Computing, Information Systems, Development Informatics & Allied Research Journal
ISBN 978-2257-44-7(Print) ISSN 2167-1710 (online)
Vol. 4. No. 3, 2013 – www.cisdijournal.net
www.iiste.org

Evaluation of Ontology for Nigerian Case Laws

[1]E. P. Ebietomere
Department of Computer Science,
University of Benin
Benin City, Nigeria
Princewill.ebietomere@uniben.edu

[2]Godspower O. Ekuobase
Department of Computer Science,
University of Benin, Benin City, 300001, Nigeria
Godspower.ekuobase@uniben.edu

[1]Corresponding Author

ABSTRACT
The problem of inaccurate search result has generated a lot of controversies leading to the advent of Web 3.0 also known as the Semantic Web. At the heart of Semantic Web is Ontology. The need to evaluate Ontology is critical to aligning the Ontology with its purpose. Besides, it provides a means of self or third party assessment of Ontologies. In order to determine the quality of a novel Ontology – Ontology for Nigerian Case Laws (ONCL), we employed the OntoQA – a metric based Ontology evaluation tool – to expose how well the Ontology is in alignment with its purpose of efficient and effective access and retrieval of Nigerian Case Laws.

Keywords: Ontology, Semantic web, Ontology evaluation.

1. INTRODUCTION

The problem with inaccurate search result has been linked to lack of proper storage structure of digital information repository as is the case with the current web also referred to as web 2.0. Ontology is a critical component of semantic web being the only way the content of the web can be marked up for automatic information processing [21,24]. Several reasons have been identified for building ontology. These include, sharing of common understanding of the structure of information among people or software agents, enabling reuse of domain knowledge, making domain assumptions explicit, separating domain knowledge from the operational knowledge and analysing domain knowledge amongst others.

In recent times, there is a proliferation of ontology in various domains including the legal domain [5,6,7,8,25]. This proliferation has made it imperative especially for users of ontology in various applications to choose the ontology that best suits their applications [24]. Choosing an appropriate ontology is herculean and end users over the years have relied on experience [24] which actually involves a grave risk. This is because the flaws in the ontology would only be discovered after the ontology has been put to use. A way out of this risk is for the ontology end users to be aware of how well the chosen ontology for their applications will fare before deploying it. The only way they can be sure of how well this chosen ontology will fare is via evaluation of the ontology.

Besides, for the effort of ontology builders not to be in futility, evaluation of ontology is critical because it will serve not only as a guide in building the ontology correctly but also in ensuring that the ontology is in strong alignment with its purpose. This paper therefore seeks to expose how to evaluate the quality of Ontology, Ontology for Nigerian Case Laws (ONCL), in line with its purpose of efficient and effective access and retrieval of case laws; so as to determine its quality as well as help expose means of improving it.

2. BACKGROUND INFORMATION

Ontology is a critical component of semantic web being the only way the content of the web can be marked up for automatic information processing [21,24]. Ontology as a concept has its root in Philosophy, where it is seen as “the science of what is, of the kinds and structures of objects, properties, events, processes and relations in every area of reality” [22]; it is from here the term ontology was borrowed to be used in Computer Science and Artificial Intelligence (AI), in particular. The most prominent definition of the term in AI is that of [14] which defined ontology as an “explicit specification of a conceptualization”. Other more explicit definitions were given by [17, 20, 23, 26, 27].

2.1 Classification of Ontologies

Several ontologies exist in literature [5, 6, 7, 8, 25]. Generally, we can classify ontology based on purpose, formality or complexity [7] but the most pronounced is the classification by purpose or ontological commitment. [27] describes three types of ontological commitments – task commitment, method commitment and domain commitment.

2.2 Building Ontology

There are several reasons for building ontology. Some of which are as follows [3, 20]:

- To share common understanding of the structure of information among people or software agents.
- To enable reuse of domain knowledge.
- To make domain assumptions explicit.
- To separate domain knowledge from the operational knowledge.
- To analyse domain knowledge.

It is important to note that in Software Engineering, lack of standard methodologies was the major issue that led to what was termed software crisis – inability of software to meet user requirements as well as software development project exceeding budget and deadline [9, 10, 11]. To tackle this crisis, Software Engineers had to come up with standard methodologies which today abound in the domain; semantic web inclusive. Today, several standard methodologies exist for the development and evaluation of ontology; some of which are highlighted as follows:

- Lenat and Guha Methodology [18] used in building CYC ontology, ontology conceived to capture shared knowledge about the world. The methodology consisted of the following steps:
  - i. Manual extraction of common sense knowledge.
  - ii. Computer aided extraction of common sense knowledge.
  - iii. Computer managed extraction of common sense knowledge.
- The Gruninger and Fox Methodology [15] that was used in construction of TOronto Virtual Enterprise (TOVE) ontology. They identified the following steps:
  - i. Specification of scenario motivating the construction of the ontology and its application.
  - ii. Developing set of competence questions that the ontology must be able to answer.
  - iii. Determining the terminologies the ontology should use.
  - iv. Specification of definition and constraints on the terminologies.
  - v. Evaluation of the ontology testing the competency question against completeness theorems [15].
- Another methodology is the KATUS methodology [4] whose focus is on development of ontologies for particular applications and employs the bottom up approach in doing that. The design process takes into consideration, the possibility of reuse of ontologies already developed for use of different applications in the same domain. Firstly, the application is specified and lists of relevant terms and tasks are provided. Secondly, a preliminary design is made according to the previous lists and specifications, which may include searching for already developed ontologies. Finally, the ontology is refined towards the final design of the application.
- Notable among the methodologies for building ontology, is the Noy and McGuinness methodology [20]. They made it clear that there is no specific correct way or methodology for developing ontologies, but did discuss general issues to consider and offered a possible process for developing ontology. The process they described was an iterative one where a rough first pass at the ontology is made and then a revision and refinement of the evolving ontology is carried out. They identified the following steps:
  Step I: Determine the domain and scope of the ontology
  Step II: Consider reusing existing ontologies
  Step III: Enumerate important terms in the ontology
  Step IV: define the classes and the class hierarchy
  Step V: Define the properties of the classes
  Step VI: Define the value of the slots
  Step VII: Create instances

2.3 Evaluating Ontology

It is necessary to rate or evaluate how well ontology is in alignment with its purpose. Several approaches exist in literature for the evaluation and validation of ontology. These approaches according to [29] can be:

(i) Gold Standard evaluation: this evaluation approach involves the comparison of a particular ontology with another ontology that is considered to be a standard. This kind of evaluation is believed to have found applications where ontologies are generated semi-automatically to tell if the process of generating the ontology is effective.

(ii) Criteria based evaluation: this takes ontology and evaluates it based on proposed criteria. Some of the criteria include consistency, completeness, conciseness, correctness, clarity, expandability, minimal ontological commitment, and minimal encoding bias amongst others. The problem with the criteria based evolution approach is in the fact such criteria as clarity and expandability cannot be measured and even some that can be demonstrated may be impossible to prove, e.g completeness of an ontology. Besides, other criteria may be extremely difficult to quantify, e.g correctness.

(iii) Task-based evaluation: this approach evaluates ontology based on the competency of the ontology in completing tasks. This approach can actually help in judging the suitability of a particular ontology for an application.

According to [24], the approaches can be Evolution based – in this case, changes in important characteristics of same ontology across different versions are tracked; Logical based – which uses rules in-built in the ontology languages the ontology builder is using and the one provided by the ontology builder to detect conflicts in the ontology [2]; or Metric based also referred to as Feature based – which gathers different types of statistics about the represented knowledge in an ontology [24]. Since, ONCL has not been deployed for real life use, we could not but adopt the Metric based approach for the Evolutionary and Logical based approaches are post deployment evaluation approaches. Examples of the Metric based approach include OntoMetric [19], AKTiveRank [1], oQual [12], OntoClean [16] and OntoQA [24].

Amongst these pre-deployment evaluation tools, OntoQA stands out because it works on populated ontologies thereby enabling it utilize knowledge represented in the instances to gain a better measure of quality of the ontology. Besides, it is simpler to use as compared to the other evaluation tools. The metrics in OntoQA are divided into categories; the category that addresses the design of the ontology schema known as the schema metrics and the category that addresses the way instances are organized in the ontology called the instance metrics [24]. The metrics are as highlighted below:

a) Schema Metrics
- Relationship Richness (RR): this metric reflects the diversity of the types of relations in the ontology. An ontology that contains only inheritance relationship conveys less information than that with diverse set of relationships. RR is given by the formula RR = |P| / |H| + |P|, where P is the number of non-inheritance relationship and H is the number of inheritance relationships.
- Inheritance Richness (IR): this is the distribution of information across different levels of the ontology's inheritance tree (fan-out of the parent classes). Ontology with a low IR shows a detailed covering of a specific domain and that with high IR shows a wide range of general knowledge with low level of detail. IR is given by IR = |H| / |C| where C is the number of classes.
- Attribute Richness (AR): this is an indicator for both the quality of ontology design and the amount of information pertaining to instance data. The more attributes (slots) that are defined, the more the knowledge the ontology is said to convey. AR is given by AR = |att| / |C| where att is the number of attributes for all classes.

b) Knowledgebase Metrics
- Class Richness (CR): this metric is an indicator of how instances are distributed across classes in the knowledgebase (KB) thereby giving an idea on how well the KB utilizes the knowledge modelled by the schema classes. High CR indicates that the data in the KB represents most of the knowledge in the schema while low CR indicates otherwise. The CR of a KB is the percentage of the number of non-empty classes (classes with instances) (C`) to the total number of classes defined in the ontology schema (C) that is to say, CR = |C`|/ |C|.
- Class Importance (CI): the importance of a class (Imp(C_i)) is the percentage of the number of instances that belong to the inheritance sub-tree rooted at C_i in the KB (Inst(C_i)) compared to the total number of class instances in the KB (CI) i.e. Imp(C_i) = |Inst(C_i)|/ |KB(CI)|. This class is important because it will help identify which area of the schema is in focus when instances are added to the KB.
- Cohesion (Coh): this is a measure of how connected the components (CC) of the graph representing the KB are. Coh = CC, where CC is the number of connected components.

These afore discussed metrics are critical to evaluating the alignment of ONCL with its focus of efficient and effective access and retrieval of case laws. Other metrics with no direct bearing with our ontology purpose in addition to these ones can also be found in [24].

3. EVALUATION OF THE ONTOLOGY FOR NIGERIAN CASE LAWS

As earlier stated, ontology evaluation is one necessary aspect of ontology development because it helps expose the quality of the ontology. After building ONCL, we found it pertinent to measure the quality of the ontology using a quantitative feature-based approach of [24] in order to expose the capability or potentials of the ONCL and expose means of improving it. We chose this metrics because it allows for evaluation at both design level and after development of the ontology, and also for simplicity sake. The scale used for the evaluation is as shown in table I and the values of the parameters used in the calculations and the results obtained are shown in table II.

Table I: Evaluation Scale.

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

Table II: Parameters and Formula.

Metric	H	C	C`	P	Att	C_i	CI	Formula (%)	Result
Relationship Richness	13	--	--	33	--	--	--	lPl/lHl+lPl	71.74
Inheritance Richness	13	46	--	--	--	--	--	lHl/lCl	28.26
Attribute Richness	--	46	--	--	23	--	--	lAttl/lCl	50.00
Class Richness	--	46	12	--	--	--	--	l C`l/lCl	26.09
Class Importance	--	--	--	--	--	14	42	Inst(C_i)/KB(CI)	33.33

The values of the parameters used in table II were obtained as follows; H, which is the number of inheritance class in the ONCL, was simply obtained by counting the number of leaf classes. P, which is the number of non-inheritance class, was obtained by subtracting the number of leaf classes from the total number of classes in ONCL.

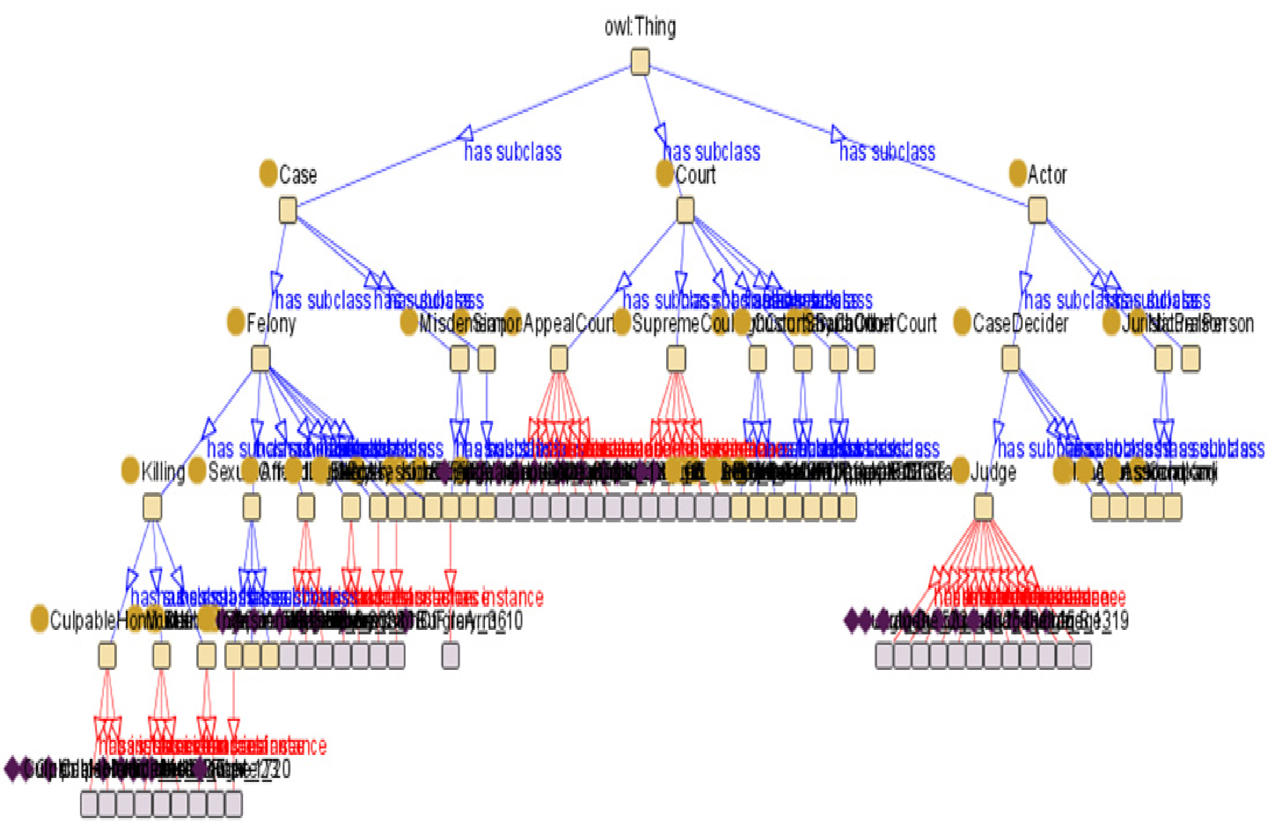

Figure 3.1: The class and individual tree for the ONCL

Att is the number of property (object, data type and annotation) defined in the ONCL. C is the number of classes defined in the ONCL. C` was obtained by counting the number of classes with instances defined. C_i is obtained at the level immediately after the root by taking the mean of the instances stemmed over that level and CI is obtained by counting the number of instances defined in the ONCL. We obtained all that from figure 3.1.

The interpretation of the results obtained in table II using the evaluation scale of table 1 is as depicted in table III. It is important to note that, the values obtained are based on the current state of the ONCL. As the ontology is extended, there is always a need for re-evaluation to give the true picture of the ontology in its current state.

Table III: Result of Evaluation.

Metric/Scale	1	2	3	4	5
Relationship Richness (RR)				✓	
Inheritance Richness (IR)		✓			
Attribute Richness (AR)			✓		
Class Richness (CR)		✓			
Class Importance (Imp(C_i))		✓			

3.1 Result Interpretation

The Relationship Richness (RR) is rated 4 because the result obtained after computation fell between [60-80]%, an indication that there are diverse relationships in the ONCL apart from inheritance relationships. This means that the amount of information conveyed by this ontology is high. The more diverse the relationships are, the more hit we are likely to make when accessing and retrieving case laws from the ONCL.

The Inheritance Relationship (IR) is rated 2, meaning low IR. This is an indication that the ONCL is deep, that is, it covers the specific area of our domain of concern in detail. This therefore implies that the result that will be returned on request will always be from the area in focus which in our situation is case laws. The Attribute Richness from our table is rated 3. This shows that the amount of knowledge conveyed by the ONCL is on the average. The higher the AR of the ONCL, the more efficient and effective the ONCL will be with regards to access and retrieval. ONCL's AR can simply be improved by defining more attributes in it.

The Class Richness (CR) of the ONCL is 2. Though this may be seen as low thereby giving an impression that the data in the KB does not represent most of the knowledge in the schema, this is true for the current state of the ONCL because less than half of the classes in the ONCL have instances defined for them. As instances are defined for the other classes in the ONCL, the CR will increase. The higher the CR of ONCL, the more efficient the ONCL will be in accessing case laws.

The Class Importance (Imp(C_i)) of the ONCL is rated 2. Since our Imp(C_i) was determined at the level immediately after the root, the instances in the ONCL is evenly distributed among three major concepts; an indication that these three areas are in focus as instances are added to the ontology; thus, access and retrieval of case laws are centred around these three concepts.

4. CONCLUSION

Using OntoQA we evaluated ONCL to determine the degree of alignment of this ontology with its purpose (i.e. the effective and efficient storage, access and retrieval of case laws). From the analyses foreshown, it is obvious that ONCL is average quality ontology in line with its purpose of effective and efficient access and retrieval of Case Laws and therefore needs to be improved upon. To do this, the AR should be improved by defining more properties in the ONCL, CR should be improved by creating more instances in ONCL. This process as exposed in this paper, can be repeated for subsequent versions of ONCL until a satisfactory level of purpose alignment is achieved.

REFERENCES

[1] Alani, H., Brewster, C. and Shadbolt, N. (2006), "Ranking Ontologies with AKTiveRank", The fifth International Semantic Web Conference, Athens, USA, 15pp.

[2] Arpinar, I. B., Giriloganathan, K. and Aleman-Meza, B. (2006), "Ontology Quality by Detection of Conflicts in Metadata", WWW 2006, Edinburg, UK, 7pp.

[3] Bench-Capon, T. J. and Visser, P. R. (1997), "Ontologies in Legal Information Systems: The Need for Explicit Specification of Domain Conceptualisations", ICAIL, Melbourne, pp. 132-141.

[4] Berneras, A., Laresgoiti, I. and Corera, J. (1996), "Building and Reusing Ontologies for Electrical Network Applications", In C. N. Caralt, "Modelling Legal Knowledge through Ontologies. OPKJ: the Ontology of Professional Judicial Knowledge", Ph.D. Thesis, Departament de Ciencia Polıtica i Dret Public, Universitat Autonoma De Bercelona, 527pp.

[5] Breuker, J., Elhag, A., Petkov, E. and Winkls, R. (2002), "Ontologies for Legal Information Serving and Knowledge Management", Legal Knowledge and Information Systems, Jurix '02: The fifteenth Annual Conference, Amsterdam, IOS Press, pp. 73-82.
[6] Breuker, J., Hoekstra, R. and Boer, A., Berg, K., Sartor, G., Rubino, R., Wyner, A. and Bench-Capon, T. (2007), "OWL Ontology of Basic Legal Concepts (LKIF-Core)", ESTRALLA, 138pp.
[7] Caralt, C. N. (2008), "Modelling Legal Knowledge through Ontologies. OPKJ: the Ontology of Professional Judicial Knowledge", Ph.D. Thesis, Departament de Ciencia Polıtica i Dret Public, Universitat Autonoma De Bercelona, 527pp.
[8] Condliffe, P., Abrahams, B. and Zeleznikow, J. (2010), "An OWL Ontology and Bayesian Network to Support Legal Reasoning in the Owners Cooperation Domain", available online at: http://www.abs.gov.au, 12pp.
[9] Ekuobase, G. O. (2004), "Scaling Process Models for Web Based application Development", M.Sc. Thesis, Department of Computer Science, University of Benin, Benin City, Nigeria, 79pp.
[10] Ekuobase, G. O., Fajuyigbe, O. and Onibere, E. A. (2009), "Process Model for Building Quality Software on Internet Time", International Journal of Pure and Applied Sciences, Vol. 5, No. 1, pp. 25-35.
[11] Ekuobase, G.O. (2006), "Software Creative Milestones", Proceedings of International Conference on Advances in engineering and Technology, Entebbe-Uganda, pp. 848-855.
[12] Gangemi, A., Catenacci, C., Ciaramita, M. and Lehmann, J. (2006), "Modelling Ontology Evaluation and Validation", ESWC 2006, LNCS 4011, pp. 140-154.
[13] Gómez-Pérez, A., Fernández-López, M. and Corcho, O. (2004), "Ontological Engineering", Grupo de Ontologías Laboratorio de Inteligencia Artificial Facultad de Informática Universidad Politécnica de Madrid, 109pp.
[14] Gruber, T. R. (1993), "A Translation Approach to Portable Ontology Specifications", Knowledge Acquisition, Vol. 43, pp 907-928.
[15] Gruninger, M. and Fox, M. S. (1995), "Methodology for the Design and Evaluation of Ontologies", Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, Montreal, 9pp.
[16] Guarino, N. and Welty, C. (2004), "An Overview of OntoClean", Handbook on Ontologies, Spring Verlag, pp. 154-159.
[17] Lassila, O. and McGuiness, D. (2001), "The Role of Frame-Based Representation on Semantic Web", Technical Reports KSL-01-02, Knowledge Systems Laboratory, Stanford University, USA.
[18] Lenat, D. B. and Guha, R. V. (1990), "Building Large Knowledge-Based Systems: Representation and Inference in the CYC Project", Addison-Wesley, Reading, Massachusetts, pp. 491-496.
[19] Lozano-Tello, A. and Gomez-Perez, A. (2004), "ONTOMETRIC: A Method to Choose the Appropriate Ontology", Journal of Database Management, Vol. 15, No.2, pp. 1-18.
[20] Noy, N. F. and McGuinness, D. L. (2001), "Ontology Development 101: A guide to Creating Your First Ontology", Stanford University, Stanford, CA, USA, 25pp.
[21] Sanchez, M. F. (2009), "Semantically Enhanced Information Retrieval: An Ontology-based Approach", Ph.D. Thesis, Escuela Politécnica Superior, Universidad Autonoma De Madrid, 254pp.
[22] Smith, B. (2003), "Ontology", In C. N. Caralt, "Modelling Legal Knowledge through Ontologies. OPKJ: the Ontology of Professional Judicial Knowledge", Ph.D. Thesis, Departament de Ciencia Polıtica i Dret Public, Universitat Autonoma De Bercelona, 527pp.
[23] Studer, R., Benjamins, V. R. and Fensel, D. (1998), "Knowledge Engineering: Principles and Methods", Data Knowledge Engineering, Vol. 25, No. 1, pp. 161-197.
[24] Tartir, S., Arpinar, I. B. and Sheth, A. P. (2007), "Ontological Evaluation and Validation", In R. Poli, M. Healy and A. Kameas (Eds), "Theory and Application of Ontology: Computer Application", Springer, Netherlands, pp. 115-130.
[25] Taveter, K., Lehtola, A., Jaaranen, K., Sorva, J. and Bounsaythip, C. (1999), "Ontology-Based Query Translation for Legislative Information Retrieval", Proceedings of 5th ERCIM Workshop, Dagstuhl, Germany pp. 47-58.
[26] Uschold, M. (1996), "Building Ontologies: Towards a Unified Methodology", Proceedings of Expert Systems, the 16th Annual Conference of British Computer Society Specialist Group on Expert Systems, 20pp.
[27] Uschold, M. and Jasper, R. (1999), "A Framework for Understanding and Classifying Ontology Applications", The Proceedings of the IJCAI-99 Workshop On Ontologies and Problem-Solving Methods (KRR5), Stockholm, Sweden, 12pp.
[28] Visser, P. R. and Bench-Capon, T. J. (1998), "A Comparison of Four Ontologies for the Design of Legal Knowledge Systems", Artificial Intelligence and Law, Kluwer Academic Publisher, pp.27-57.
[29] Yu, J., Thom, J. A. and Tam, A. (2007), "Ontology Evaluation Using Wikipedia Categories for Browsing", CIKM, Lisboa Portugal, pp. 223-232.