International Journal of Scientific and Technological Research
ISSN 2422-8702 (Online)
Vol 4, No.10, 2018

An Analysis of Educational Linked Data

Yasemin Gultepe
Department of Computer Engineering,
Kastamonu University, Kastamonu/Turkey
E-mail: yasemingultepe@kastamou.edu.tr

Abstract
In the last few years, the scale and variety of data sets published in accordance with the Linked Data (LD) principles have increased and at the same time have led to the emergence of a broad range of educational conformity. However, sufficient information about the status, scope and scope of the current educational Linked Data still seems to be incomplete. In this study, the scope and area of higher educational linked data on the web are analyzed. This study provides an overview of the creation of data on higher education related to data sets and vocabulary, data collection and processing, Resource Description Framework (RDF) and link building. By using Linked Data Principles, Linked Data can be queried from multiple sources at once, and all data shares can be combined without the need for a single common schema. This study sheds light on the basic requirements of linked data applications at national and organization levels in higher education environments.

Keywords: Linked data, semantic web, education, educational data.

1. Introduction

The developments in information and communication technologies have led to the change in data searching and access behaviors of individuals in the structure of data access systems. The possibilities offered by web technologies made it possible to access data without time and space limitations. Access to quality data generated by higher education systems through central information systems brings many problems with it. These problems can be listed as follows: a) failure to integrate these data with web resources, b) difficulties in determining the relationships between data exchange between other institutions, structures and information systems other than the higher education institutions themselves, c) difficulties in making management operations, such as updating existing standards, principles, rules and approaches to information provision and access.

Semantic web technologies can provide potentially powerful tools for developing higher education teaching in complex, fast-working subject areas (Carmichael and Jordan, 2012). Along with the semantic web, web content has to be associated with each other and can be expressed by sentences, the web environment has been transformed into a large database, the machines can be asked questions, and machines can find answers to each other by answering questions and create a user-centered distributed structure. In this case, universities must establish a structure based on semantic web technologies to manage data. This data needs to be collected, associated and shared on the web in order to get better queries and results by computers and individuals. In this context, the information transferred to the semantic network environment produced by universities has been adapted and produced for this environment.

Linked data, which is considered to be the realization of the semantic web today, is based on the fact that the data is published on the web in the form of data sets compatible with the RDF standard and associated with the RDF connections of different data sets (Berners-Lee, 2006). Linked data technology for the development of innovative applications in the field of education involves linking data from different sources based on open standards and using open data schemes (Keßler et al., 2009). The data sets are mostly based on data from the open university website. The open university linked data platform is one of the most known and large data sets on the linked data cloud. Data sets related the course information, scientific publications, audio and video format teaching materials implemented at the Open University. Linked data enables a publication distribution systems (postcasts), open education resources, social networking site videos, university organization structure, libraries to connect their collections with

resources out on the web.
Linked data is a link path between learning-teaching materials, sharing and re-use, and learning-teaching materials of participants/students (Stefan et al., 2012). Learning-teaching materials are linked with lectures, multimedia resources, scientific studies and research resources.
The aim of this study is to define and conceptualize the institutions and procedures in an academic institution, to form the basic concepts of curriculum for the teaching and learning concepts in a higher education environment. The developed educational ontology aims to be highly transferable and reusable to other universities.

2. Context and Related Works

In this section, basic information about semantic web, linked data and Linked Open University Data (LOUD) will be discussed. The topic of how to access existing ontologies will be covered in the title of ontologies. The use of the linked data and LOUD will form the contents of the other sections.

2.1 Semantic Web

The semantic web concept is put forward by Tim Berners-Lee, who designs and finds structures such as Uniform Resource Identifier (URI), Hyper Text Transfer Protocol (HTTP) and Hyper Text Markup Language (HTML), which form the basis of today's web, and is considered to be the web's future step for the full potential use of the existing web environment (Berners-Lee et al., 2001). The semantic web is not a new and separate web, but an extension of the existing web where people and people can work in collaboration, with well-defined meanings to the information. The main purpose of the semantic web is to develop standards and technologies that can easily read and link the well-defined and linked information and services on the machines and make them understandable by the machines.
Ontology offers the ability to define the characteristics of an e-learning system for higher learning through its components (Synak et al., 2009). It contains all the necessary concepts and requirements that form the basis for a web-based e-learning system.

2.2 Linked Data

The semantic web offers a concept of Linked Data developed by World Wide Web Consortium (W3C) to enable it to transform into an intelligent database on a global scale. Each information is modeled for a particular meaning with the linked data. This information is linked to each other. It creates a three-dimensional navigation.
The most concrete application of the linked data is the Linking Open Data Project (Berners-Lee et al., 2001). In this project, the data published in various formats on the internet is aimed at connecting each other by adapting to the five star rule (Berners-Lee, 2006; Heath and Bizer, 2011) proposed by Tim Berners Lee. The first star is to make information available on the web under an open licence. All subsequent stars aim to make it easier to discover, use and understand data. The second star is assigned to be available as machine-readable structure data. The third star is given to data that users do not require a special software package to analyze. The fourth star goes to data that uses W3C standards, such as RDF and Query Language for RDF (SPARQL). Finally, the fifth star is to connect your data to other datasets. The data cloud in this project is growing day by day. Even in 2018, the dataset currently contains 1,229 datasets with 16,125 links (McCrae, 2018).
The four design principles of the linked data (Bizer et al., 2009): 1) URI is used in the resource definition. An RDF statement is a data source defined by a URI. 2) URI is used to define things. The data source can be accessed via HTTP. HTTP URIs are used for searching these things by people and machines. Under the linked data, resources are usually defined using HTTP URIs and the use of other URI schemes such as URNs, DOIs is avoided. HTTP URIs offer a globally unique name without the central government to provide access to information about resources on the web. 3) Inference procedures are carried out by using inquiry standards on RDF data sources. In this way, semantic relations are determined between the data in the web environment. 4) It includes links to other URIs so that more resources can be discovered. Educational applications that can use the semantically related data on the linked data cloud within the framework of their requirements will also play an important role in the use of the semantic web in the field of education. Linked Data enables data clusters to be combined in a data scale to overcome data sustainability issues of many educational applications, including the research the efficiency of the newest Information and Communication Technology (ICT) tools in education.
In 2012, Ruiz-Calleja et al. (2012), a Linked Data-based approach is proposed, which allows the integration of training ICT tool registries with the Web of Data. Thus, it was possible to share the

descriptions of ICT tools between third parties and to alleviate existing data sustainability problems (Ruiz-Calleja et al., 2012).
In this sense, Talis Aspire can be used as the Linked Data application, which helps to create and manage lists of educational resources. The application is written in PHP, which is supported by the Talis Platform to store, manage and access Linked Data, and is used daily by tens of thousands of students in countless universities on a daily basis (Clarke, 2009).
Another important project, ASSESS, is used as a resource for creating automated training materials for Linked Data. A bridge has been established between the natural language and the RDF by using the innovative RDF and entity summarization technology. RDF data was used to generate exams covering different types of questions on user-defined domain names (Bühmann et al., 2015).

2.3 Linked Open University Data
Linked Open Data (LOD) is an important mechanism for information management and integration because it facilitates innovation and information creation from interlinked data (Bauer and Kaltenböck, 2012). Linked Open University Data can be applied to the university information domain using semantic web and linked data technology. Thus, The LOD aims to create interlinked semantic data on the basis of university information. In addition, the unied inner and inter school provides an opportunity for inquiry and comparison.
On the Open University's Open Linked Data platform, data from various institutional archives of the university are linked and made available for reuse (Ma et al., 2011). The data can be queried on the SPARQL endpoint on this platform. The data sets include publications, qualifications, lessons and audio / video materials that are currently produced at the Open University. All of these data can be used via standard formats (RDF and SPARQL). The data sets can be listed as follows (Enrico et al., 2015): Open Educational Resources, Scientific Production, Social Media, Organizational, Data from Research Projects, Metadata and Document. *Open Educational Resources* includes metadata data about educational resources produced or jointly produced by the Open University. *Scientific Production* is the metadata for the scientific production of the Open University. *Social Media* is content that is hosted by social media websites. *Organizational data* is collected from internal repositories and is first made public as linked data. *Data from research projects* are data from research projects. *The metadata and documents* are the data fields that are separated into the schema description and documentation.
In 2017, Fleiner et al. (2017), a new ontology called an Ontology for Linked Open University data (OLOUD), which supports the development and publication of Linked Open University data sets and the applications placed on these defined data sets (Fleiner et al., 2017). The field of OLOUD ontology consists of open data to be published in the university. OLOUD offers a high-level model covering multiple training-related usage cases that catalyze linked data production and consumption in the field. OLOUD includes classes and features to define Organizations, People, Roles and Publications, Subjects Courses and other Events along with temporal and spatial description.
Linked universities (LinkedUniversities.org) is a practical community working in cooperation on developing standards. As a collaborator, it aims to create an agreed upon vocabulary. In general, an initiative of the European universities trying to present their public data as linked data (Linked Universites Hub, 2018).
LinkedUp aims to exploit the large number of publicly available data available on the Web, especially by educational institutions and institutions. The vision of the LinkedUp was to perform personalized education of global impact based on open web data and knowledge (Dietze et al., 2013).

3. Material and Methods
All semantic web users may not be at the level of information that will create an ontology from the beginning. There are other disadvantages to this method, even if all users are assumed to be at this level. One of them is a huge loss of time and labor because each user has created his or her own ontology from the beginning. Another drawback is that information sharing between users with different ontologies becomes incapacitated due to the lack of common concepts between ontologies.
The power of semantic web ontologies lies in sharing, as the number of users sharing the same terms and ontologies increases, the benefits of ontologies will also increase (Hendler, 2001). Semantic web ontologies will not be major ontologies, i.e. all the concepts a user needs will not be defined in a single ontology. Instead, a unique ontology will be created by combining concepts from ontologies of many different fields. It is also currently a challenging activity to identify a set of data that contains resources on a given topic. Moreover, the lack of up-to-date and precise descriptive information has further

increased this difficulty. The combination of concepts means that the conceptual value of each concept is taken from the ontology to which it belongs when it is needed.

3.1. Datasets

In this study, a linked dataset has created for defining the Computer Engineering department curriculum. Different data sets and ontologies have been used to provide information on the education system to be shared and interoperable within the scope of the Bologna Reform. This study covers curricula, subjects, courses, semesters and educators. More than one ontology is needed for these title definitions. This study was based on FOAF (Friend of a Friend), AIISO (Academic Institution Internal Structure Ontology), BIBO (Bibliographic Ontology), Bowlogna ontologies. These ontologies have been expanded by various additions within the scope of the developed scenario.

The FOAF ontology consists of classes such as person and name, surname and email address of those classes (Brickley and Miller, 2014). Although classes and properties can be defined with RDFS, complex relationships between objects cannot be modeled with RDFS. For this reason, FOAF is prepared with OWL which has more qualifications. In this study, FOAF has been used to specify the characteristics of university staff and their connection with other staff.

AIISO provides the classes and features used to define the internal organizational structure of academic organizations (Styles and Shabir, 2008). This ontology focuses on the organizational structure of a university. A course content can be modeled using AIISO ontology. In this study, http://purl.org/vocab/aiiso/schema#Module is associated with education:Teaching_Unit defined in the educational (created in the study) ontology.

BIBO is a vocabulary for describing citations and bibliographic references associated with the course, such as books or alumni (D'Arcus and Giasson, 2009). BIBO provides concepts and features for identifying citation and biographical references. Ontology was created using RDF language. Dublin Core Ontology (DCMI, 1998) is used to identify biographical sources. Dublin Core element set can be used as a metadata scheme for information exchange and integration between digital sources. For example, BIBO ontology was used to describe the book.

- *<ISBN: 1614996229> <rdf:type> <bibo:Book> .*
- *<ISBN: 1614996229> <dc:title> "Publishing and Consuming Linked Data".*
- *<ISBN: 1614996229> <dc:creator> "L. Rietveld".*
- *<ISBN: 1614996229> <bibo:isbn10> "1614996229".*
- *<ISBN: 1614996229> <bibo:isbn13> "978-1614996224".*

Bowlogna Ontology is used to model an academic environment proposed by bologna reform (Demartini et al. 2013). This ontology develops standardization to improve the linkage between academic systems and to improve training workflows, communication and collaboration between universities.

3.2. Ontology Development Tool

Ontologies are combined using the Protégé ontology editor (Protégé, 2018). Ontologies are defined visually by the Protégé ontology development editor's graphical interface. The desired area can be modeled in this way. It also provides great convenience in ontology development, facilitating the development of ontologies and reduces the possibility of errors. Protégé can work with a wide range of web tools and technologies developed for ontology development, management, evaluation, inference, etc.

3.3. Methods

The process for this study consists of three stages. In the first stage the theoretical constructs of this study were developed using literature. In the second stage the integration of data from multiple, distributed and heterogeneous (different) data sources. Gathered data often comes from heterogeneous sources; therefore, integration activities are needed and very important. This means the ability to integrate information and functionalities of different Information Technology (IT) systems. In the third stage, data are linked together to form a large data center network. That's the web of linked data. The basic concepts of university information network for this study are shown in Figure 1.

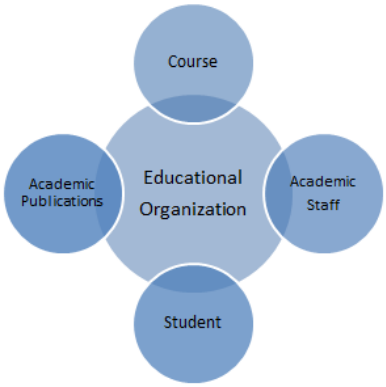

Figure 1. The basic concepts for university information network.

Course describes the courses that are given by lecturers and students must attend. A course planning is made based on the definition of the Bologna process of the organizations. Some features related to the course; *hasName*, *isTeacherOf*, *isAssistantOf*. EducationalOrganization defines an organization designed to provide educational services and learning environments. When creating an ontology, first of all, it is the determination of super classes and sub classes. The sub classes of EducationalOrganization are defined as *University*, *Institute*, *Faculty* and *Department*. Academic staff is the name given to individuals with academic title in all levels of a university. In this study, FOAF ontology was extended to publish personal information about academic staff. Depending on the academic levels, it can be divided into the following subcategories: *Professor*, *AssociateProfessor*, *Lecturer*, *ResearchAssistant*. In this study, the FOAF ontology structure has been extended to define academic profiles based on academic staff profile characteristics. There is an *TeachesIn* relation between the AcademicStaff class and the Course class. Student who is studying at a university or other place of higher education. There is a *takesCourse* relationship between the students and the classes. AcademicPublications is a kind of scientific writing. There are different subclasses: *Article*, *Book*, *Periodical*, *Proceedings* and *Thesis*.

Especially on the basis of Bowlogna ontology and other ontologies described in the DataSets (3.1) section supply the requirements of this study. The linked data, which expands by connecting with the meaningful information, becomes a navigated data set on it. Within the scope of this study, various additions to these field ontologies have been made. The extended ontology consists of more than 100 classes, 85 object properties and 79 datatype attributes. Property hierarchy of ontology; Object and datatype are displayed in two ways. Object properties that are not directly included in the concept list are defined by considering the relationships between classes. For example, there must be an inheritance relationship between classes, which is shown in ontology with the object property named relationship. Figure 2 shows the images of the classes of ontology.

First, educational ontology was determined by using an external data set. Then, the mappings between these nodes and external relational properties are defined. The mapping between external data sets and educational ontology includes mapping between nodes, data properties, and object properties. Semantic matching is an ontology-based kind of information search process. Semantic mapping returns semantically matched information with concepts defined in a particular ontology. The advantage of semantic mapping is that semantically related results can be returned to the user when an exact match cannot be found. With semantic matching, the user is provided access to the most semantic information in response to the user's request.

The aim of this study is to define and conceptualize the institutions and procedures in an academic institution, aiming to model the basic concepts of curriculum for the teaching and learning concepts in a higher education environment. The developed educational ontology aims to be transferred to other universities and be reusable.

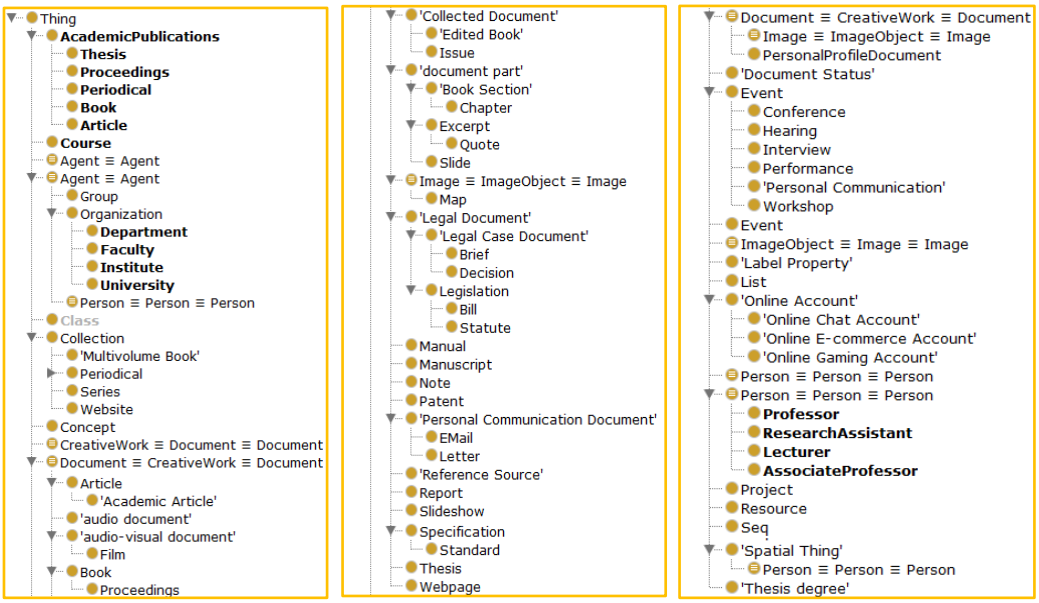

Figure 2. Extended ontology to represent the basic concepts for university information network.

5. Conclusion

In recent years, there has been a large increase in the number of projects, initiatives and academic studies conducted on data and academic linked data on the web. At the same time, important steps have been taken in the standards developed for the inquiry and search of semantic search. The main purpose of the publication of educational data as Linked Data is to ensure the reusability of data and interoperability between data sources.

Thanks to the linked data; 1) The meaning of the information can be uploaded. 2) A connection can be made between these meaningful information. 3) New information and results can be reached automatically thanks to these links established between the information.

An educational ontology was created in this study. Using the created educational ontology, mappings are defined between a subset and other linked data sets. Ontology mapping allows independent linking of ontologies and the interoperability of heterogeneous resources. Linked data is a new step in this direction. Ontology was expanded as a result of questioning on the open datasets attached to extract samples from these sources.

The importance of the study will emerge considering the requirements in the publication of the linked data and the elimination of the deficiencies at the institutional level. It sheds light on the basic requirements of linked data applications at national and institutional levels. It is a reference to be used as a base at the corporate level to publish linked data.

Acknowledgement

This research has been supported by Kastamonu University Scientific Research Projects Coordination Department. Project Number: KÜBAP-01/2013-65.

References

Bauer, F. and Kaltenböck, M. (2012). Linked Open Data: The Essentials. *Semantic Web Company*.

Berners-Lee, T., Hendler, J. and Lassila, O. (2001). The Semantic Web. A New Form of Web Content That Is Meaningful to Computers Will Unleash a Revolution of New Possibilities. *Scientific American*, 284, 1-5.

Berners-Lee, T. (2006). Linked Data – Design Issues, 2006.

Bizer, C., Heath, T. and Berners-Lee, T. (2009). Linked Data - The Story So Far. *International Journal on Semantic Web and Information Systems*, 5(3), 1-22.

Brickley, D. and Miller, L. (2014). FOAF (The Friend of a Friend) Vocabulary Specification 0.99, [Online] available: http://xmlns.com/foaf/spec/ (November 01, 2018).

Bühmann, L., Usbeck, R. and Ngomo, A-C.N. (2015). ASSESS-Automatic Self-Assessment Using Linked Data, ISWC 2015.

Carmichael, P. and Jordan, K. (2012). Semantic Web Technologies for Education-Time for a 'Turn to Practice'?. *Technology Pedagogy and Education*, 21(2), 152-169.

Clarke, C. (2009). A Resource List Management Tool for Undergraduate Students Based on Linked Open Data Principles. In Proceedings of the 6th European Semantic Web Conference, Heraklion, Greece, 2009.

D'Arcus, B. and Giasson, F. (2009). Bibliographic Ontology Specification, [Online] available: http://bibliontology.com/ (November 01, 2018).

DCMI. (1998). Dublin Core Metadata Element Set, Version 1.0: Reference Description.

Demartini, G., Enchev, I., Gapany, J. and Cudré-Mauroux, P. (2013). The Bowlogna Ontology: Fostering Open Curricula and Agile Knowledge Bases for Europe's Higher Education Landscape, *Journal Semantic Web*, 4(1), 53-63.

Dietze, S., Herder, E., Drachsler, H., d'Aquin, M. Wall, S. and Parodi, E. (2013). LinkedUp-Linking Web Data for Education Project, Project Networking Track at ESWC 2013.

Enrico, D., Mathieu, d., Alessandro, A. and Stuart, B. (2015). The Open University Linked Data-data.open.ac.uk. *Semantic Web Journal*, 7(2), 183-191.

Fleiner, R., Szász, B. and Micsik, A. (2017). OLOUD-An Ontology for Linked Open University Data, *Acta Polytechnica Hungarica,* 14(4), 63-82.

Heath, T. and Bizer, C. (2011). Linked Data-Evolving the Web into a Global Data Space. *Morgan and Claypool Publishers*, 1st edition.

Keßler, C., d'Aquin, M. and Dietze, S. (2009). Linked Data for Science and Education, IOS Press, 1-5.

Linked Universites Hub, [Online] available: http://linkeduniversities.org (November 01, 2018).

Ma, Y., Xu, B., Bai, Y. and Li, Z. (2011). Building Linked Open University Data-Tsinghua University Open Data as a showcase, JIST 2011: The Semantic Web, 385-393.

McCrae, J.P. (2018). The Linking Open Data Cloud Diagram, [Online] available: http://lodcloud.net (November 01, 2018).

Protégé. (2018). A Free, Open-source Ontology Editor and Framework for Building Intelligent Systems, [Online] available: https://protege.stanford.edu/ (November 01, 2018).

Ruiz-Calleja, A., Vega-Gorgojo, G., Asensio-Pérez, J.I., Bote-Lorenzo, M.L., Gómez-Sánchez, E. and Alario-Hoyos, C. (2012). A Linked Data approach for the discovery of educational ICT tools in the Web of Data, *Computers & Education*, 59, 952-962.

Stefan, D., Quing, Y.H., Daniela, G., Eleni, K., Nikolas, D. and Davide, T. (2012). Linked Data Education: Interlinking Educational Resources and the Web of Data. The 27th ACM Sysmposium on Applied Computing (SAC-2912).

Styles, R. and Shabir, N. (2008). Academic Institution Internal Structure Ontology, [Online] available: http://vocab.org/aiiso/ (November 01, 2018).

Synak, M., Dabrowski, M. and Kruk, S.R. (2009). Semantic Web and Ontologies, In: Kruk S.R., McDaniel B. (eds) *Semantic Digital Libraries*, Springer, Berlin, Heidelberg.