

THAI CUSTOM INFORMATION SHARING ON THE INTERNET BY LINKED DATA TECHNIQUES

Somphong Wathanti¹, Wirapong Chansanam^{*2}, Kulthida Tuamsuk³

¹Ph.D. Student in Information Studies Program, KhonKaen University, Thailand

^{2,3}Information Science Department, KhonKaen University, Thailand

E-mail: *2wirach@kku.ac.th

Received: 20.05.2020

Revised: 17.06.2020

Accepted: 06.07.2020

Abstract

This paper propose a set of preparation for Linked Data creation and publication the Thai custom information on the Internet based totally on ontology construction, collectively with one pilot mannequin example in the area of Thai custom in northeast of Thailand. By presenting distinctive descriptions of every task in the generation and e-book processes, these training advantage each private and public agencies that work with data about custom and lifestyle related different area in Thailand in generating Linked Data formal ready-existing information and in launching the generated information in accordance to the contemporary standards. This instance assistances the audience from one-of-a-kind businesses to reap better understanding into the procedures of Linked Data creation and publication, as a consequence confirming the easiest exceptional of the outputs of these practises.

Keywords--Thai Custom, Information Sharing, Ontology, Linked Data

© 2020 by Advance Scientific Research. This is an open-access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)
DOI: <http://dx.doi.org/10.31838/jcr.07.08.281>

INTRODUCTION

Thai usual twelve (Heet-Sibsong) is a way of life called in the northeast place of Thailand. Heet, or tradition, that people in the northeast have inherited, is one of the treasured and intangible cultural heritage of human beings in the northeast acknowledged as the custom twelve and lifestyle fourteen, that consists of twelve customs in twelve months, nicely regarded as the "Heet-Sibsong", which is a subculture that has historically been the guidelines for living below the coexistence of society are known as "Kong-Sibsee", divided into two types for elite and for the customary people. In the past, not solely the way of life, the significance of residing as a community practice, it is additionally a social norm that when practiced will be lead to peace. So that everybody in Thai society should respect and observe. Therefore, Heet-Sibsong and Kong-Sibsee are likes a law of society that everybody ought to abide because they have described the obligation of anybody in the society, with the belief that any society will comply with Heet-Sibsong and Kong-Sibsee will continue to be keep calm for all. The information administration gadget of the Thai custom ("Heet-Sibsong") by using the explanatory framework for internet site technology nonetheless lacks the management machine to be capable to hyperlink the information showing the Thai custom in a variety of dimensions. That can give an explanation for or illustrate the things that are the equal and exceptional in every lifestyle and location, along with the lack of a gadget for creating links facts to relevant sources of knowledge. Ontology is a way to assist organize knowledge systems and understanding retrieval to help with the understanding collection, understanding exchange, and information reuse (Noy and McGuinness, 2001). This paper provides the result of creating the records sharing on the Internet about the Thai customized with the aid of applying a Linked Data technique based on ontology building.

ILLUSTRATIONS

Linked Data is the procedure of following a set of satisfactory performs for publication and linking structured information on the Web (Bizer et al., 2009). It is the ultimate step in the 5-star deployment scheme for linking data. The word "linked data" for that reason mentions to statistics which is posted on the Web and, aside from being computing device readable, it is also

connected to other outside datasets. The multiplied rate of application of Linked Data excellent applies has lead the Web to evolve into an international statistics area comprising billions of assertions, where both documents and records are linked. The advancement of the Web allows the examination of new associations between records and the ensuing improvement of new applications. Chotirat et al. (2011) current a tenet for format improvement and ontology knowledge-based testing for automatic evaluation of information online. According to the Ontology Life Cycle, researchers have studied the factors associated to the analysis of online information and regarded key phrases that affect evaluation of content and developed an ontology knowledge-based. This used to be developed the usage of Hozo-Ontology Editor, which is a freeware program for infrastructure development and then evaluation was once conducted. Results of the comparison by means of specialists show that the ontology knowledge-based is atan appropriate degree (mean =3.61, S.D= 0.16). Therefore, this ontology understanding-based shows effectivity with this specific issue and may have to be applied for growing and computerized evaluation of information online appropriately for further research. Chansanam et al. (2014) presented the improvement of an ontologyon the knowledge of trust subculture in the Greater Mekong subregion (GMS). The ontology used to be carefully designed to specify the principles relevant to intangible and tangible cultural heritage and the members of the family among them. The expertise domain in this work focuses on the cultural context and implicit attributes of the GMS as an initial case study. To further illustrate the attainable of the developed ontology, a semantic search application was carried out and then evaluated by way of experts. On the contrast processes, various intricate queries have been designed in order to absolutely utilize the relations among ontological classes, and the outcomes have been returned accurately. The evaluation proved that the ontology was nicely described in aspects of its hierarchical shape and members of the family from intermediate thought layers. Butte and Puarungroj (2016) hadimproved the semantic web lookingfor cultural records in Loei province, two research objectives aimed to 1) enhance an ontology-based on cultural information in Loei province; 2) boost a semantic looking out device for utilising the developed ontology. The system used to

be supposed to solve the problem that presented daily cultural information which cannot be obtained from a single source and the statistics obtained cannot meet the requirements. The research used the ontology as a principal thing in the form of OWL (Web Ontology Language). The looking out system was once carried out via the use of SPARQL. The outcomes show that by way of integrating the developed ontology with Lunar Calendar, the ontology ought to higher describe which means of subculture and could furnish users with the required information. The device was once evaluated primarily based on consumer pleasure which suggests that the delight regarding the semantic internet and usefulness is exceptionally at a very high level. The machine was once evaluated to measure its overall performance by using examining the system's and F-measure, precision, and recall. The end result shows that the device performed well with the precision = 0.77, recall = 0.83 and F-measure = 0.80. Chaikhambung and Tuamsuk (2016) developed a semantic ontology of information on ethnic groups in Thailand via examining content material from relevant statistics resources the use of classification theory. Research techniques comprised of 3 steps: 1) identifying the reason and scope of ontology, 2) developing of ontology the usage of the Hozo Ontology Editor, and 3) comparing of the ontology by using interviewing with the experts. The research has resulted in a semantic ontology of expertise area of ethnic organizations in Thailand which consists of 137 classes, divided into 16 primary instructions including Ethnic groups, Self-call name, History, Language, Religion, Beliefs, Customs and rituals, Dress, Arts, Performing arts and plays, Distinctive habits, Habitat, Subsistence, Social organization, Economic system, and Adaptation. The Ethnic group class is a classification that mates all predominant instructions to show the expertise content material in every component of every ethnic group. The ontology assessment determined that the standard ontology development system was gorgeous in 5 processes: Definition, scope, and reason of development (score = 0.89), Definition of concepts/classes (score = 0.78), Definition of properties (score =0.70), Definition of instances (score = 0.84), and software and the ontology improvement approach in the future (score = 1.00). The key idea in the back of the Linked Data paradigm is to observe this very standard structural design, no longer to internet documents, but to real data. Linked Data is typically summarised by way of the so-called Linked Data principles (Berners-Lee, 2019): 1) use URIs as names for things, 2) let the URIs resolve to pages that report or support the interpretation of that thing, 3) furnish an RDF description of the identical aspect at the equal URI, 4) consist of links to different associated data. Linked Data consequently described represents a conservative extension of the document Web with two new features: URIs are used to refer no longer only to Web documents however to things more commonly (principle 1), and URIs unravel to distinctive representations of a thing relying on who is asking; a human reader is to be directed to a conservative Web

document with human-readable content, whereas an electronic mediator is to be offered a machine-readable description of the same element in the structure of an RDF document.

METHODOLOGY

The information-sharing mission has the objective of making hyperlinks in the data of RDF (Ferrara, Nikolov, & Scharffe, 2011& Hussain et al., 2020) and it can be completed in numerous sequential phases with the aid of the use of the RDF dataset and the ontology as inputs. The first step is to perceive instructions whose cases can be the concern of linking, while the second step is to become aware of records sets that might also incorporate instances for the in the past identified classes. The third step is to choose the tools for execution the process. The distinctive tools for linking data occurred, and personally tool has its benefits and be responsible for multiple functionalities for assured matching tasks. Nevertheless, in this study, the linking can be applied by using hand (e.g., when the generated dataset is small, or when the quantity of instances to hyperlink is now not a whole lot more), and the next step is now not always implemented. The fourth step is using that tool in order to acquire links. The one of a kind tools are used in a different way and each tool requires conformation from the person in a particular form. Tools that can be used for data linking put in force include OpenRefine, Jena and Fuseki (Apache Jena, 2019), and Jetty and Pubby (Cyganiak, & Bizer, 2008) with the reconciliation service of the RDF extension. One liked aid to take into account when classifying correlated datasets for linking is the Thai custom dataset list, which incorporates Linked Open Data sets associated to Thai custom in the northeast of Thailand. The training whose situations can be topics of linking have been recognized first. In this study, there were two developed Thai custom ontology for vocabulary, the identified training include TCO:Custom:Custom_name. DBpedia (Jr, 2009) and WikiData (Vrandečić, & Krötzsch, 2014) were a database encompassing the geared up content material from the Wikipedia information in RDF format; it was recognized as the dataset that might include associated cases of the formerly noted classes. Since in the TCO RDF dataset, there was a huge wide variety of cases that can be linked, and due to the fact OpenRefine has been used for converting the data into RDF, we have carefully chosen OpenRefine for accomplishment the connecting task. The OpenRefine RDF extension can accomplish the search venture and locate the hyperlinks which can be represented in an remoted column in OpenRefine and mapped with the correlated RDF instances using the prefixed property such as owl:sameAs. When such mapping is created, the hyperlinks perform in the RDF information exported by OpenRefine. In total, there are 23 distinct instances in the TCO RDF dataset for which OpenRefine counseled hyperlinks to instances in DBpedia and WikiData. After accomplishment the connecting process, 23of cautioned instances have been connected to DBpedia and WikiData.

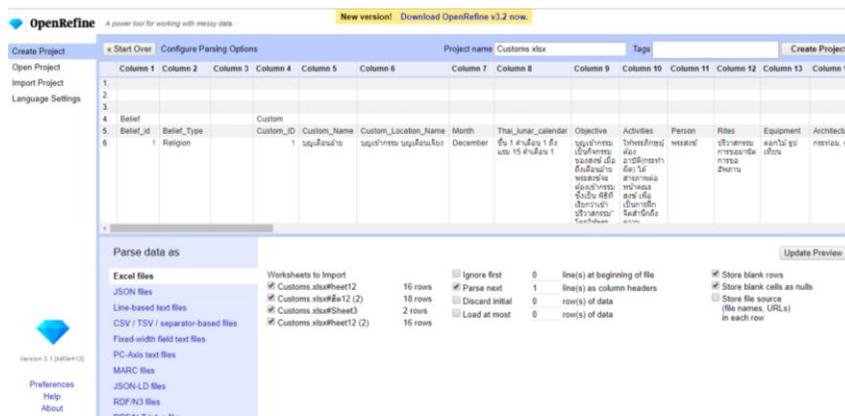


Figure 1. Reconciliation linking data by OpenRefine and RDF extension

For the semantic illustration of Thai custom knowledge we pick out up the recognized thinking from the semantic looking out community of the definition of OWL (Hitzler et al., 2019) or RDFS (W3C, 2019) vocabularies that describe positive elements of the tradition domain. Such factors usually incorporate structural components (e.g. custom, tradition, activities).

Matching ontologies often provide annotation vocabulary that permits the definition of similarly metadata. In the following we first describe suitable vocabularies for the representation of structural northeast regular aspects of a Thai custom.

A linking data development upon these vocabularies we introduce a novel ontology that exploits structural Thai custom elements to facilitate direct get entry to core knowledge entities like element overviews or repair procedures. At this factor, the Thai custom expertise structure already receives four out of five stars. The addition of annotation vocabularies completes the part with the achievement of Thai custom 5-star.

RESULT

The aggregate of datasets and interoperability universes (Web-based offerings and Linked Data), supported by way of the semantics furnished via the links, is one of the main blessings of our methodology.

This makes it simpler to address troubles associated with global data, which require a cross-disciplinary strategy and lets in unlocking data storage tower facilitating their reuse throughout groups and communities. We have chosen to store the RDF dataset into a specialised RDF depository; in specific, the data have been kept into an Apache Jena Fuseki Server. It is quint

essential to have in consciousness that in this phase, the data are accessible on the local Web. As well, the ontology developed for the TCO has been posted online. In order to allow HTTP get admission to to the data, a front-end of LOD has been chosen and organized. In specific, we have selected the Jetty and Pubby employment of the LOD API fixation.

This frontend ensures get entry to via HTTP to our data and enables content mediation to allow customers to request the data in a number of formats.

The remaining section in the practice has been to allow get right of entry to the RDF shop settled up in the first phase. For this, we have configured our Apache Jena Fuseki keep to be manageable thru the SPARQL HTTP protocol and have enabled public access.

This open-access allows anyone to question our repository the usage of the SPARQL language, however it is essential to be aware that this access should be restricted the use of widespread HTTP security machineries and an extra particular configuration of the depository. The file comprising the Thai custom RDF data is also handy on line as shown in Fig. 2 below.

Linked Open Data is an information link via Uniform Resource Identifier (URI) identifier, which one URI is used to pick out only one thing, such as lunar calendar, activities, custom synonym, custom name, custom meaning, etc. in create links to URI with open expertise can request the URI through the internet browser, which will display data associated to that is in the standard RDF structure and can be linked to different URIs.

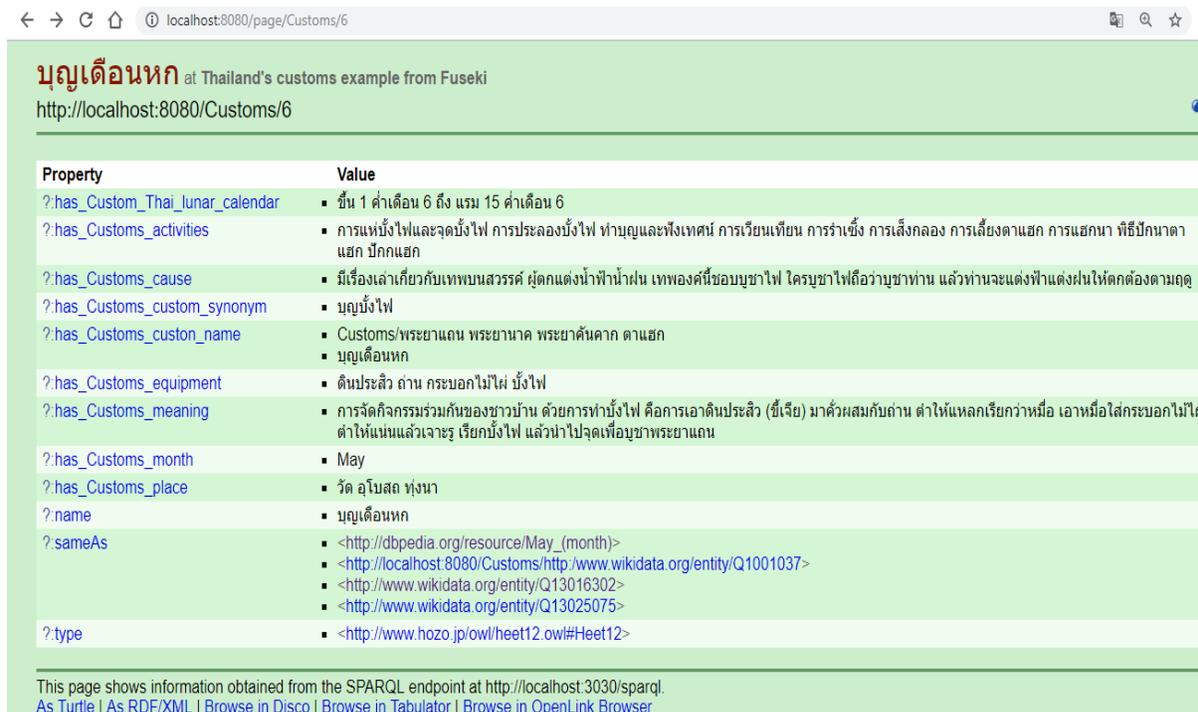


Figure 2. An example of Thai custom linked open data page

We provide instruction and high-quality practices on architectural methods to publishing Linked Data; selecting URIs and vocabularies to classify and define resources; deciding what data to return in a description of a useful resource on the Web, see Fig.1; methods for automatic connecting of records sets; and investigating and correcting methodologies for Linked Data arrangements. We give an overview of current Linked Data

functions and then look at the architectures that are used to devour Linked Data from the Web, alongside current equipment that allow these.

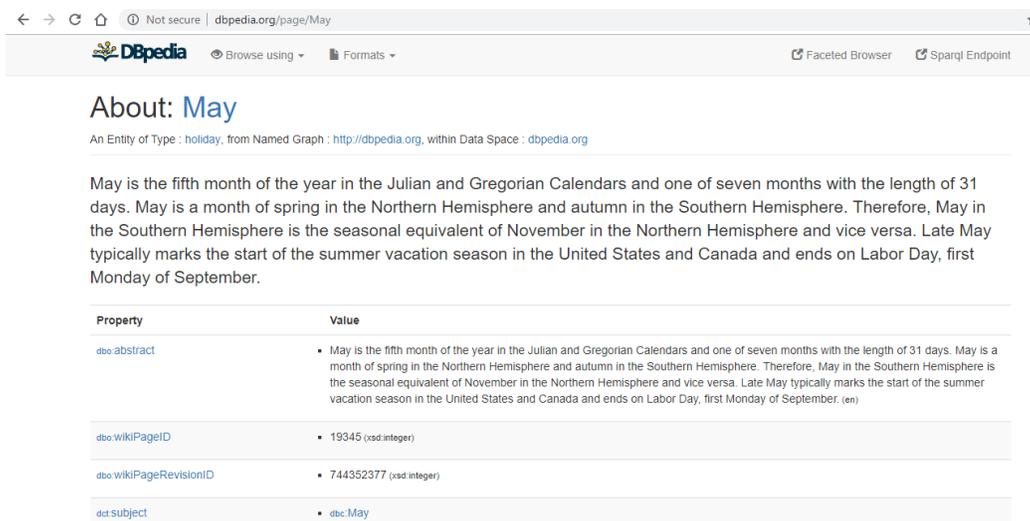


Figure 3. DBpedia linked data

All Thai custom RDF written in Turtle need to be used interior the query language part of the SPARQL Protocol and RDF Query Language (SPARQL), which makes use of a Turtle/N3 fashion syntax for the Triple patterns and for RDF triples in the

CONSTRUCT clause. This approves using RDF written in Turtle to enable forming "queries by using example", using the data to make a preliminary question which can then be edited to use variables where bindings are wanted, see Fig. 4.

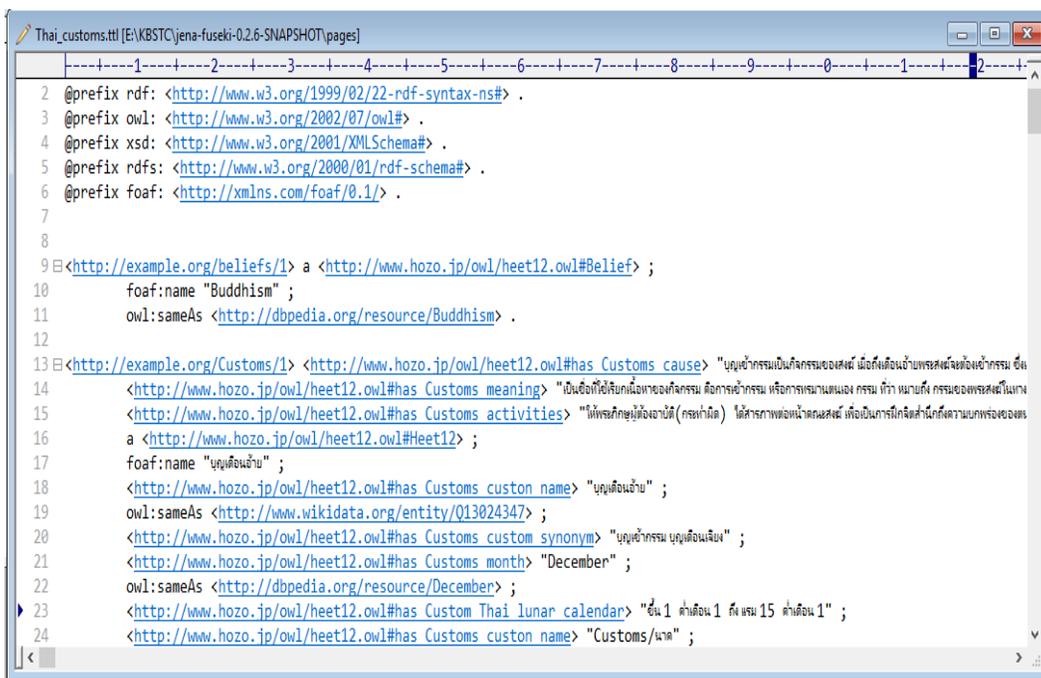


Figure 4. A snap short of Thai custom Turtle file format

EVALUATION

The objective of application-based processes in linking with querying is, of course, to aspect for similarities instead than for accuratysimilar values, that is, to introduce tender as an alternative than human evaluation. For put into effect so, various characteristics for evaluating ontology may want to be taken into account (Poveda-Villalón, Gómez-Pérez, & Suárez-Figueroa, 2014). In order to elevate out this movement, it is appropriate to use Linked Data semantic search evaluation, reasoners and syntax validators. For looking out existing ontologies, the Thai custom ontology, Linked Open Vocabularies, and search engines can be used together. The Thai custom ontology is focused on ontologies for the Thai custom Linked Data and associated

domains. This classify system affords the user with curated data and Linked Data evaluation features. In addition to the hassle of discovering a beneficial measuring principle, a project is to devise a principle of application-based contrast that is practical in connection with question processing. The application-based assessment technique is an evaluation of performance through the use of the application. Application by trying out the linkage of understanding from the open data link system (Link Open Data), the contrast result of the average assessment is a high level (mean = 3.97). When separated in each aspect, it is observed that in a high stage in all aspects including, content (mean= 4.00), accuracy (mean= 4.33), layout (mean= 3.50), handy to use

(mean= 3.83), timeliness (mean= 3.83), and perceived usefulness (mean= 4.33).

CONCLUSION

To acquire perfection when applied to Linked Open Data technology, we adjusted the Thai custom ontology in accordance to the experts' recommendation of the ontology before using it. Therefore, the ontologies that have been evaluated and modified according to the recommendations of experts are in the form of an ontology file that can be used as an intermediate of explanation as a substitute of knowledgeable to understand the principles in the scope of the Thai custom. However, due to the ontology of knowledge of the Thai custom derived from the research, it is no longer complex sufficient to be applied to the semantic web technology. We experimented it the use of Linked Data methods in order to explainable the ontology greater without difficulty and to help in addition information sharing development. We proceeded to convert the file format to use in open data linking. The structure of the language is primarily based on the structure of the RDF/XML and generated to Turtle structure that can be used to join to other assets backyard (Linked Open Data).

ACKNOWLEDGEMENTS

This research supported by the Digital Humanities Research Group (DHRG) and the Graduate School, KhonKaen University, Thailand.

REFERENCES

1. Apache Jena. (2019). Fuseki: serving RDF data over HTTP. Retrieved May 2019, from https://jena.apache.org/documentation/serving_data/fuseki1
2. Berners-Lee, T.(2019). Linked Data. (July 2019) URL: <http://www.w3.org/DesignIssues/LinkedData.html>.
3. Butte, P &Puarungroj, W. (2016). The Development of Semantic Web for Searching Cultural Information InLoei Province. *Information Technology Journal*, 12(2), 33-41.
4. Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked data — The story so far *International Journal Semantic Web Information Systems*, 5 (3), pp. 1-22
5. Chaikhambung, J. &Tuamsuk, K. (2016). Semantic Onlogy of Knowledge on Ethnic Groups in Thailand. *IccsNT 2016- The 5th International Conference on Computer science and Network Technology*, 10-11 December 2016, Changchun Normal University Changchun, China.
6. Chansanam, W., Tuamsuk, K., Kwiecien, K., Ruangrajitpakorn, T., Supnithi, T.(2014).Development of the belief culture ontology and its application: Case study of the GreaterMekongsubregion. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8943, pp. 297-310. doi: 10.1007/978-3-319-15615-6_22
7. Chotirat, W., Boonrawd, P., & Na Wichian, S. (2011). Developing an Ontology Knowledge Based for Automatic Online News Analysis. *Information Technology Journal*, 7(14), 13-18.
8. Cyganiak, R. &Bizer, C. (2008). Pubby - A linked data frontend for SPARQL endpoints. Retrieved May 2019, from <http://wifo5-03.informatik.uni-mannheim.de/pubby/>
9. Ferrara, A., Nikolov, A. &Scharffe, F. (2011). Data linking for the Semantic Web, *International Journal of Semantic Web and Information Systems*, 7(3), 46-76.
10. Hitzler, P., Krötzsch, M., Parsia, B., Patel-Schneider, P.F., Rudolph, S. (eds.): *OWL 2 Web Ontology Language: Primer*. W3C Recommendation (October 2019), available at <http://www.w3.org/TR/owl2-primer/>
11. Hussain, A., Manikanthan, S.V., Padmapriya, T., Nagalingam, M. (2020). Genetic algorithm based adaptive offloading for improving IoT device communication efficiency. *Wireless Networks*, 26 (4), pp. 2329-2338.
12. Jr, T. T. (2009). The DBpedia data set. Retrieved May 2019, from <http://wiki.dbpedia.org/Datasets>
13. Noy, N. F. & McGuinness, D. L. (2001). *Ontology development 101: A guide to creating your first ontology*. Retrieved January 2019, from https://protege.stanford.edu/publications/ontology_development/ontology101.pdf
14. Poveda-Villalón, M., Gómez-Pérez, A. & Suárez-Figueroa, M. C. (2014). Oops! (Ontology pitfall scanner!): An on-line tool for ontology evaluation. *International Journal on Semantic Web and Information Systems*, 10(2), 7-34.
15. Vrandečić, D. &Krötzsch, M.(2014). Wikidata: A Free Collaborative Knowledge Base. *Communications of the ACM*.
16. W3C: RDF Schema 1.1 – W3C Recommendation. (October 2019), available at <http://www.w3.org/TR/rdf-schema/>