



TACO: An Ontology in Turkish for Identifying and Controlling Plant Pests, Weeds and Diseases

TACO: Bitki Zararlıları, Yabani Otlar ve Hastalıkların Tespiti ve Kontrolü İçin Türkçe Bir Ontoloji

Övünç Öztürk 

Manisa Celal Bayar Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümü, Manisa, TÜRKİYE

Corresponding Author / Sorumlu Yazar*: ovunc.ozturk@cbu.edu.tr

Abstract

While ensuring a more sustainable production, because of reduced chemical usage it is more complicated to control plant pests, diseases and weeds in smart agriculture. For this reason, it is of great importance to detect pests, diseases and weeds at the earliest stage. It is important that both farmers and the artificial intelligence applications developed for agricultural control should be able to detect these organisms and to know the agricultural control methods. Semantic technologies and ontologies provide machine interpretable information and solutions for heterogeneity. This study presents the Turkish Agricultural Control Ontology (TACO), which is built in Turkish and contains information about plant pests, diseases and weeds common in Turkey. The contributions of the study are that it is the first Turkish ontology built in this field and that the methods of agricultural control are included within the scope of the ontology. According to the commonly used ontology evaluation metrics, TACO is predominantly characterized as a deep classification taxonomy. In addition, it was concluded that the classes in the ontology have an evenly distributed and sufficient number of class individuals.

Keywords: ontology, artificial intelligence, agricultural control, plant pests, plant diseases, weeds

Öz

Akıllı tarımda daha sürdürülebilir bir üretim sağlanırken, kimyasal kullanımının azalması nedeniyle bitki zararlıları, hastalıkları ve yabancı ot kontrolü daha karmaşık hale gelmektedir. Bu nedenle zararlı, hastalık ve yabancı otların erken aşamada tespit edilmesi büyük önem taşımaktadır. Hem çiftçilerin hem de tarımsal mücadele için geliştirilen yapay zeka uygulamalarının bu organizmaları tespit edebilmesi ve tarımsal mücadele yöntemlerini bilmesi önemlidir. Semantik teknolojiler ve ontolojiler, makine tarafından yorumlanabilir bilgiler ve heterojenlik için çözümler sağlar. Bu çalışmada Türkiye'de yaygın olarak görülen bitki zararlıları, hastalıkları ve yabancı otlar hakkında bilgiler içeren Türkçe olarak oluşturulmuş Türk Tarımsal Kontrol Ontolojisi (TACO) sunulmaktadır. Çalışmanın katkıları, bu alanda yapılan ilk Türk ontolojisi olması ve tarımsal mücadele yöntemlerinin ontoloji kapsamında yer almasıdır. Sıkça kullanılan ontoloji değerlendirme metriklerine göre TACO, ağırlıklı olarak derin bir sınıflandırma taksonomisi olarak nitelendirilmiştir. Ayrıca ontolojideki sınıfların eşit olarak dağılım gösteren, yeterli sayıda sınıf örneğine sahip olduğu sonucuna varılmıştır.

Anahtar Kelimeler: ontoloji, yapay zeka, tarımsal kontrol, bitki zararlıları, bitki hastalıkları, yabancı otlar

1. Introduction

Today, with the increasing world population, smart agriculture applications that increase productivity and reduce resource consumption is gaining attention. There are many ongoing studies in this field. This subject has also been studied intensively in Turkey recently [1]. Smart agriculture practices aim yield increase; reduced chemical usage; disease, pest and health status monitoring and automated agricultural production. One of the challenges in implementation of smart farming practices is to provide the necessary information needed by the applications.

There are huge amounts of heterogeneous, unstructured and non-machine interpretable data, which are presented to third parties in various formats. Ontologies and semantic technologies are useful tools to integrate and harmonize data from different sources.

A number of dictionaries and ontologies have been defined for different purposes for the agricultural community, most of which are hosted in the AgroPortal ontology repository [2]. Some important ontologies related to this work, both in the AgroPortal and in other sources, are listed below;

Plant Ontology (PO) [3]: An ontology collection developed by the Plant Ontology Consortium. These ontologies describes the anatomical structures, growth and development stages of the organisms in the Viridiplantae group. PO is designed for use in multiple applications, including genetics, genomics, phenomics, developmental biology, taxonomy and systematic studies, semantic applications and education.

Plant Trait Ontology (PTO) [4]: An ontology that describes phenotypic characters in plants. Each phenotypic character is a distinguishable feature, characteristic, quality or phenotypic characteristic of a developing or mature plant.

Plant Experimental Conditions Ontology (PECO) [5]: a structured, controlled vocabulary describing treatments, growing conditions, and/or types of studies used in plant biology experiments.

Plant Stress Ontology (PSO) [6]: describes the biotic and abiotic stresses that a plant may encounter.

IDOPlant Ontology [7]: an ontology of infectious plant diseases.

Agronomy Ontology [8]: provides terms from the agronomy domain that are semantically organized and can facilitate the collection, storage and use of agronomic data.

AGROVOC [9]: a Linked Open Data Set about agricultural concepts, terms, definitions and relationships.

CropPestO [10]: an ontology covering crops, related pests and diseases, their associated symptoms, and suggested control methods. The ontology has been built in English and labelled in Spanish.

Ontology of Crop Pest Control [11]: defines a general model of crop pest control that contains related datasets on crops, pests and pest control measures.

Plant-Pathogen Interactions Ontology (PPIO) [12]: an ontology describing plant-pathogen interactions.

Pests in Crops and their Treatments Ontology (PCT-O) [13]: an ontology developed to explain the relationship between crops, pests and treatments. It contains 462 products, 549 pests and 42397 treatments.

PestOn [14]: an ontology for the domain of pesticide products so that their characteristics and features can be easily accessed, interoperable, and jointly usable by food system stakeholders.

The ontology in this study presents crops, associated pests (insects, diseases, weeds, nematodes, and mammals) and pest

management. Table 1 compares ontologies in the agricultural field. IDOPlant, Agrovoc, PCT-O, CropPestO and Ontology of Crop Pest Control are ontologies that describe plant pests. However, the gaps of these ontologies can be listed as follows:

- None of these ontologies includes information about weeds.
- There is no Turkish language support other than Agrovoc.
- They contain no or very limited information on crop pest control, other than PCT-O or Crop Pest Control Ontology.

The ontology that is closest to the presented ontology is Crop-Pest Ontology [10]. However, Crop-Pest Ontology does not provide Turkish language support. TACO has been defined in Turkish Language and contains information about diseases, pests, weeds especially seen in Turkey. Another similar and comprehensive study is the AGROVOC dictionary [9]. AGROVOC offers Turkish support. However, TACO also includes pest control methods. TACO also has a richer pesticide and weed content and the Turkish equivalents of the terms are based on expert opinions. However, in the tables and figures in the article, we used the English equivalents of the Turkish terms in the ontology to provide readability for non-Turkish readers of the article.

At this point, it would be appropriate to remind that the content of the ontology was created entirely by utilizing and adhering to the technical instructions published by the Turkish Ministry of Food, Agriculture and Livestock [15].

The next section describes the method followed and the materials used by this study. It contains the following subsections; ontology reuse, classes and class hierarchy, properties and instances. Section 3 presents the results and discussion. It contains the following two subsections: ontology statistics and experimental evaluation of the ontology. Finally, section four concludes the article with some useful recommendations.

Table 1. Comparison of ontologies for the agricultural community.

ontology	# entities	scope	language
PO	2018	plant anatomical entities, plant structure development stages	English
PTO	5260	phenotypic traits in plants	English
PECO	3119	treatments, growing conditions, and/or study types used in plant biology experiments	English
PSO	3762	major types of plant stress	English
IDOPlant	660	infectious plant diseases	English
Agronomy Ontology	3736	practices, techniques, and variables used in agronomic experiments	English
Agrovoc	41016	all areas of interest to FAO	42 languages
CropPestO	12404	plant pests and diseases	Spanish, English
Ontology of Crop Pest Control	1151	crop pest control	English
PPIO	2508	plant-pathogen interactions	English
PCT-O	43408	pests and suitable treatments	English
PestOn	16000	pesticide product information	English
TACO	1036	plant diseases, pests and weeds	Turkish

2. Material and Method

TACO ontology was built using OWL ontology language and Protégé Ontology Editor [16].

2.1. Ontology Reuse

It is planned to publish the TACO ontology in the Linked Open Data Cloud (LOD) [17] after obtaining the necessary permissions from TAGEM. In accordance with Linked Open Data standards, the terms in the TACO will be matched with those in AGROVOC, which is a standard vocabulary that provides information about organisms, plants and their products.

The matching process (Figure 1) is carried out by mapping the equivalent classes using "owl:equivalentClass" construct provided by the OWL ontology modeling language [18].

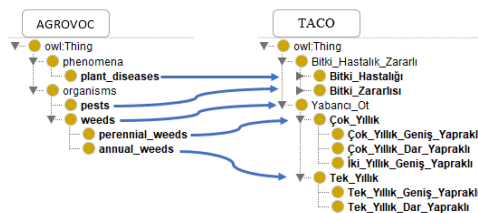


Figure 1. Mapping AGROVOC and TACO.

2.2. Classes and Class Hierarchy

To build the class hierarchy of the TACO ontology, first the common terms in the domain are listed. Then draft of the hierarchy is completed by adding concepts that are more specific. Figure 2 depicts the partial class hierarchy of TACO and the individual counts of the classes shown. The total count of the

classes in the ontology is 133 and there are 131 “subclassOf” relations between these classes.

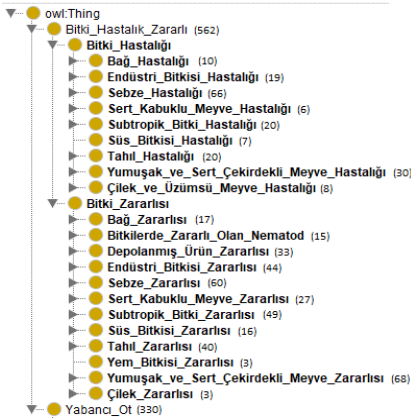


Figure 2. Partial hierarchy of TACO.

2.3. Properties

Table 2. Data properties defined in TACO.

owl:DataProperty	rdfs:domain	rdfs:range	#assertions
warfare	Plant_Disease_Pest U Weed	xsd:string	
biological_warfare	Plant_Disease_Pest U Weed	xsd:string	128
biotechnological_warfare	Plant_Disease_Pest U Weed	xsd:string	16
physical_warfare	Plant_Disease_Pest U Weed	xsd:string	21
physical_and_chemical_warfare_combination	Plant_Disease_Pest U Weed	xsd:string	6
quarantine_measure	Plant_Disease_Pest U Weed	xsd:string	41
chemical_warfare	Plant_Disease_Pest U Weed	xsd:string	570
conventional_warfare	Plant_Disease_Pest U Weed	xsd:string	1922
mechanical_warfare	Plant_Disease_Pest U Weed	xsd:string	70
scientific_name	Plant_Disease_Pest U Weed	xsd:string	893

3. Results

3.1. Ontology Metrics

The TACO ontology defines a total of 133 classes, 891 individuals and 7324 axioms. Figure 4 shows the statistical information of TACO obtained from the Protégé Ontology Editor.

Ontology metrics:	
Metrics	
Axiom	7324
Logical axiom count	5035
Declaration axioms count	1034
Class count	133
Object property count	0
Data property count	9
Individual count	891
Annotation Property count	3
Class axioms	
SubClassOf	131

Figure 4. The statistical information of TACO.

The data properties (between the classes and the literals) and the assertions of these properties are depicted in Table 2. There are 9 data properties and 3677 assertions of these properties in total.

2.4. Instances

891 individuals are defined in TACO. Figure 3 shows an example class individual and its attribute values. Table 3 depicts the individual examples of the some important classes in the ontology.

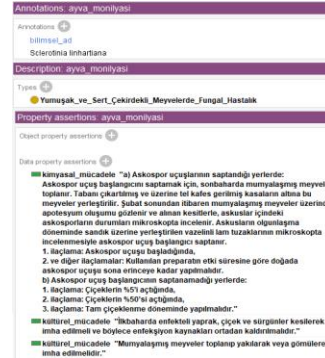


Figure 3. An example class individual.

3.2. Experimental Evaluation of the Ontology

In this section, TACO ontology is evaluated using the metrics from OntoQA framework [19], which is one of the most used ontology evaluation tools. QntoQA metrics are classified as schema and knowledge base metrics.

Within the scope of this study, relationship diversity (RD), attribute richness (AR) and schema deepness (SD) metrics from schema metrics and class utilization (CU), class connectivity (CC), class importance (CI), relationship utilization (RU), relationship importance (RI) and average population (AP) from knowledge base metrics are used.

Relationship Diversity (RD): shows the percentage ratio of rich relations between classes to all relations between classes. Rich relations are obtained by excluding hierarchical (subClassOf) relationships. The RD value for TACO was calculated as 0 because there are not any relationships between classes (except subClassOf) in TACO. Attribute Richness (AR): indicates the number of attributes per class. For TACO, this metric is calculated as 9/133 ≈ 0,067.

Table 3. Individuals defined in TACO.

Class	Examples of Instances
Vineyard_Disease	vineyard_anthraxnose, esca, vineyard_powdery_mildew, vineyard_downy_mildew...
Vineyard_Pest	vineyard_cicada, vineyard_thrips, vine_leafroller_tortrix, june_beetle...
Plant_Pest_Nematode	wheat_gall_nematode, dagger_nematode, potato_cyst_nematode...
Stored_Product_Pest	angoumois_grain_moth, mill_moth, khapra_beetle, flour_mite...
Industrial_Plant_Disease	cercospora_leaf_spot, late_leaf_spot...
Industrial_Plant_Pest	anise_moth, black_bean_aphid, pink_bollworm...
Vegetable_Disease	potato_late_blight_fungus, bean_rust...
Vegetable_Pest	artichoke_moth, carrot_fly, psychid_moth, colorado_potato_beetle...
Hard_Shelled_Fruit_Disease	leaf_spot_of_pistachio, almond_canker, walnut_anthraxnose...
Hard_Shelled_Fruit_Pest	twig_borer_moth, almond_seed_wasp, pear_blight_beetle, european_fruit_lectanium...
Subtropical_Plant_Disease	blue_mold_rot, citrus_storage_moulds...
Subtropical_Plant_Pest	cottony_cushion_scale, florida_wax_scale, olive_moth, cottony_camellia_scale...
Ornamental_Plant_Disease	begonia_mildew, rose_rust...
Ornamental_Plant_Pest	rose_shoot_sawfly, european_brown_scale...
Grain_Disease	stinking_smut, barley_smut, corn_smut, wheat_yellow_rust, rice_blast_disease...
Grain_Pest	european_rabbit, italian_tree_cricket, wheat_thrips, italian_locust, sunn_pest...
Feed_Crop_Pest	six_belted_clearwing, alfalfa_weevil...
Soft_and_Hard_Fruit_Disease	honey_fungus, quince_monilia, crown_gall...
Soft_and_Hard_Fruit_Pest	hawthorn_mite, brown_tail_moth, apple_rust_mite...
Berry_Fruit_Disease	raspberry_spur_blight, strawberry_powdery_mildew, common_spot_of_strawberry...
Strawberry_Pest	two_spotted_red_spider, carmine_spider_mite...
Weed	garden_vetch, clovegrass, cockspur, cornflower, johnsongrass, deathcap...

As the number of attributes per class increases, the quality of the modeled ontology also increases. A decrease in the number indicates that the number of attributes belonging to the classes is low and the classes are not extensively modeled.

Schema Depth (SD): shows the average number of subclasses per class. The SD value for TACO is calculated as $131/133 \approx 0.984$. The interpretation of the result is highly dependent on the structure of the ontology. While the schema depth is expected to be low in ontologies modeling a very specific field, the schema depth of the ontology generally increases as the modeled field expands.

Class Usage (CU): shows the number of instantiated classes divided by the number of all classes. The CU value for TACO is calculated as $133/133=1$.

Class Connectivity (CC): shows the total number of relationship instances of the class with other class instances. The CC values for TACO is 0 because there are not any relations between instances of classes.

Class Importance (CI): It shows the ratio of the number of samples belonging to the class and its subclasses to the total number of samples. The CI values for some classes in TACO were calculated as follows: Plant_Disease ($\approx 0,209$), Plant_Pest ($\approx 0,421$), Vineyard_Disease ($\approx 0,11$), Weed ($\approx 0,37$), Subtropical_Fruit_Pest ($\approx 0,055$), Grain_Disease ($\approx 0,22$). This metric, along with class connectivity, serves to understand the important classes in the ontology.

Relationship Usage (RU): shows the ratio of the number of relationships used by instances of a class to the number of

relationships defined for the class. The RU value is calculated as 1 for all classes in TACO. In other words, all relations defined for the class were used by the instances of the class. If the result is low, it means the relations are not used enough at the instance level. A higher value of this metric (closer to 1) means that relationships defined at the schema level are also used at the instance level.

Relationship Importance (RI): shows the ratio of the number of instances of a relationship to all relationship instances in the ontology. The RI values for relationships in TACO were calculated as follows:

biological_warfare	($\approx 0,035$),
biotechnological_warfare	($\approx 0,004$),
physical_warfare	($\approx 0,006$),
physical_and_chemical_warfare_combination	($\approx 0,002$),
quarantine_measure	($\approx 0,011$),
chemical_warfare	($\approx 0,155$),
conventional_warfare	($\approx 0,523$),
mechanical_warfare	($\approx 0,019$),
scientific_name	($\approx 0,243$).

Average Population (AP): is obtained by dividing the number of class individuals in the knowledge base by the number of classes. This metric indicates whether the instance count is sufficient to represent all of the knowledge in the schema. The AP value for TACO was calculated as $\approx 6,67$.

These metrics were evaluated with the methodology presented in [20]. The percentage of the important metrics selected according to this methodology were scored as shown in Table 4. Table 5 shows the selected OntoQA metrics and their values for TACO (in percentage format).

According to the evaluation results, the RD value was calculated as 1. This result shows that "rich relationships" in ontology are

not modeled as comprehensively as “hierarchical relationships”. Considering that TACO is a biological classification taxonomy, this can be considered an expected result.

The SD value was calculated as 5. The result is high as expected. TACO models plant diseases, pests and weeds. This extensive domain knowledge results in the formation of a deep ontology taxonomy.

AR value was calculated as 1. This result indicates that the number of attributes per class is not high. Since ontology focuses

Table 4. Evaluation Scale of the Ontology Metrics.

Scale	1	2	3	4	5
Score	[0-20]%	[20-40]%	[40-60]%	[60-80]%	[80-100]%

Table 5. Scoring the metric values of TACO.

Metric	Value	Scale
RD	%0	1
SD	%98	5
AR	%7	1
CU	%100	5
Cl _{avg}	%2	1

4. Results and Discussion

This article presents an ontology, namely TACO for plant diseases, pests and weeds, especially common in Turkey. Thus, a knowledge base that can be used in artificial intelligence-based smart agriculture applications has been built. TACO has been evaluated using OntoQA which is one of the most widely used ontology evaluation tools.

There is a need for the increase and development of smart computer solutions aiming disease, pest and weed control. In this sense, the contributions of the study to the research area can be summarized as follows:

- A fundamental study has been carried out on the increase of Turkish agricultural databases and their linking with other data sources,
- Studies on the compatibility and economic efficiency of the pesticides for disease, pest and weed control are gaining importance. The creation of relevant extensions of the presented ontology is critical for these studies,
- The ontology-based data will facilitate the integration of software systems on agribusiness with information in the field of disease, pest and weed control.

TACO is the first ontology on pest and weed control in Turkish. In addition to pest control, it provides information on weed control and provides a complementary study of agricultural control.

As a future work, it is planned to expand the TACO ontology according to the evaluation results. In addition, it is intended to extend the ontology with the images of class instances. Another future work is to make the ontology accessible on the LOD cloud via the REST API. It is also planned to add the multi-language support feature so that the widespread impact of the work will be increased.

Ethics committee approval and conflict of interest statement

This article does not require ethics committee approval.

This article has no conflicts of interest with any individual or institution.

References

on classes and class instances, this can be considered as an expected result.

The CU value was calculated as 5. This result states that all classes defined in the schema are used at the individual level. Finally, the average CI value of 1 indicates that all classes in the ontology have approximately equal importance. When the TACO ontology is examined, it will be seen that the distribution of class samples is roughly equal. This low average value of CI is due to the nature of the ontology.

[1] Tarım ve Orman Bakanlığı, 2019. Akıllı Tarım Platformu. <http://www.akillitarim.org/tr/> (Date Of Access: 22.02.2023)

[2] Jonquet, C., Toulet, A., Arnaud, E., Aubin, S., Yeumo, E.D., Emonet, V., Graybeal, J., Laporte, M.A., Musen, M.A., Pesce, V., Larmande, P. 2018. AgroPortal: A vocabulary and ontology repository for agronomy: Computers and Electronics in Agriculture, Vol. 144, p. 126-143. DOI: 10.1016/j.compag.2017.10.012

[3] Jaiswal, P., Avraham, S., Ilic, K., Kellogg, E.A., McCouch, S., Pujar, A., Reiser, L., Rhee, S.Y., Sachs, M.M., Schaeffer, M., Stein, L., Stevens, P., Vincent, L., Ware, D., Zapata, F. 2005. Plant Ontology (PO): a Controlled Vocabulary of Plant Structures and Growth Stages: Comparative and Functional Genomics, Vol. 6(7-8), p. 388-397. DOI: 10.1002/cfg.496

[4] Arnaud, E., Cooper, L., Shrestha, R., Menda, N., Nelson, R.T., Matteis, L., Skofic, M., Bastow, R., Jaiswal, P., Mueller, L., McLaren, G. 2012. Towards a reference Plant Trait Ontology for modeling knowledge of plant traits and phenotypes. International Conference on Knowledge Engineering and Ontology Development, 4-7 October, Barcelona, 220-225.

[5] European Bioinformatics Institute, 2017. Plant Experimental Conditions Ontology. <https://biportal.bioontology.org/ontologies/PECO> (Date Of Access: 22.02.2023)

[6] European Bioinformatics Institute, 2020. Plant Stress Ontology. 2020. <https://biportal.bioontology.org/ontologies/PLANTSO> (Date Of Access: 22.02.2023)

[7] Walls, R.L., Smith, B., Elser, J., Goldfain, A., Stevenson, D.W., Jaiswal, P. 2012. A plant disease extension of the infectious disease ontology. International Conference on Biomedical Ontology, 21-25 July, Graz, 1-5.

[8] Aubert, C., Buttigieg, P.L., Laporte, M.A., Devare, M., Arnaud, E. 2017. Agronomy Ontology. <http://purl.obolibrary.org/obo/agro.owl> (Date Of Access: 22.02.2023)

[9] Caracciolo, C., Stellato, A., Morshed, A., Johannsen, G., Rajbahndari, S., Jaques, Y., Keizer J. 2013. The AGROVOC Linked Dataset: Semantic Web Journal, Vol. 4(3), p. 341-348. DOI: 10.5555/2786071.2786087

[10] Rodriguez-Garcia, M.A., Garcia-Sanchez, F. 2020. CropPestO: An Ontology Model for Identifying and Managing Plant Pests and Diseases: Communications in Computer and Information Science, Vol. 1309, p. 18-29. DOI: 10.1007/978-3-030-62015-8_2

[11] Onkov, K. 2020. Ontology of Crop Pest Control. 3rd International Conference on Information Science and Systems, 19-22 March, Cambridge, 8-12.

[12] Iglesias, A.R., Aranguren, M. E., González, A.R., Wilkinson, M.D. 2013. Plant Pathogen Interactions Ontology (PPIO). International Work-Conference on Bioinformatics and Biomedical Engineering, 18-20 March, Gradana, 695-702.

[13] Lacasta, J., Lopez-Pellicer, F.J., Espejo-García, B., Noguera-Iso, J., Zarazaga-Soria, F.J. 2018. Agricultural recommendation system for crop protection: Computers and Electronics in Agriculture, Vol. 152, p. 82-89. DOI: 10.1016/j.compag.2018.06.049

- [14] Medici, M., Dooley, D., Canavari, M. 2022. PestOn: An Ontology to Make Pesticides Information Easily Accessible and Interoperable: Sustainability, Vol. 14, p. 66-73. DOI: 10.3390/su14116673
- [15] TAGEM, 2018. Zirai Mücadele Teknik Talimatları Kitapları Cilt 1-Cilt 6. T.C. Tarım Orman ve Hayvancılık Bakanlığı, Ankara, 1841 p.
- [16] Noy, N., Ferguson, R., Musen, M. 2000. The knowledge model of Protégé-2000: combining interoperability and flexibility, 12th International Conference on Knowledge Engineering and Knowledge Management, 2-6 October, French Riviera, 17-32.
- [17] Bizer, C., Vidal, M.E., Skaf-Molli, H. 2018. Linked Open Data. p 2096-2101. Liu, L., Özsu, M.T., ed. 2018. Encyclopedia of Database Systems, Springer Nature, Berlin, 4866 p.
- [18] Welty, C., McGuinness, D.L., Smith, M.K. 2004. OWL Web ontology language guide. <https://www.w3.org/TR/owl-guide/> (Date Of Access: 22.02.2023)
- [19] Tartir, S., Arpınar, I.B., Moore, M., Sheth, A.P., Aleman-Meza, B. 2005. OntoQA: Metric-based ontology quality analysis. IEEE Workshop on Knowledge Acquisition from Distributed, Autonomous, Semantically Heterogeneous Data and Knowledge Sources, 27 November, New Orleans, 45-53.
- [20] Ebietomer, E.P., Ekuobase, G.O. 2013. Evaluation of Ontology for Nigerian Case Laws: Computing, Information Systems, Development Informatics & Allied Research Journal, Vol. 4(3), p. 1-6.