 17(7), 161–169.
- [15] McGreal, R. (2004). Learning objects: A practical definition. *The International Journal of Instruction Technology & Distance Learning*, 1(9), 21–32. Retrieved from http://www.itdl.org/journal/sep_04/article02.htm
- [16] Meeker, M., & Wu, L. (2013). 2013 Internet Trends. Retrieved from <http://www.kpcb.com/insights/2013-internet-trends>
- [17] Munzner, T. (2009). Visualization. In *Fundamentals of Computer Graphics* (pp. 675–720).
- [18] Shneiderman, B. (1996). The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *IEEE Symposium on Visual Languages* (pp. 336–343).
- [19] Weibel, S. (2005a). The Dublin Core: A Simple Content Description Model for Electronic Resources. *Bulletin of the American Society for Information Science and Technology*, 24(1), 9–11. doi:10.1002/bult.70
- [20] Weibel, S. (2005b). The State of the Dublin Core Metadata Initiative: April 1999. *Bulletin of the American Society for Information Science and Technology*, 25(5), 18–22. doi:10.1002/bult.127

Received: September 2013.

Accepted: December 2013.

About the Author

Zorica Bogdanović

University of Belgrade, Faculty of Organizational Sciences
zorica@elab.rs



Zorica Bogdanović received her PhD degree with the thesis titled "Business intelligence in adaptive e-education" in 2011, at the Faculty of Organizational Sciences. As an assistant professor at the Faculty of Organizational Sciences she is involved in teaching courses covering the area of E-business, Mobile business, Simulation and simulation languages, Internet technologies, Internet of things, Internet marketing, E-education, E-government, and E-health. Zorica Bogdanović is a secretary of the IEEE Computer Society CO-16 chapter.

Marijana Despotović-Zrakić

University of Belgrade, Faculty of Organizational Sciences
maja@elab.rs



Marijana Despotović-Zrakić received her PhD degree with the thesis titled "Design of Methods for Postgraduate E-education Based on Internet Technologies" in 2006 at the Faculty of Organizational Sciences. She is currently an associate professor, teaching courses in E-business, Internet technologies, Cloud computing, Mobile business, Simulation and simulation languages, Virtual reality, Internet of things, Internet marketing and E-education. Marijana Despotović-Zrakić is a vice chair of the IEEE Computer Society CO-16 chapter.

Miloš Milutinović

University of Belgrade, Faculty of Organizational Sciences
milosm@elab.rs



Miloš Milutinović is currently a PhD student at the Laboratory for Electronic Business, Faculty of Organizational Sciences. He received his master studies degree with the thesis titled „Developing a Mobile Application for Learning the Japanese Language Based on Learning Objects” in 2012 and is currently involved in teaching courses covering the area of E-business, Internet technologies, Internet marketing, Internet of things, and M-business.

Mihajlo Anđelić

University of Belgrade, Faculty of Organizational Sciences
mihajlo88@gmail.com



Mihajlo Anđelić is a final year B.Sc. student at the Faculty of Organizational Sciences and is currently employed as a Quality Assurance Engineer on a project concerning the development of the next generation of medical software. His current professional interests include Internet technologies, visualization, e-business, e-learning, e-government, m-health, programming paradigms and embedded electronics.

Suzana Milinović

Falc East d.o.o,
suzi.milinovic@gmail.com



Suzana Milinović received her M.Sc. degree in 2013, with the thesis titled "Implementation of an Internet marketing plan for a shoe factory", at the Faculty of Organizational Sciences. She is currently working as a general manager for "Falc East", a factory manufacturing Italian footwear. She is interested in exploring the web as a marketing platform.