A Conceptual Model for Ontology Quality Assessment

A Systematic Review

Editor(s): Name Surname, University, Country
Solicited review(s): Name Surname, University, Country
Open review(s): Name Surname, University, Country

R.S.I.Wilson[a,b,*], J.S. Goonetillake[b] , W.A. Indika[c] and Athula Ginige[d]
[a]*Department of Computer Science and Informatics, Uva Wellassa University, Badulla, Sri Lanka*
[b]*Department of Information Systems Engineering, University of Colombo School of Computing, Colombo 07, Sri Lanka*
[c]*Department of Computer Science, University of Ruhuna, Matara, Sri Lanka*
[d]*School of Computer, Data and Mathematical Sciences, Western Sydney University, NSW, Australia*

Abstract. With the continuous advancement of methods, tools, and techniques in ontology development, ontologies have emerged in various fields such as machine learning, robotics, biomedical informatics, agricultural informatics, crowdsourcing, database management, and the internet of things. Nevertheless, the nonexistence of a universally agreed methodology for specifying and evaluating the quality of an ontology hinders the success of ontology-enabled systems in such fields as the quality of each component is required for the overall quality of a system and in turn impact the usability in use. Moreover, a number of anomalies in definitions of ontology quality concepts are visible, and in addition to that, the ontology quality assessment is limited only to a certain set of characteristics in practice even though some other significant characteristics have to be considered for the specified task. Thus, in this research, a comprehensive analysis was performed to uncover the existing contributions specifically on ontology quality models, characteristics, and the associated measures of these characteristics. Consequently, the characteristics identified through this review were classified with the associated aspects of the ontology evaluation space. Furthermore, the formalized definitions for each quality characteristic are provided through this study from the ontological perspective based on the accepted theories and standards. Additionally, a thorough analysis on the extent to which the existing works have covered the quality evaluation aspects is presented and the areas further to be investigated are outlined.

Keywords: ontology quality, ontology quality model, ontology quality characteristics, systematic review

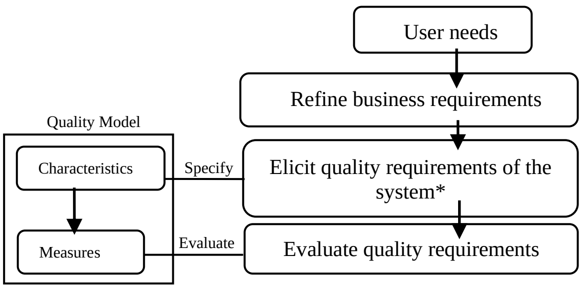
1. Introduction

Ontology is a specified conceptualization of the world that is represented for some purpose [91]. More specifically, an ontology has been defined as "a formal, explicit specification of a shared conceptualization" in [81] merging the two prominent definitions of Gruber [91]: "ontology is an explicit specification of a conceptualization" and Borster [94]: "an ontology is a formal specification of a shared conceptualization". In which the conceptualization is an abstract and simplified view of the world that consists of concepts that are expected in the world being represented, and relationships among them [87] also named as a representational vocabulary in [71,91]. A formal specification is machine-interpretability in a way that a machine can understand and process the specified conceptualization. This would enable ontology is to be used for several purposes such as to make meaningful communication among their agents/users, facilitate interoperability, knowledge sharing and reuse among subsystems and, to distin-

[*]Corresponding author. E-mail: shyama@uwu.ac.lk

guish the domain knowledge from operational knowledge. These can be viewed as capabilities of ontologies [64,71] and consequently, ontologies are increasingly incorporated in information systems. However, the above capabilities would not be a reality unless the good quality ontology is produced at the end of the development. To this end, it would be beneficial if a universally accepted methodology for the quality evaluation of ontologies is available, however not so far, as the way it is in software engineering. For instance, the availability of widely agreed quality standards in software engineering such as ISO/IEC 9126 [40], ISO/IEC 9241-11 [38], ISO/IEC 25012[37], ISO/IEC 25010 [39], and IEEE software quality standards [36] provide immense support in developing good quality software.

Moreover, one trivial internal quality issue would cause multiple quality issues in the end product which may result in a high cost of debugging and fixing the issues and in turn loss of consumers' trust. For instance, it was revealed that the ontology-enabled system in the medicine domain has caused 55% loss of information due to one missing explicit definition in the ontology [8, 9, 13, 22, 77]. Thus, it is required to sufficiently evaluate both the quality of the ontology content during the development, as well as, the ontology as a whole after the development [8]. This has been viewed as two aspects of ontology evaluation in [84] namely intrinsic and extrinsic that have been further discussed in Section 3.



*Based on the overall system quality requirements, further quality requirements of each component of the system can be derived.

Fig. 1. Abstract process of specifying quality requirements

Furthermore, the quality requirements to be assessed under each aspect (i.e., intrinsic and extrinsic) can be derived from the business requirements also called users' needs in [69,12]. Then, the derived quality requirements with respect to the aspects can be evaluated to ensure whether the required quality is being achieved across the ontology development. To this end, a quality model, that consists of a proper set of quality characteristics and associated measures, provides the basis for specifying the quality requirements and in turn for evaluating the quality requirements of an ontology [37]. Figure 1 shows the abstract process of deriving the quality requirements and how a quality model is associated with that process. The more fine-grained approach is elaborated in Section 3.

There are countable quality models that have been proposed for ontology quality assessments. However, none of them are widely accepted in practice due to some of them are generic such as the semiotic metric suit [2], the quality model of Gangemi et al [5], and OQuaRE [4], thus, it is required to additional efforts to customize and add absent characteristics which are significant to the specified context. Moreover, some of the quality models are specific to the context such as OntoQualitas,[63] and the quality model of Zhu et al. [35], Therefore, these models do not guarantee work well for other contexts rather than its specified context. Thus, the researchers and practitioners face the difficulties as mentioned follows due to the non-existence of an agreed quality model;

(i.) Difficulty in determining the required set of quality characteristics and in turn the relevant quality measures to achieve the specified quality requirements [72].
(ii.) Difficulty in differentiating the quality characteristics and quality measures as their definitions in the literature are vague and the terminologies have been used interchangeably [82]. This also has a negative impact on (i.) above.
(iii.) Inadequate knowledge in assessing the complete set of quality characteristics applicable for an ontology since the ontology quality assessment in practice has got limited only to a certain set of characteristics irrespective of many other characteristics which would become applicable and have been theoretically proposed [3].

Thus, there is in fact a requirement to perform a thorough analysis on the domain of ontology quality assessment. As a result, a systematic review was performed to streamline the findings of the existing approaches and thereby producing a conceptual quality model that is useful for researchers and practitioners to understand the characteristics, attributes, and measures to be considered in different aspects of ontology quality assessment.

The rest of the article is structured as follows; Section 2 presents the survey methods that we have followed. Section 3 provides preliminaries and conceptualization of ontology quality assessment to understand the concepts that have been described in the rest of the sections. The overview of the exiting ontology quality models is presented in Section 4. In Section 5, the quality characteristics, associated quality measures, and definitions for each quality characteristic are clearly defined. Furthermore, the relations between the characteristics and comparison of the existing approaches are also discussed in this section. Finally, the survey is concluded with highlights of future works.

2. Survey Methodology

Firstly, a traditional theoretical review was performed by retrieving the related ontology quality surveys and the literature reviews, to explore the ontology quality evaluation and then, to identify the possible quality models, including significant ontology quality characteristics and measures, and the relations among the characteristics. Few attempts have been carried out to present such contributions [11, 45, 48-49,79]. However, they are not comprehensive surveys and the discussions were limited to a set of characteristics proposed in [5,9-10,29,92]. In addition to that, few comprehensive surveys on ontology evaluation have been identified (see Table 1) in which the main focus is on ontology evaluation in broader aspects, not specifically for ontology quality problems.

Table 1

The existing survey studies in ontology evaluation

Article	Description
Vrandecic, 2010 [29]	**Databases**: Not specified **Results**: Presented the overview of domain- and task-independent evaluation of an ontology by emphasizing the quality assessment related to the aspects: vocabulary, syntax, structure, semantics, representation, and context.
Gurk et al, 2017 [85]	**Databases**: Three databases: ScienceDirect, IEEE Xplore Digital Library, ACM Digital Library (the selected articles have not been mentioned) **Results**: Performed a survey to find the ontology quality metrics for a set of characteristics that are defined in ISO/IEC 25012 Data Quality Standard (This work is limited only to the inherent quality and inherent-system quality)
Degbelo, 2017 [3]	**Databases**: the Semantic Web Journal, the Journal of Web Semantics (2003-2017) **Results**: Performed the review on quality criteria and strategies which have been used in the design and implementation stages of the ontology development. Presented the gaps between theory and practice of ontology evaluation.
McDaniel and Storey, 2019 [61]	**Databases**: Web of Science journals, approximately 170 articles **Results**: Performed the review on the evaluation of domain ontology. Classified and discussed the existing ontology evaluation research studies under the five classes: Domain/Task fit, Error-checking, Libraries, Metrics, Modularization.

We conducted the systematic review by following the methodologies described in [17,74] with the objective of addressing the issues (i, ii, and iii) stated in Section 1. Thus, the review includes a summarization of ontology quality problems and models: recognition of characteristics and measures. Eventually, the aim is to present a conceptual model that provides a basis for researchers to retrieve the quality characteristics upon the quality requirements. To achieve these objectives, the following research questions were derived;

- What ontology quality models are proposed for ontology quality assessment?
- What quality problems are discussed in the previous approaches with respect to ontology quality?
- What are the ontology quality characteristics and measures assessed in the previous approaches?

Then, the search terms were identified based on the research questions such as *quality*, *ontology*, *ontology quality, assessment*, *evaluation*, *approach*, *criteria*, *measures*, *metrics*, *attributes*, *characteristics*, *methods, and methodology,* and trial searches were executed using various combinations of the terms. Thereafter, the following combination was used to retrieve the relevant papers from the selected digital databases, journals, and search engines (see Figure 2).

- [ontology AND [Quality OR Evaluation OR Assessment]
- [[Ontology AND Quality] AND [Criteria OR Measures OR Metric OR Characteristics OR Attributes]
- [Ontology AND Quality] AND [Models OR Approach OR Methodology]]

These search combinations were performed according to the instructions given in each digital database under the advanced search. The following inclusion and exclusion criteria were defined to select the candidate studies at the screening and eligibility phases (see Figure 2).

- Inclusion Criteria: *Studies that are in English, published during the period (2010 – 2021), were peer-reviewed. studies discussed ontology quality assessment, empirical or theoretical studies.*
- Exclusion Criteria: *Studies published as a short paper (less than 6 pages), tutorial, poster, and report. Studies did not present the rigorous approach to achieve the defined quality assessment*

objectives. Studies do not directly relevant to the research questions

During the screening phase, the candidate studies were retrieved by performing the keywords search on title and abstract and eliminated the duplicates using the reference management tool (i.e., Mendeley). Then, the abstracts were reviewed to filter the relevant studies. There are some cases wherein the abstract is vague and difficult to understand the real contribution. In this case, the papers were forwarded to the second round for further review of the introduction and conclusion as suggested in [74]. Altogether, 128 papers were forwarded for the full paper review. During this stage, forward and backward citations were searched and included the appropriate studies to report. It has been identified that some papers have two versions: journal and conference entitled to same authors. In such a situation, the journal paper was selected. Finally, thirty (30) papers mentioned in APPENDIX A were selected for the reporting phase.

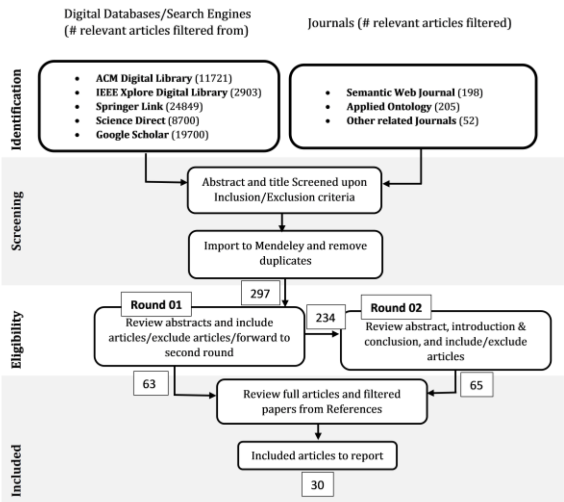


Fig. 2. Main steps of the review process and number of articles retrieved

3. Preliminaries and conceptualization

In the context of quality, a number of ad hoc definitions and inconsistent terminologies appear in the literature leading to terminology misapplications and misinterpretations [23]. Thereby, in this section, the concepts and the terms that are relevant to our discussion are briefed to avoid miscommunication and thus to maintain consistency. Additionally, the conceptualization of ontology quality assessment is provided based on the existing theories.

Quality Model. Quality models provide the basis for specifying quality requirements and evaluating the quality of an entity (i.e., software, tools, ontology, part of the software) [37,39]. In software engineering, quality models are twofold, relational and hierarchical [7]. The relational model presents the correlation among the quality characteristics and this type of model has not been proposed for the ontology so far in the existing works. Only, the hierarchical models can be found which have usually four levels (see Figure 3). The top-level (i.e., the first level) consists of a set of characteristics that are further decomposed into sub-characteristics at the next level (i.e., the second level). The third level would consist of associated attributes of characteristics/sub-characteristics. These attributes have measures, which are lay at the bottom level (i.e., the fourth level), that can be used to assess the entity either quantitively or qualitatively. Thus;

Characteristics (i.e., Criteria): describe a set of attributes. An attribute is a measurable physical or abstract property of an entity [18]. In the ontology quality context, an entity can be a set of concepts, properties, or an ontology [30].

Measure (i.e., metrics): describes an attribute formally and assesses them either quantitively or qualitatively [93].

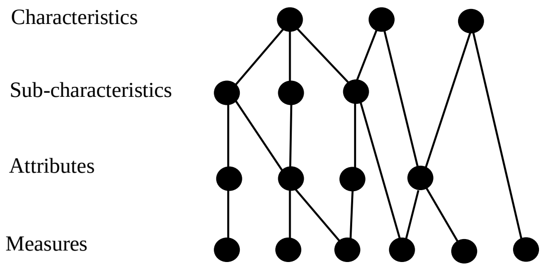


Fig. 3. Association between Characteristics, Attributes, and Measures

There are no definite requirements to have sub-characteristics for each characteristic or else characteristics/sub-characteristics should have attributes. Thereby, a characteristic can directly be related to measures without a set of attributes (see Figure 3). Moreover, an attribute/a measure can be related to many characteristics. For instance, *cognitive complexity* is one of the ontology characteristics that is associated with several attributes such as *size*, *depth*, *breadth*, and *fan-outness* [6] and each attribute has a set of measures. If the size attribute is considered, the related measures are the number of classes, number of relations, and number of instances. Thus, the associated levels of the quality models are <characteristics, attributes, measures>.

Moreover, the mentioned measures can also be used to assess the cohesion attribute of the modularity characteristic [35,89]. If we consider the consistency characteristic of an ontology, it can be further derived into two sub-characteristics namely, internal consistency and external consistency [10,35]. Internal consistency contains direct measures such as circularity errors, the number of multi-parent classes, and the number of missing disjointness [10]. Thus, the associated levels are <characteristics, sub-characteristics, measures>. However, external consistency can be measured through attributes interpretability, clarity, precision [53,63], and each attribute has a set of measures that have been elaborated in Table 6. Thus, it covers levels <characteristics, sub-characteristics, attributes, measures>.

Moreover, sub-characteristics also can be further derived into another set of sub-characteristics also called primitive characteristics [15] as the way it is in OQuaRE [4]. Additionally, it can be observed that some of the quality models consist of another higher-level called dimensions or factors on top of the characteristics-level that have been proposed to classified the characteristics/attributes which are related to similar aspects of the ontology. For instance, the quality model of Gangemi et al. [5] contains three dimensions namely structural, functional and usability-related. Similarly, the quality model of Zhu et al. [35] has classified the characteristics under the content, presentation, and usage dimensions. However, there is no clear definition is given for the dimension in literature, thus, the term: dimension has been used interchangeably to define the term characteristics. For instance, in [14,83], the dimension is defined as a characteristic. However, in our study, we use the terms in the order: dimensions, characteristics, sub-characteristics, attributes, and measures to describe the structure of a quality model whenever it is necessary.

Ontology Layers (i.e., Levels). The researchers have discussed the ontology quality layer-wise, or level-wise by concerning an ontology as a multi-layered vocabulary [9,29,42,48-49]. Initially, three layers to be focused on have been proposed in [9], namely: content, syntactic & lexicon, and architecture. Later, this was expanded by including the layers: structural and context [29,42].

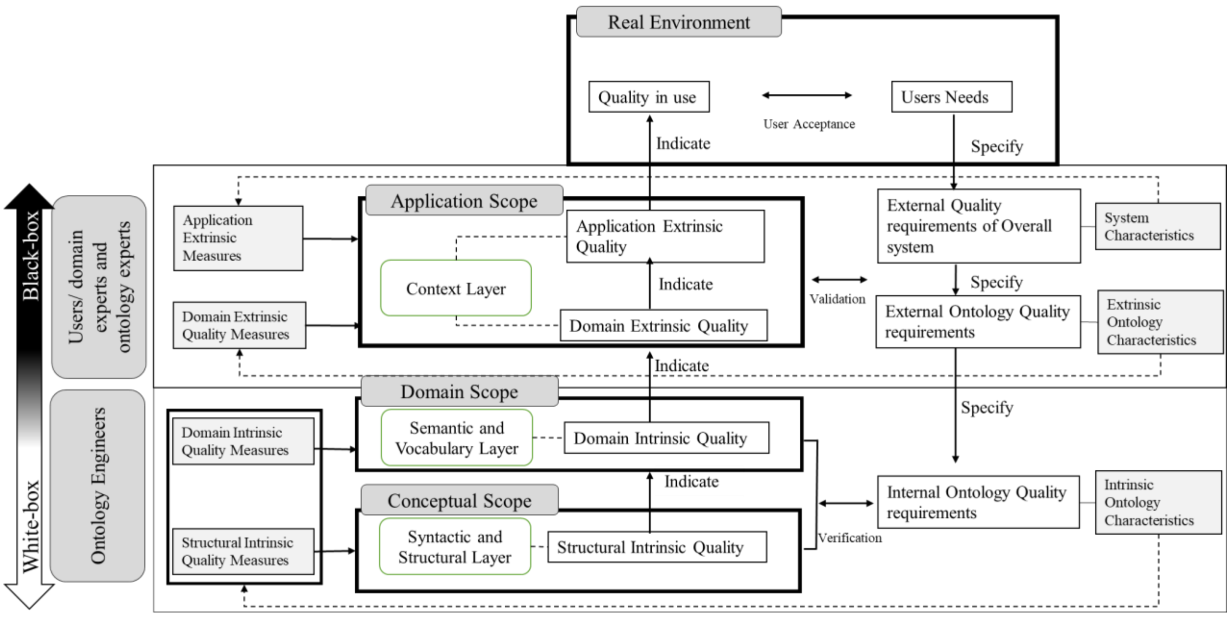


Fig. 4. Conceptualization of ontology quality evaluation

The syntactic layer considers the properties related to the language that is used to represent the ontology formally. The structural layer/architectural layer focuses on the *is-a* relationship (i.e., hierarchical relationships) which is more important in the ontology modeling against the other relations. Moreover, it considers the design principles and other structural features required to represent the ontology. The lexicon layer is also named as vocabulary or data layer [29,42], which takes into account the vocabulary that is used to describe concepts, properties, instances, and facts. The non-hierarchical relationships and semantic elements are considered under the semantic layer. To evaluate the vocabulary and semantic layer, some understanding of domain knowledge is required. The context layer concerns the application scope that the ontology is built for. It is important to assess whether the ontology confirms the real application requirements as a component of an information system or a part of a collection of ontologies [42]. We have performed a separate analysis on the ontology evaluation layers and its relationship between the evaluation approaches (i.e., methods) such as *task-based, human-based, data-driven,* and *golden-based* that can be found in [82] for more detail.

Ontology evaluation Space. A system would normally consist of many components. The quality of each component of a system would influence the overall quality of a system that is often critical for business success. Thereby, the quality requirements of each component should be traced from business requirements that reflect the needs of users (Figure 1). This can be viewed as a connected flow starting from users' needs to the internal quality of a product. For instance, the users' needs can be defined as a set of requirements to be achieved by a system when it is in use. This is also defined as quality in use [37,69]. Then, this set of requirements determine the external quality requirements of each component of a system, accordingly, determines the internal quality requirements. These quality requirements can be expressed as a set of characteristics with associated measures, then, these characteristics can be evaluated to ensure whether the desired quality in each stage is achieved (see Figure 4).

For an ontology-enable system, ontology quality requirements also need to be distinctly identified from the overall quality requirements of the system [12]. Based on that, the external and internal quality requirements to be achieved from the ontology can be derived. This has been discussed under the aspects of ontology evaluation space in [32,72] namely extrinsic and intrinsic aspects of ontologies. Thus, Figure 4 presents the flow of specifying the ontology quality requirements with the associated evaluation aspects when an ontology is a component of a system. It should be noted that the quality requirements related to the other components of the ontology-enabled system have not been elaborated in Figure 4.

The extrinsic evaluation focuses on two types of aspects namely domain extrinsic and application extrinsic. In these aspects, ontology quality is evaluated as a component of a system and after deploying the ontology in an application respectively. This has been considered as context-layer evaluation in [29, 42]. Thus, the ontology quality requirements elicited from the business requirements of the overall system are evaluated under these aspects [12,65]. Traditionally this evaluation is called "black-box" or "Task-based" testing [32,46] as the quality assessment is performed without peering to the ontology model also defined as ontology validation in [87]. Thus, ontology is assessed to confirm whether the right ontology was built [29,87], mainly this is performed with the involvement of users and domain experts.

The intrinsic evaluation focuses on the ontology content quality to be achieved concerning the intended use of the system, also that would indicate the extrinsic quality. To this point, ontology is considered as an isolated component separated from the system. At this stage, ontology quality can be evaluated under two aspects: domain intrinsic and structural intrinsic [32] mainly done by ontology engineers. This is similar to "white-box" testing in software engineering [46]. In the structural intrinsic aspect, the focus is given to the syntactic and structural requirements (i.e., syntactic and structural layers) which involve the specified conceptualization such as language compliance, conceptual complexity, and logical consistency. In the domain intrinsic aspect, ontology quality is evaluated with reference to the domain knowledge that is required for the intended purpose of the system [32]. Thus, it would consider the semantic and vocabulary layers of the ontology. Through the intrinsic aspects, verification is being done to ensure whether the ontology is built in the right way [29,87].

In each aspect of the ontology evaluation space mentioned so far can be evaluated using a set of characteristics together with associated measures, thus, it ensures whether the defined quality requirements are being achieved. The characteristics and related measures in each aspect are further elaborated in Section 5.

4. Ontology Quality Models

4.1. Quality models in Software Engineering

Many well-accepted quality models are available in software engineering such as McCall's quality model [43]; Boehm's quality model [15]; Garvin's quality model [26], the data quality model: ISO/IEC 25012[37], and the ISO 9126 quality model [40].

McCall's quality model has proposed eleven quality factors (i.e., can also be considered as dimensions) from the product perspectives which further have been divided into a set of characteristics and measures. Thus, it provides a hierarchal model containing the three levels. Boehm's quality model is a hierarchical model that is similar to McCall's model. However, it provides a broader range of characteristics without defining the measures. The top-level contains the characteristics: *As-is utility, Maintainability, Portability* based on that, the intermediate and low levels describe the related sub-characteristics and primitive characteristics respectively.

Garvin's quality model has described eight quality dimensions (i.e., factors) specifically for the product quality which can be adapted to software engineering. Moreover, the dimensions have been defined at the abstract level, and some of them are subjective, thus it is difficult to measure.

The data quality model ISO/IEC 25012 has defined fifteen characteristics that have to be taken into account when assessing the quality of data products is built on [37]. The characteristics further have been categorized into two types namely inherent data quality and system-dependent data quality. The inherent data quality describes the characteristics of data such as accuracy, completeness, consistency, credibility, and currentness that have the intrinsic potential to satisfy stated and implied needs. The rest of the characteristics such as accessibility, compliance, confidentiality, efficiency, precision, traceability, understandability, availability, portability, and recoverability, come under the system-dependent data quality that must be preserved within a computer system when data is used under specified conditions.

ISO 9126 model (ISO/IEC 9126:1991) is a widely accepted software product quality model that was produced in 1991 by the International Organization for Standardization (ISO). Later, it has been extended as ISO/IEC 9126-1:2001 that includes four parts:

- Part 1: Quality model
- Part 2: External Measures
- Part 3: Internal Measures
- Part 4: Quality in use Measures

Part 1: the quality model comprises with six characteristics related to the internal and external software product quality which have further been decomposed into sub-characteristics. The measures of each characteristic/sub-characteristic have been defined under Parts: 2, 3, and 4 of the standard.

Furthermore, SQuaRE (**S**ystems and software **Qu**ality **R**equirements and **E**valuation) has been introduced by redesigning the ISO 9126 model with the ISO 14598 series of standards because the ISO 9126 model contains issues, as a result of the advancement of the information technologies and the changes in its environment [40]. SQuaRE comprises of the same set of characteristics that were defined in ISO 9126 with several amendments as described in [40].

From the ontological perspective, there are no agreed models as the way it is in software engineering. However, significant contributions have been made to the field that has been elaborated on in the subsequent section.

4.2. Quality models for ontologies

Initially, a framework has been proposed in [9] to verify that developers are building a correct ontology with a set of characteristics namely *soundness, correctness, consistency, completeness, conciseness, expandability,* and *sensitiveness,* which have been adopted in later research works and developments [29,51,61,63]. Thereafter, significant methods [64,70-71] and tools such as OntoTrack [67], OntoQA [89], OntoMetric [19], SWOOP [67] have been proposed enabling quality assessment with several sets of characteristics mainly focusing on the intrinsic aspect. Afterward, several attempts have been taken to provide a generalized quality model, significantly, the semiotic metric suit [2], the quality model of Gangemi [5], and OQuaRE [4].

Semiotic metric suite [2]: Semiotic theory [80] has taken as the foundation to this model as the ontology has a semiotic nature. It describes "the study of the interpretations of signs and norms" with six levels namely syntactic, semantic, pragmatic, social, physical, and empiric [80]. In the semiotic metric suite, a set of attributes have been classified related to the ontology quality evaluation under the above aspects excluding physical and empiric levels, that can also be considered as dimensions as follows;

- **Syntactic**: Lawfulness, Richness
- **Semantic**: Interpretability, Consistency, Clarity

– **Pragmatic**: Comprehensiveness, Accuracy, Relevancy
– **Social**: Authority, History

Even though the authors have mentioned that this model has been constructed solely by focusing on the intrinsic aspect, the attributes: *accuracy, relevancy,* and all attributes in the *social* dimension take an extrinsic nature. To the reason that, under the pragmatic and social levels, the ontology is assessed as a whole and measure based on the external factors such as domain experts' knowledge [54], and external documents: usage logs, page ranking, details of external links with other ontologies [61].

The quality model of Gangemi et al [5]: by considering an ontology as a semiotic object "*including graph objects, formal semantic spaces, conceptualizations, and annotation profiles*", the three main dimensions: *structural*, *functional*, and *usability-related* have been proposed. The *structural* dimension consists of thirty-two measures related to the topological, logical, and meta-logical characteristics of an ontology [6]. Primarily, it covers syntactic and formal semantic. Under the *functional* dimension, the possible attributes: *precision*, *Recall (i.e., coverage*), and *accuracy* have been proposed to assess the conceptualization specified by the ontology with respect to the intended use. The measures of the *usability-related* dimension focus on the metadata (i.e., annotations) about the ontology and its elements related to the communication context of an ontology. To this end the attributes: *presence, amount, completeness,* and *reliability* have been described under three levels: *recognition* annotation, *efficiency* annotation, and *interfacing* annotation.

OQuaRE [4]: By adopting the standard SQuaRE (ISO/IEC 25000:2005) [39], a quality model has been proposed to evaluate the ontology quality, that comprises the same set of characteristics defined in the standard ISO/IEC 25000:2005 such as *functional adequacy, reliability, operability, maintainability, compatibility,* and *transferability*. In addition to that, the *structural* characteristic has been included in the model to assess the inherent topological characteristics of an ontology. Furthermore, these characteristics have been decomposed into several sub-characteristics considering the ontological point of view including a set of associated measures which are available online[1].

Based on our survey, OQuaRE [4] has not appeared in the later research. The reason would be the proposed sub-characteristics are subjective and difficult to apply in practice. Moreover, it has not focused on the semantic nature of the ontology that is an essential component when processing meaningful interpretations. [29,35,63]. However, noticeable studies are available related to the other two models: semiotic metric suit[2] and the quality model of Gangemi et al [5].

OntoKeeper [54] and DoORS [60] are semiotic-metric-suite-driven initiatives, in which OntoKeeper has automated the metric suite by taking into account only the intrinsic aspect. DoORS is a web-based tool for evaluating and ranking ontologies. It has been developed by extending the semiotic metric suite [2] with a set of characteristics: *structure*, *precision* (i.e., instead of clarity), *adaptability, ease of use,* and *recognition*.

The quality model of Gangemi et al. [5] has been adopted in multiple studies [29,62,78,90]. OOPS! has used the model to classify the proposed common pitfalls which can occur during ontology development. Furthermore, the dimensions have been extended to another level (see Table 2). The authors in [90] have used the structural measures defined in the model of Gangemi et al. [5] to evaluate the ontology cognitive ergonomic. Moreover, a theoretical model proposed in [29] contains eight characteristics: *Accuracy, Adaptability, Clarity, Completeness, Computational efficiency, Conciseness, Consistency,* and *Organizational fitness*, which have been derived through literature reviews including the set of characteristics proposed in Gangemi et al's list. The quality model proposed in the ONTO-EVO^{A}L approach has been constructed mainly based on Gangemi et al's model that contains two dimensions: content and usage with six characteristics (see Table 2) [78].

In addition to that, few efforts have been made in designing quality models for a specific purpose such as OntoQualitas [63], the quality model for semantic descriptions of web services [35], and SemQuaRE [33]. OntoQualitas provides a set of characteristics and related measures to evaluate the quality of ontologies which are built upon the purpose of exchanging information between heterogeneous systems. The ROMEO methodology [51] is the basis that OntoQualitas follows to derive the measures relating to the quality characteristics. The applicability of the characteristics to the intended context was empirically evaluated. Focusing on the context of semantic descriptions of web services, the authors in [35] have presented a quality model which comprises the four levels namely: *aspects, attribute, factors,* and *metrics* (Table 2). This can be viewed as dimensions, charac-

[1] http://miuras.inf.um.es/evaluation/oquare/Metrics.html

teristics, sub-characteristics/attributes, and measures as per our definitions in section 3.

SemQuaRE [33] quality model has not specifically been defined for ontologies; however, it is for assessing the quality of semantic technologies. For instance, *ontology engineering tools*, *Ontology matching tools*, *Reasoning systems*, *Semantic search tools*, *and Semantic web services*. Notably, this model has been constructed following the bottom-up approach, Thus, the model is constructed by analyzing the measures that were derived in the existing studies and mapped them until reaching the top-level characteristics. All of the other models so far discussed have followed a top-down approach, in which the characteristics are determined based on the ontological theories, from that, the possible sub-characteristics and measures are derived.

The quality models for ontologies that we have explored through the survey are summarized in Table 2. All of them have a hierarchical structure and there is no significant work on constructing a relational (i.e., non-hierarchical) model that shows the correlation between characteristics. In Table 3, the characteristics/attributes proposed in the ontology quality models in [35,60,63,78] have been mapped with the respective aspects (i.e., *structural intrinsic, domain intrinsic, extrinsic (domain/application)*, and *quality in use*).

Table 2

The existing quality models for ontologies identified through the survey

Model	Description	Dimensions	Characteristics and Sub-Characteristics/attributes

Quality model in ONTO-EVO^{A}L [78], 2010	**Structure**: Hierarchical model
	Base model: No specific model is used
	Purpose: To assess the quality of evolving ontology
	Approach: Top-down approach: the certain characteristics have been derived from [6,29,89]
	Evaluation: Empirical evaluation has not been presented.
		Content	Complexity
			Cohesion
			Conceptualization
			    Semantic Richness
			    Attribute Richness
			    Inheritance Richness
			Abstraction

		Usage	Completeness
			    Precision
			    Recall
			Comprehension

OQuaRE [4], 2011	**Structure**: Hierarchical model **Base model**: The SQuaRE standard **Purpose**: To rank, select, compare and assess the ontologies. **Approach**: Top-down approach: based on the ontological theories **Evaluation**: Empirical evaluation has been performed on two applications: Ontologies of Units of Measurement[2]and Bio ontologies[3]	-	Structural Functional Adequacy Reliability Operability Maintainability Compatibility Transferability (More sub-characteristics of these are available online[4])

OntoQualitas,[63], 2014	**Structure**: Hierarchical model
	Base model: No specific model is used
	Purposed: To evaluate the quality of an ontology whose purpose is the information interchanges between heterogeneous systems
	Approach: Top-down approach: adopted the ROMEO methodology [51], then the criteria have been derived from [2,10,70]
	Evaluation: Empirical evaluation has been performed on ontologies of enterprises interchange Electronic Business Documents,
		-	Language conformance
			Completeness
			Conciseness
			Correctness
			    Syntactic Correctness
			    Semantic Correctness
			    Representation Correctness
			Usefulness

Quality model in OOPS! [62], 2014	**Structure**: Hierarchical model **Base model**: the model of Gangemi et al. [5] **Purpose**: To classify the identified pitfalls **Approach**: Top-down approach **Evaluation:** the model of Gangemi et al. [5]	Structural	Correctness Modeling Completeness Ontology Language conformance
		Functional	Requirement Completeness Content Adequacy

[2] http://miuras.inf.um.es/evaluation/oquare/UOM/ContMetricsUOM.html
[3] http://miuras.inf.um.es/evaluation/oquare/CTO/MetricsBiOntology.html
[4] http://miuras.inf.um.es/evaluation/oquare/Contenido.html

	has been extended to classify the pitfalls, thus, no evaluation has been provided.	Usability-related	Ontology Understanding Ontology Clarity
Quality model of Zhu et al. [35], 2017	**Structure**: Hierarchical model **Base model**: No specific model is used **Purpose**: To evaluate ontology for Semantic Descriptions of Web Services **Approach**: Top-down approach **Evaluation:** Empirical evaluation has been performed on the five weather ontologies	Content	Correctness Internal Consistency External Consistency Compatibility Completeness Syntactic Completeness Semantic Completeness
		Presentation	Well-formedness Conciseness Non-redundancy Structural Complexity Size Relation Modularity Cohesion Coupling
		Usage	Applicability Definability Description Complexity Adaptability Tailorability Composability Extendibility Transformability Efficiency Search Efficiency Composition Efficiency Invocation Efficiency Comprehensibility
Quality model of McDaniel et al. [60], 2018	**Structure**: Hierarchical model **Base model**: The semiotic metric suit **Purpose**: To rank ontologies **Approach**: Top-down approach **Evaluation:** empirical evaluation has been performed by selecting ontologies from the Bio Portal ontology repository	Syntactic	Lawfulness Richness Structure
		Semantic	Consistency Interpretability Precision
		Pragmatic	Accuracy Adaptability Comprehensiveness Ease of use Relevance
		Social	Authority History Recognition

However, the models: OQuaRE [4] and the quality model in OOPS! [62] has not been included in Table 3. To the reason that the OQuaRE [4] characteristics have been defined by considering the ontology as a software artifact, thus it is difficult to distinctly mapped with the ontological evaluation aspects and OOPS has adapted an existing model to classify a set of pitfalls, and not specifically provide a set of characteristics concerning the quality requirements.

OntoQualitas and the model in ONTOEVO^{A}L have the characteristics related to the intrinsic extent, in which the characteristics: *completeness* (coverage), *conciseness* (Precision), *representational correctness*, are attached with the domain that the ontology considered, thus for the evaluation where some understanding of the domain is needed.

The quality model of Zhu et al. model [35] concerns both intrinsic and extrinsic aspects. However, the sub-characteristics of *applicability* and *efficiency* have been specifically defined for the semantic web services context. When comparing these models, the characteristics proposed in the quality model of McDaniel et al. [60] cover many aspects of the ontology quality evaluation. In which, *authority*, *history*, and *recognition* reflect user satisfaction. For instance, *authority* considers the number of linkages with other ontologies, *history* considers the number of revisions made and how long it is an activity available publicly, *recognition* considers the number of times the ontology is downloaded and the reviews given on to the ontology. Thus, these attributes are useful to understand to what extent the ontology is accepted by the community, then the positive values of the attributes imply the user satisfaction of the ontology.

Table 3

The characteristics/attributes of the existing ontology quality models mapping with the evaluation aspects

Model	Structure Intrinsic	Domain Intrinsic	Extrinsic (Domain/ Application)	Quality in use
Quality model in ONTO-EVO^{A}L [78], 2010	Complexity Cohesion Conceptualization Abstraction Comprehension	Completeness (Precision and recall)		
OntoQualitas, [63], 2014	Language conformance Completeness Conciseness Syntactic Correctness Semantic Correctness	Completeness: Coverage Conciseness: Precision Representation correctness	Usefulness	
Quality model of Zhu et al. [35], 2017	Internal consistency Well-formedness Structural Complexity Modularity	External Consistency Compatibility Completeness Conciseness	Applicability Adaptability Efficiency Comprehensibility	
Quality model of McDaniel et al. [60], 2018	Lawfulness Richness Structure	Consistency Interpretability Precision Comprehensiveness	Accuracy Adaptability Ease of use Relevance	Authority History Recognition

In addition to that, by adopting the GQM (Goal-Question-Metrics) methodology [93], the study [51] employed in providing approaches to derive the measures of quality characteristics tracing from ontology requirements. In which, goals are the ontology requirements or roles which are gradually refined into questions/sub-questions which reflect the respective quality characteristics to be measured. To this end, quality models act as a complementary component that supports deriving measures with respect to the characteristics reflected from each question. Thus, the proposed approaches would not be effective without a quality model that presents a set of characteristics and corresponding measures.

5. Classification of ontology quality characteristics

A number of characteristics and measures have been discussed in the selected papers. From that, thirteen (13) significant characteristics: *compliance, complexity, internal consistency, modularity, conciseness, coverage, external consistency, accuracy, relevancy, functional completeness, comprehensibility, adaptability, efficiency*, were identified and grouped under the four evaluation aspects also can be viewed as dimensions namely: *Structural intrinsic*, *Domain intrinsic*, *Domain extrinsic* and *Application extrinsic* which were derived based on the ontology evaluation space as described in Section 3. Moreover, this model has the same nature as ISO/IEC 25012-Data Quality model [37], nevertheless it consists of two main categories, *inherent data quality*, and *system-dependent data quality*. However, for ontology quality, three main categories were defined: *inherent ontology quality, domain-dependent ontology quality, and Application-dependent ontology quality* (see Table 4). The *structural intrinsic aspect* attaches with the *inherent ontology quality. The domain intrinsic and domain extrinsic aspects* were grouped under the *domain-dependent ontology quality*. Finally, the *application extrinsic* aspect was mapped with the *Application-dependent ontology quality*.

In addition to the identified characteristics through the survey, a set of characteristics that can be applied to ontology were adopted from ISO/IEC 25012 standard namely *currentness, credibility, accessibility, availability, and recoverability* [37]. Moreover, another two time-related characteristics: *timeliness* and *volatility* were identified with the characteristic: *currentness*, in which, timeliness depends on both currentness and volatility characteristics [14]. Altogether, twenty (20) characteristics were mapped with the ontology evaluation space. Each of them is elaborated in the following sections and definitions were derived upon the discussed theories and the standard ISO/IEC 25012. Moreover, the associated measures for each characteristic have been presented at the end of the respective sections.

Table 4

Ontology quality model

Dimensions	Structural Intrinsic	Domain Intrinsic	Domain Extrinsic	Application Extrinsic
Characteristics	Compliance Complexity Internal Consistency Modularity	Conciseness Coverage External Consistency	Accuracy Relevancy Functional Completeness Comprehensibility Timeliness (Currentness, Volatility) Credibility	Adaptability Efficiency Accessibility Availability Recoverability
	Inherent Ontology Quality	Domain-Dependent Ontology Quality		Application Dependent Ontology Quality

5.1. Structural Intrinsic characteristics (Inherent ontology quality)

The structural intrinsic aspect considers the characteristics related to the language that is used to represent knowledge and the associated inherent quality of an ontology. Thus, the evaluation does not depend on the knowledge of the domain that an ontology is being modeled. As well, the characteristics can be quantitively evaluated and are automatable. Therefore many artifacts can be seen in this aspect such as OntoQA [89], OOPS[62], OntoMetrics [19].

5.1.1. Compliance

In this study, the term compliance is used to denote language conformity and syntactic correctness. This is an important feature and prerequisite for other analyses [2,31]. In ontology quality, many terms are there with related to this definition. Neuhaus et al. [31] described compliance under the *craftsmanship* dimension as "whether an ontology is well-built in a way that adheres to established best practices". Rico et al. [63] defined compliance using the term *syntactic correctness* by adopting the definition provided in [2] as "the quality of the ontology according to the way it is written". They have used the same measures of attributes: *lawfulness* and *richness* proposed by Burton-Jones et al. [2] to gauge the syntactic correctness. Similarly, Zhu et al. [35] stated the term: *well-formedness* as "syntactic correctness with respect to the rules of the language in which it is written". Poveda-Villalón et al. [62] discussed the modeling decision pitfalls that are related to this aspect which determine whether developers use the primitives provided by the ontology implementation languages in a correct way. Based on the provided definitions, we defined compliance as;

Definition 1 (Compliance). *Compliance refers to the ontology is constructed in accordance with defined standards, conventions, or rules in the ontology representation language*.

5.1.2. Complexity

Complexity describes the topological properties of an ontology [5] also referred to as cognitive complexity [21]. There is no definition found in the selected papers, nevertheless, several attributes have been described such as depth, breadth, fan-outness, category size, semantic variance, etc., (see Table 5). Complexity is an important characteristic that provides pieces of evidence of redundancy, reliability, efficiency of an ontology [27,35,90]. For instance, Sánchez et al. [27] stated that "the larger the topological features (i.e., average and variance of the taxonomic depth, and the maximum and variance of the taxonomic breadth) are, the higher the probability that the ontology is a reliable one". On the other hand, increased complexity affects searching efficiency. For instance, Evermann et al. [44] show that it would take a long time to search instances when the level of categories (i.e., concepts) is increased. Based on the facts, the complexity can be defined as;

Definition 2: (Complexity). *The extent of the cognitive complexity of an ontology.*

5.1.3. Internal Consistency

There are two types of consistencies such as *internal consistency* and *external consistency* that can be associated with different perspectives: *ontology perspective* and *real-world perspective* respectively. Later one associate with the domain knowledge, thus it is discussed under the domain intrinsic aspect. Internal consistency is viewed as an inherent point of view also discussed as *logical consistency* [6,31], thus the evaluation is independent of domain knowledge, and Zhu et al. [35] defined this as "whether there is no self-contradiction within the ontology". Hnatkowska et al. [16] have adopted the definition provided in [9,10], "consistency refers to whether it is possible to obtain contradictory conclusions from valid input data". This definition has been provided considering both mentioned perspectives, in which the internal consistency is defined as "the formal and informal definitions have the same meaning" and a set of errors such as the *circularity* error and *partition* errors are highlighted related to it. Poveda-Villalón et al. [62], have been described the pitfalls that affect the internal consistency also grouped as the critical pitfall that essentially needs to be fixed. The OOPS tool is employed to detect such pitfalls. In addition to that, reasoners can infer the logical consequences underlying knowledge representation and detect the internal inconsistencies such as circularity, tangledness, disjointness [6,9,10,90]. We adopted the definition given in [14] to defined internal consistency in the ontology point of view as follows;

Definition 3: (Internal Consistency). *Internal consistency refers to the ontology is free of logical/formal contradictions with respect to particular knowledge representation and inference mechanisms.*

5.1.4. Modularity

From the ontological point of view, a module is defined as "any subgraph *sg* of a graph *g*, where the set of graph elements S_1 for *sg* is such that $S_1 \subseteq S$, S is a collection of elements in the graph *g*" [6]. In the selected papers, the article [86] defined a module as "a part of an ontology that is partitioned or extracted from an original ontology related to a certain topic". If the ontology contains multiple unrelated topics, then the ontology has multiple modules [95]. Thus, Ma et al. [95] suggested decomposing the ontology into multiple ontologies to serve each topic and enhance cohesion.

In the selected papers, only two papers are given definitions for modularity. Duque-Ramos et al. [4] defined modularity as "the degree to which the ontology is composed of discrete components such that a change to one component has a minimal impact on other components". Zhu et al. [35] stated: "how well the ontology is decomposed into smaller parts, to make it easier to understand, use, and maintain". Other articles described the attributes related to modularity mainly cohesion, and coupling [35,52,84, 86,88]. Many of the measures of these attributes are defined upon the structural notion. For instance, measurements such as the number of classes, the number of root classes, the depth of inheritance tree, and the number of relationships between instances. In addition to the structural measures, the authors in [84], proposed behavioral measures for modularity by relating the characteristics: knowledge encapsulation and coverage. Importantly, Ma et al. [95], defined a set of ontology cohesion metrics based on ontological semantics notion that focuses on the possible inconsistency associated with modules of the ontology such as the number of minimally inconsistency subsets, and the average value of axioms inconsistencies. Oh et al. [86] defined the coupling measures (i.e., the number of disconnected relations of each module) that also can be used to check the consistency between modules and the original ontology. Based on the facts, the generic definition provided in [4] was adopted as follows;

Definition 4: (Modularity). *Modularity refers to the degree to which the ontology is composed of discrete subsets* (i.e., *modules of a graph, sub-graphs) such that a change to one component has a minimal impact on other components.*

Table 5

Measures of characteristics related to the structural intrinsic aspect

Characteristic	Attribute	Measures
Compliance	Lawfulness	The ratio of the total number of breached rules in the ontology is divided by the number of statements in the ontology [2, 54, 60, 63].
	Richness	The ratio of the total syntactical features used in this ontology divided by the total number of possible features in the ontology [2, 54, 60, 63].
Complexity	Size	The number of classes, number of attributes, number of binary relationships, number of instances [35]. number of nodes in ontology graph, maximal length of the path from a root node to a leaf node, number of leaves in ontology graph; number of nodes that have leaves among their children, number of arcs in ontology graph. [6,90].
	Depth	Absolute depth, average depth, minimal depth, maximal depth; dispersion of depth; dispersion of depth divided by the average depth [6,90].
	Breadth	Average breadth; average relation of adjacent levels breadth; maximal relation of adjacent levels breadth; the ratio of dispersion of relations of adjacent levels breadth to the average relation of adjacent levels breadth [6,90].
	Fan-outness	The average number of leaf-children in a node, the maximal number of leaf-children in a node, minimal number of leaf-children in a node; dispersion of a number of leaf-children in a node [90].
	Relationship Richness	The ratio of the number of (non-inheritance) relationships (P), divided by the total number of relationships defined in the schema (the sum of the number of inheritance relationships (H) and non-inheritance relationships (P)) [1,59,88]. The ratio of the number of relationships that are being used by instances I_i that belong to C_i ($P(I_i,I_j)$) compared to the number of relationships that are defined for Ci at the schema level ($P(C_i,C_j)$) [88]
	Inheritance Richness	The average number of subclasses per class [88]
	Attribute Richness	The average number of attributes (slots) per class [1,88]
	Class Richness	The ratio of the number of non-empty classes (classes with instances) (C) divided by the total number of classes defined in the ontology schema (C) [88]
	Semantic Variance	Given an ontology O, which models in a taxonomic way a set of concepts C, the semantic variance of O is computed as the average of the squared semantic distance d (·,·) between each concept $c_i \in C$ in O and the taxonomic Root node of O. The mathematical expression of the semantic variance can be found in [27]
Internal Consistency	-	The number of cycles in an ontology, the number of nodes that are members of any of cycles divided by a number of all nodes of an ontology graph [6,90], Tangledness: number of nodes with several parents, the ratio of a number of nodes with several parents to a number of all nodes of an ontology graph; the average number of parent nodes of a node, [6,90]., Partition errors: subclass partition omission, exhaustive subclass partition omission [10,63]
Modularity	Cohesion	The number of ontology partitions, Number of minimally inconsistent subsets, Average value of axiom inconsistencies [95]. The number of root nodes, maximal length of simple paths, the total number of reachable nodes from roots, the average depth of all leaf nodes [35], Average number of connected components (classes and instances) [88,78]
	Coupling	The ratio of the number of hierarchical relations that are disconnected after modularization to the total number of relations, The ratio of the number of disconnected non-hierarchical relations to the total number of relations after ontology modularization [86], the total number of relationships instances of the class have with instances of other classes [88], the number of class in external ontologies which referenced by the discussed ontology [52]

5.2. Domain intrinsic characteristics

In the domain intrinsic aspects, mainly consider whether ontology representation is modeled with respect to the real world. Thus, some domain understanding is required to assess the characteristics that come under this aspect. Consequently, the characteristics are domain-depended and evaluation can be automated with more effort while employing domain knowledge [31].

5.2.1. Conciseness

The articles [16,35,63] have adopted the definition for conciseness from [10], that is "an ontology is concise if it does not store any unnecessary or useless definitions if explicit redundancies do not exist between definitions, and redundancies cannot be inferred using other definitions and axioms". In which, three types of redundancies have been explained, (i) *Grammatical redundancy errors,* which occur when more than one explicit definition exists in an ontology related to the hierarchical relation either directly or indirectly. For instance, direct repetitions are: defined the *is-a* relation twice in between the same source and target classes, defined the *instances-of* relation twice between the same instance and class. Indirect repetitions, e.g., the definitions: A is a subclass of B, as well as having A is a subclass of C and C is a subclass of B. Similarly, this can occur in the instances level. (ii). *Identical formal definition of some classes*: is appeared when the same formal definition has been used to define two or more classes. (iii) *Identical formal definition of some instances:* is appeared when the same formal definition has been used to defined two or more instances. Additionally, Rico et al [63] have checked redundancies by measuring the precision of an ontology with respect to the standard ontology (i.e., frame of reference) that ensures the ontology does not consist of any unnecessary or useless information. The definition frequently used in the previous studies was adopted in our study as;

Definition 5: (Conciseness), *Conciseness refers to whether all the information included in the ontology is useful and precise. Thus, in an ontology, explicit redundancies do not exist between definitions, and redundancies cannot be inferred using other definitions and axioms.*

5.2.2. Coverage (i.e., Completeness)

Completeness from a real-world perspective is considered as coverage in [1,52,66]. Authors have used the terms completeness and coverage interchangeably to describe completeness from a real-world perspective [63]. The authors in [16,63] have adopted the definition given by Gomez-Perez[10] as "an ontology is complete if and only if; all that is supposed to be in the ontology is explicitly set out in it or can be inferred, and each definition complete". Zhu et al. [35] stated completeness is "the number of elements in the standard (i.e., frame of reference) that are covered by the candidate ontology". As well, Ouyang et al. [52] stated that "the coverage is a number of concepts and relations with regards to the ontology set (i.e., frame of reference)". There are some incompleteness errors, which can be detected under the characteristics: *compliance* and *internal consistency*, are pre-requisite to the completeness in the domain intrinsic aspect such as *missing disjointness, partition errors, missing domains, and ranges, necessary and sufficient conditions, existential* and *universal restrictions* [10,62,63]. Thus, in the domain intrinsic aspect, completeness is the coverage of structure and design (i.e., concepts, instances, relations, and constraints) that can be determined concerning the domain knowledge being modeled. Thus, the definition was derived as;

Definition 6: (Coverage). *Coverage refers to the degree to which an ontology covers the information in the domain that the ontology is being modeled.*

5.2.3. External consistency

Zhu et al. [35] defined external consistency as "whether the ontology is consistent with the subject domain knowledge". Vrandečić [29] stated that external consistency can also be named coherence. Hnatkowska [16] and Rico et al. [63] adopted the consistency definitions provide by Gomez-Perez [10] which is "consistency refers to whether it is possible to obtain contradictory conclusions from valid input definitions". Further, they defined, "a given definition is individually consistent if and only if; the formal and informal definitions are metaphysically consistent with respect to the real-world". According to this definition, Gomez-Perez [10] describes *incorrect semantics* that checks whether the ontology representation is correct with the real-world model. For instance, the incorrect semantic classifications are (i) *dog* class is defined as a subclass of *house*, (ii) *pluto* who is a *dog* in the real-world defined as an instance-of *house*. The same definition has been adopted in [75] for semantic correctness.

Neuhaus et al. [31] defined fidelity: "whether the ontology represents the domain correctly, both in the axioms and in the annotations that document the ontology for humans". This definition is also similar to the definitions provided to the external consistency in [35] and semantic correctness in [9,10,75]. Thus, all definitions are referred to the same characteristic and can only be evaluated relative to the subject domain, thus it is required to define the frames of reference to assess the ontology. Based on the provided definitions, we derived external consistency/semantic accuracy as;

Definition 7: (External Consistency/Semantic Correctness). *External consistency refers to the degree to which the information in an ontology is free from contradiction and is coherent with respect to the real world*.

Table 6

Measures of characteristics related to the domain intrinsic aspect

Characteristic	Attributes	Measures
Conciseness		The ratio of the number of classes with the same formal definition as other classes in the ontology divided by the number of classes in ontology, the ratio of the number of instances with the same formal definition as other instances in the ontology divided by the number of instances in the ontology, The ratio of number of redundant subclass-of relations in the ontology divide by the number of hierarchical relations, The ratio of the number of redundant non-hierarchical relations in the ontology divided by number of non-hierarchical relations, The ratio of the number of redundant instance-of relations in the ontology divided by number of instance-of relations in the ontology [63]
	Precision	The ratio of number of classes matches between the candidate ontology and the classes in a frame of reference (i.e., standard ontology) divided by the number of classes in ontology (this measure can be extended for other entities instead of classes such as relations, features, instances) [63]
Coverage	Recall	The ratio of the number of matching entities (i.e., class, relations, instances, terms in case of data extracted from a corpus) between candidate ontology and the standard ontology divided by the total number of entities in the standard ontology [35,63,66]
	Precision	The ratio of the number of classes matches between the candidate ontology and the classes in a frame of reference (i.e., standard ontology) divided by the number of classes in ontology (this measure can be extended for other entities instead of classes such as relations, features, instances) [34,66]
	F-measure	The harmonic means between the Recall and Precision metrics [66]
External Consistency	Vagueness	The ratio of the number of ontological elements (classes, relations, and data types) that are vague, divided by the total number of elements, the ratio of the number of vague ontological elements that are explicitly identified as such, divided by the total number of vague elements, the degree to which the ontology's users disagree on the validity of the (potential) instances of the ontology elements [73]
	Interpretability	The ratio of the number of terms that have a sense listed in an independent authority divided by the total number of terms used to define classes and properties in the ontology [54,60,63]
	Clarity/Precision	The ratio of the total number of terms used to define classes and properties in the ontology divided by the number of definitions for terms in an independent authority that occur in the ontology [54,60]
	Semantic Richness	The ratio of correct concepts, Average ratio of correct instances, Average ratio of correct attributes, Average ratio of correct relations [35]

5.3. Domain extrinsic characteristics

From the *domain extrinsic* point of view, quality evaluation is performed taking into account an ontology as a whole without peering into the internal structure and design. Thus, it is checked whether the ontology meets the requirements for its intended use also defined as the fitness of the ontology in [31]. Moreover, the characteristics associated with this aspect are functional and subjective as the evaluation is performed with respect to the specified tasks and requirements of users (i.e., ontology consumers' views, domain experts, application users, agents of intelligent systems). Neuhaus et al. [31] stated that "successful answers to competency questions provide evidence that the ontology meets the model requirements that derive from query-answering based functionalities of the ontology".

5.3.1. Accuracy (Functional Correctness)

The selected papers have not provided definitions for accuracy under this aspect. However, The authors in [54,60] have adopted measures from the semiotic metric suit of Burton-Jones et al. [2], they have defined "accuracy is whether the claims an ontology makes are true". Similar to this definition, Duque-Ramos et al. [4] defined *precision* under functional adequacy as "the degree to which the ontology provides the right or specified results with the needed degree of accuracy". The authors in [34] have used answers provided to the competency questions to assess functional correctness. ISO/IEC 25010 provides a definition for functional correctness as "the degree to which a product or system provides the correct results with the needed degree of precision" [39]. This can be adopted to ontology at this stage (in the operational environment) as the ontology acts as a part of an information system.

Definition 8: (Accuracy/Functional Completeness). *Accuracy refers to the degree to which an ontology provides the correct results with the needed degree of precision*.

5.3.2. Relevance

From the selected articles, only Amith et al. [54] have provided a definition for relevance as "fulfillment of a specific use case". McDaniel et al. [60] adopted measures for relevance from [2], in which relevancy is defined as "whether the ontology satisfies the agent's specific requirements". The ISO/IEC 25010 standard defines *functional appropriateness* which is similar to the definition of relevance "the degree to which the functions facilitate the accomplishment of specified tasks and objectives". This definition has been adapted for ontologies and it can be measured by assessing the percentage of ontology successfully answering the competency questions [34,54].

Definition 9: (Relevancy). *Relevance refers to the degree to which the ontology provides information to accomplish the specified tasks and user/agent requirements.*

5.3.3. Functional completeness

Only Poveda-Villalón et al. [62] have discussed the completeness in the functional aspect from the selected articles. They defined completeness as "the coverage of the requirements specified in the ontology requirement specification documents by the ontology". Fox and Grüninger [68] defined functional completeness as "can the ontology represent the information necessary for a function to perform its task". Furthermore, they stated that "the functional completeness of an ontology is determined by its competency". For that, the success of answering the competency questions can be analyzed related to a particular function (application). Similarly, in data quality, completeness is defined as "the extent to which data are of sufficient breadth, depth, and scope for the task at hand" [83]. For software quality, The ISO/IEC 25010 standard defines functional completeness as "the degree to which the set of functions covers all the specified tasks and user objectives". Based on the facts, we derived functional completeness for ontologies as;

Definition 10: (Functional Completeness). *Functional completeness refers to the degree to which the ontology covers information with respect to the specified tasks and user/agent requirements.*

5.3.4. Comprehensibility

Zhu et al. [35] defined comprehensibility of an ontology as "whether human readers can easily understand the semantic description" given in the ontology. The authors of the articles [1,31,78] described comprehension as the level of annotations that facilitate understanding the ontology. To assess the comprehensibility, the measures: *the average number of annotated classes, the average number of annotated relations,* and *the average number of instances per class*, have been used. In the linked data quality assessment, Zaveri et al. [14] show comprehensibility has been interchangeably used with understandability. Same for ontology, Poveda-Villalón et al. [62] and McDaniel et al. [60] evaluate understandability using the same measures similar to [1] such as the *number of annotations per term in the ontology, the number of missing and misusing annotation*s. Basically, these are usability-related measures (i.e., level of annotations) as described in [6] which are related to the communication context of an ontology, also it has been named as ease of use in [60]. Thus, these measures are useful for ontology consumers to understand the ontology when adopting it for an intended purpose. Thus, we derived the definition for comprehensibility as;

Definition 11: (Comprehensibility). *Comprehensibility refers to the degree of annotations (i.e., metadata) of an ontology and its elements that enable users (i.e., ontology consumers) to understand the appropriateness of the ontology for a specified task.*

Table 7

Measures of characteristics related to the domain extrinsic aspects

Characteristic	**Measures**
Accuracy	The ratio of a number of false/true statements in the ontology is divided by the number of statements in the ontology [2,54,60]. The number of competency questions that are correctly answered [34].
Relevance	The ratio of the type of syntax relevant to the user is divided by the number of statements in the ontology [2,60], the percentage of adherence for the competency questions [54].
Functional Completeness	There is no a significant measure that has been defined for completeness in the functional aspect. Thus, based on the definitions, the measure "number of competency questions that have provided complete answers concerning the specified task" was derived.

Comprehensibility	The ratio of annotated classes is divided by the total number of classes, the ratio of annotated instances is divided by the total number of instances, the ratio of annotated semantic relations (object properties) is divided by the total number of semantic relations [1,6,60,78].

5.4. Application extrinsic characteristics

Under the application extrinsic aspect, the ontology quality assessment is performed by considering the ontology as a component of a system. Thus, it attaches with capabilities of reasoning tools, computer systems' components: hardware/software, the environment in which the ontology is used, and the tasks to be performed. Thus, the characteristics in this aspect are evaluated from the application perspective.

5.4.1. Adaptability

Zhu et al. [35] defined adaptability of an ontology in the context of web services as "how easily the ontology can be changed to meet the specific purposes of developing a particular web service". Moreover, the sub-factors such as *tailorability*, *composability*, *extendibility*, and *transformability* have been defined under adaptability. McDaniel et al [60] adopted the definition from Vrandečić [29] which is "adaptability measures how well an ontology anticipates, how its future uses and whether it provides a secure foundation which is easily extended and flexible enough to react predictably to small internal changes". Originally, this definition has been provided by Gomez-Perez [10] to refer to *expandability* and *sensitiveness*. Based on that, Vrandečić [29] stated that these terms are synonyms for adaptability. In addition to that, Gangemi et al. [6] defined adaptability in the term of flexibility as "an ontology that can be easily adapted to multiple views". In which modularity and partition are defined as the attributes related to adaptability. Based on the provided definitions, we derived adaptability as;

Definition 12: (Adaptability). *Adaptability refers to the effort required to change (i.e., add, remove, update) the ontology definitions without altering the definitions that are already verified.*

5.4.2. Efficiency

In this study, the term efficiency refers to ontology computational or performance efficiency. The articles [1,20,41] adopted the definition of computational efficiency given by Gangemi et al. [6] as "an ontology that can be successfully/easily processed by a reasoner". On the other way, it is the response time and memory consumption utilized by reasoners when answering queries, classification, or checking consistency. Gangemi et al. [6] proposed measures related to computational efficiency are disjointness ratio, tangledness, circularity, and restrictions. Additionally, Evermann et al. provided empirical evidence to prove that semantic distance and category size (i.e., the number of instances) influence to search efficiency of an ontology. Similarly, Bouiadjra et al. [1] claimed that the measures of the size such as the average number of classes, the average number of sub-classes per class, the average number of relations, the average number of relations per class, can be adopted to assess the efficiency. Duque-Ramos et al. [4] defined two sub-characteristics of performance efficiency, (i) *Response time*: "the degree to which the ontology provides appropriate response and processing times and throughput rates when performing its function, under stated conditions". (ii) *Resource Utilization*: "the degree to which the application uses appropriate amounts and types of resources when the ontology performs its function under stated conditions". Based on the fact, the definition for efficiency of an ontology was derived as follows;

Definition 13: (Efficiency). *Efficiency refers to the degree to which the ontology can be processed and provide the expected level of performance by utilizing the appropriate amount and types of resources in a specific context of use.*

Table 8

Measures of characteristics related to the application extrinsic aspects

Characteristic	Attributes	Measures
Adaptability	-	The sum of the average number of ancestors for the leaves in an ontology and the ratio of the number of leaves to the total number of classes in an ontology [60].
Efficiency	Size	The sum of the average number of classes, the average number of sub-class per class, the average number of relations, the average number of relations per class, and the average ontology size [1].
		Response time, the number of resources utilized [4]. disjointness ratio [6], The number of cycles in an ontology, the number of nodes that are members of any of the cycles is divided by a number of all nodes of an ontology graph [6,90], Tangledness: number of nodes with several parents, the ratio of a

		number of nodes with several parents to a number of all nodes of an ontology graph; the average number of parent nodes of a node [6]

5.5. Time-related characteristics

Batini et al. [23] proposed three types of time-related dimensions such as currentness, timeliness, and volatility which are interchangeably discussed in the literature. Zaveri et al. [14] adopted these characteristics for linked data quality assessments and shown timeliness depends on characteristics: currentness and volatility. Each of these characteristics was grouped under the dataset dynamicity aspect. Similarly, these three characteristics are applicable for the ontology as the knowledge that is considered for the representation can constantly change and expand in the real domain [76,88]. For instance, in the medical domain, constantly new diseases and respective treatments can be discovered. If an ontology-enabled application is developed to share such knowledge, then the ontology integrated with that application should be updated once the new disease knowledge is discovered and it should be made available to the respective users. Thus, we discussed the three time-related characteristics concerning the dynamic nature of an ontology, and possible measures are presented in Table 9.

5.5.1. Currentess

For data quality, currentness is defined in ISO/IEC 25012 as "the degree to which data has attributes that are of the right age in a specific context of use". Batini et al. [23] defined currentness as "how promptly data are updated". For linked data, Zaveri et al. [14] adopted the same definition given by Batini et al. [23]. There is no definition specifically defined for ontology. However, ontology evolution is required over time while domain information is changed as mentioned above. Thus, we adopted the same definition provide by Batini et al. [23] to assess how ontology is updated with current information available in the domain.

Definition 14: (Currentness). *Currentness refers to how promptly information is updated in the ontology.*

5.5.2. Volatility

Batini et al. [23] stated volatility is "the frequency with which data vary in time". Similarly, for ontologies, Murdock et al. [47] defined volatility as "a measure of the amount of change between two or more different versions of a populated ontology". Stvilia [21] defined it as "the amount of time the content of an ontology remains valid" and can be measured by calculating the average update rate of the ontology. Volatility is a property that assesses the stability of an ontology. Information such as dates of birth, places of birth, manufacturing dates have zero volatility and information like stock exchange prices have a high degree of volatility, thus it is valid for a very short time.

Definition 15: (Volatility). *Volatility refers to the frequency of the content of an ontology that remains valid.*

5.5.3. Timeliness

It is not enough to ontology is up-to-date, but also the information is needed to provide on time for the specified task. In this aspect, there is no definition found in the literature of ontologies, thus we adopted the definition provided by Batini et al. [23] for ontologies.

Definition 16: (Timeliness). *Timeliness refers to how current data are available in ontology for the task at hand.*

5.6. Characteristics Adapted from ISO/IEC 25012

5.6.1. Credibility

Credibility is the quality of being trusted and believe in [25]. For data quality, credibility is defined as "the degree to which data has attributes that are regarded as true and believable by users in a specific context of use" [37]. The term believability has been used interchangeably to describe credibility. Wang et al. [83] defined believability as "the extent to which data are accepted or regarded as true, real, and credible". The product acceptance by the community depends on how it can be trusted. From an ontological point of view, McDaniel et al. [59] have discussed quality attributes that determine community acceptance such as authority and history. These are measures of social quality which initially were defined in the semiotic metric suite that answers the question "Can ontology be trusted" [2]. In which authority is defined as "the extent to which other ontologies/ ontology consumers' rely on it". History is another attribute that is evaluated upon the number of times ontology has been used, the number of years ontology in the public library, and the number of revisions made. By adapting the definition provided in ISO/IEC 25012 [37], we derived definition for credibility in the ontological point of view as;

Definition 17: (Credibility). *Credibility refers to the extent to which an ontology is accepted or regards as true and believable in a specified context of use.*

5.6.2. Accessibility

Ontology is not directly accessed by the users, however through applications. To this point, it is required to consider how easy ontology can be used for the specified purpose. There is no agreed definition is provided in the selected articles. However, Gangemi et al. [6] defined it as "an ontology that can be easily accessed for effective application". Vrandečić [29] defined this as organizational fitness and adopted the same definition provided by Gangemi et al. [6]. The supported attributes are modularity, less logical complexity, annotation, and accuracy with respect to the specified task. However, none of the authors provided the measures to those attributes and have not defined how those attributes are associated with accessibility. Based on the given definition for accessibility from the ontological perspective and the ISO/IEC 25012 definition, we derived the definition for accessibility as;

Definition 18: (Accessibility). *Accessibility refers to the extent to which an ontology can be easily accessed through an application.*

5.6.3. Availability

Duque-Ramos et al. [4] defined availability for ontologies by assuming ontology as a software artifact in an application as "the degree to which a software component (language, tools, and the ontology) is operational and available when required for use". Poveda-Villalón et al [62] stated that "ontology not availability" is a critical pitfall in a situation like a system that entirely depends on the ontology. Based on the facts, the definition for availability was derived as;

Definition 19: (Availability). *Availability refers to the availability of an ontology when required for a specified task.*

5.6.4. Recoverability

Duque-Ramos et al. [4] defined recoverability as "the degree to which the ontology can re-establish a specified level of performance and recover the data directly affected in the case of a failure". None of the other selected papers have discussed recoverability, nevertheless, it is important to consider because ontologies are evolving due to changes in the needs of the application, changes in the domain, changes in conceptualization, and changes in the explicit specification [58,88]. Thus, keeping different versions is useful to track changes, detect invalid modifications, detect inconsistencies, and reestablish the specified level of quality in the case of failure [76,88]. Thus, the definition for recoverability was derived by adapting the definition provided in ISO/IEC 25012.

Definition 20: (Recoverability). *Recoverability is the degree to which the ontology can maintain consistency and preserve a specified level of operations and quality, even in the event of failure, in a specific context of use.*

Table 9

Measures of characteristics adapted from ISO/IEC 25012

Characteristic	**Attributes**	**Measures**
Timeliness	-	Difference between the last modified time of the original/target sources and the last modified time of the particular knowledge in the ontology [14,24]
Curretness	-	Average class currency [21] (This measure can also be extended for other entities of an ontology)
Volatility	-	Average update rate [21] The amount of change between two or more different versions of a populated ontology [47]
Credibility	Authority	The number of other ontologies that link to the target ontology, the number of shared terms there are within those linked ontologies [59]
	History	the number of times ontology has been used, the number of years the ontology is in the public library, and the number of revisions made [59]
Accessibility		The measures of characteristics: complexity and modularly (see Table 5), comprehensibility (see Table 7), and semantic accuracy (see Table 6) can be adapted [6].
Availability		Investigated the ontology is available online as RDF nor as HTML for a specified work (Yes/ No) [62]
Recoverability		There are no specified measures that have been provided. This can be qualitatively evaluated by considering an ontology as a software artifact [4].

Table 10

The respective characteristics discussed in each of the selected approaches with the evaluation aspects

Approaches/ Characteristics	Compliance	Complexity	In. Consistency	Modularity	Conciseness	Coverage	Ex. Consistency	Accuracy	Relevancy	Functional Completeness	Comprehensibility	Timeliness	Volatility	Currentness	Credibility	Adaptability	Efficiency	Accessibility	Availability	Recoverability
Evermann et al. [44],2010		**x**																		
Ma et al. [95],2010				**x**																
Djedidi et al. [78], 2010		**x**		**x**		**x**					**x**									
Tartir et al. [88], 2010		**x**		**x**																
Oh et al. [86], 2011				**x**																
Duque-Ramos et al. [4], 2011	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**						**x**	**x**	**x**	**x**	**x**
Murdock et al. [47], 2011							**x**						**x**							
Bouiadjra et al. [1], 2011		**x**				**x**					**x**						**x**			
Ouyang et al. [52], 2011				**x**		**x**														
Gavrilova et al. [90], 2012		**x**																		
Schober et al. [28], 2012											**x**									
Haghighi et al. [75], 2013			**x**		**x**	**x**	**x**			**x**						**x**				
Neuhaus et al. [31], 2013	**x**						**x**	**x**	**x**	**x**	**x**					**x**				
Sánchez et al. [27], 2014		**x**																		
Poveda-Villalón et al. [62], 2014	**x**		**x**			**x**	**x**				**x**								**x**	
Alexopoulos et al [73], 2014							**x**													
Rico et al. [63], 2014	**x**		**x**		**x**	**x**	**x**													
Batet et al. [55], 2014		**x**																		
Lantow [20], 2016		**x**		**x**																
McDaniel et al. [59], 2016															**x**					
Hnatkowska et al. [16], 2017			**x**		**x**	**x**														
Zhu et al. [35], 2017	**x**	**x**	**x**	**x**	**x**	**x**	**x**				**x**					**x**	**x**			
Tan et al. [34], 2017	**x**		**x**					**x**	**x**											
Ashraf et al. [41], 2018		**x**	**x**			**x**				**x**										
McDaniel et al. [60], 2018	**x**	**x**				**x**	**x**	**x**	**x**		**x**				**x**	**x**				
Kumar et al. [84], 2018		**x**		**x**		**x**														
Demaidi et al. [66], 2019						**x**														
Amith et al. [54], 2019	**x**	**x**					**x**	**x**	**x**						**x**					
Franco et al. [57], 2019		**x**																		

5.7. Relationships between characteristics

It is vital to explore dependencies in between the characteristics as one characteristic may affect another characteristic either negatively or positively. Up to now, there are no relational (i.e., non-hierarchical) models that show the correlations between ontology quality characteristics have been proposed, also it may vary from domain to domain. However, significant discussions on certain characteristics and attributes have been made in the literature. Evermann et al. [44] stated that the searching time is increased when the category size (i.e., the number of instances per class) and semantic distance are high. Moreover, the authors in [27,55] have used complexity attributes: semantic variance of the taxonomies depth and breadth for evaluating semantic accuracy as they are positively correlated, which was verified in [56]. Franco et al. [57] have analyzed the correlation between structural measures that is useful for researchers to ignore the measures that show the same impact in the evaluation. Additionally, in the context of semantic descriptions of web services, Zhu et al. [35] made possible assumptions such as; (i) conciseness, structural complexity, and modularity affect adaptability; (ii) conciseness would reduce the complexity of service descriptions; (iii) efficiency may be reduced if the structural complexity is higher. Similarly, Sánchez et al. [27] claimed that reliability is relatively high when an ontology consists of more topological features. However, valid empirical evidence is required to confirm such correlations. For instance, it is declared that tangledness negatively affects computational efficiency in [6], however, Yu et al. [51] empirically proved that ontology tangledness positively impacts efficiency when an ontology is specified for browsing articles. Thus, it is required more empirical evaluations to confirm the correlation between the characteristics instead of superficial investigations [29].

5.8. Comparison with the selected approach in the survey

In our survey, we selected 30 papers for the review, in which, sixteen journals, eleven conference papers, and three chapters were included. The distribution of articles across the years is presented in Figure 5.

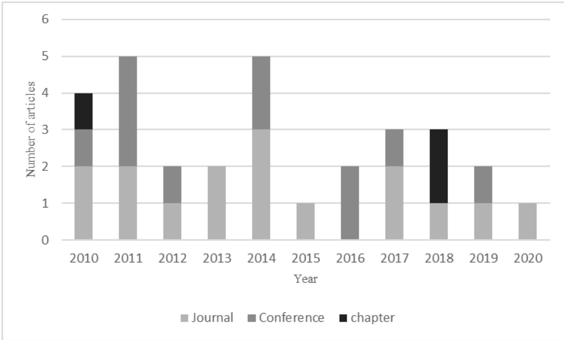


Fig. 5. Distribution of the selected papers across the years

In Table 10, the number of characteristics considered in each approach (i.e., articles) was mapped along with the characteristics defined under the evaluation aspects in Table 4. The authors in [31] have not specifically discussed a set of characteristics instead provided five high-level characteristics covering a set of questions attached to the ontology quality assessment namely *intelligibility* (i.e., can humans understand the ontology correctly?), *fidelity* (i.e., does the ontology accurately represent its domain?), *craftsmanship* (i.e., is the ontology well-built and are design decisions followed consistently?), *fitness* (i.e., does the representation of the domain fit the requirements for its intended use?), and *deployability* (i.e., does the deployed ontology meet the requirements of the information system of which it is part?). Under each level, a concise description of quality to be focused on has been provided. The article [41] has evaluated U ontology by adapting eight characteristics proposed in [29], even though it is unclear how adaptability, conciseness, and fitness were evaluated.

In a summary, the least discussed characteristics in the approaches are credibility, volatility, efficiency, accessibility, availability, and recoverability. None of the approaches considered timeliness and currentness for ontology quality assessments. There are frequent discussions on the characteristics compares to others in descending order: complexity, coverage, external consistency, internal consistency, modularity, compliance, and comprehensiveness. Most of them come under the structural intrinsic aspect and less attention has been provided to the extrinsic aspects of the ontology particularly after the ontology is deployed in a system.

6. Conclusion and future works

This article presented the results of the systematic review on ontology quality assessments. Primarily, a quality model for ontology quality assessment was proposed (see Table 4) with the formalized definitions of the characteristics and related measures. There are eighteen characteristics that mainly were identified with respect to the four aspects of the ontology evaluation space. In which timeliness was further linked with another two characteristics: currentness and volatility. Thus, altogether twenty definitions were derived. The proposed quality model (see Table 4) would provide an underpinning for ontology quality assessment and further experiments are required for a more complete model with a balanced set of characteristics, thereby it can be adapted for any domain with minimum amendments. Additionally, it is vital to empirically explore the effect (i.e., positive, negative) on changes in one characteristic to another that has not been so far discussed. Instead of that, currently, researchers have made assumptions about the correlation between the characteristics that cannot be acknowledged without rigorous experiments as they can be varied in the context where the ontology is built for. Based on the comparison of the previous works, it can be observed that none of the quality models and approaches covered all characteristics in the ontology aspects, nevertheless, a wide range of characteristics have been discussed in OQuaRE [4]. However, it does not support the evaluation of the semantic features of an ontology and the proposed attributes are subjective. Moreover, there is limited evidence related to the quality evaluation in the extrinsic aspect of ontologies, thus, more research on the extrinsic edge is required. In the next step, the aim is to empirically evaluate a set of ontologies that are modeled for a specified task with the identified characteristics and afterward, observe the correlation between the proposed characteristics concerning a specified task in a selected domain.

References

[1] A. Bachir Bouiadjra, S. Benslimane, FOEval: Full ontology evaluation, in: 2011 7th International Conference on Natural Language Processing and Knowledge Engineering, 2011, pp. 464–468.

[2] A. Burton-Jones, V.C. Storey, V. Sugumaran, P. Ahluwalia, A semiotic metrics suite for assessing the quality of ontologies, Data & Knowledge Engineering. 55 (2005) 84–102.

[3] A. Degbelo, A Snapshot of Ontology Evaluation Criteria and Strategies, in Proceedings of the 13th International Conference on Semantic Systems (Semantics2017). Association for Computing Machinery, New York, NY, USA, 2017, pp. 1–8.

[4] A. Duque-Ramos, J.T. Fernández-Breis, R. Stevens, N. Aussenac-Gilles, OQuaRE: A SQuaRE-based approach for evaluating the quality of ontologies, Journal of Research and Practice in Information Technology. 43 (2011) 159–176.

[5] A. Gangemi, C. Catenacci, M. Ciaramita, J. Lehmann, A theoretical framework for ontology evaluation and validation., in Semantic Web Applications and Perspectives (SWAP) – 2nd Italian Semantic Web Workshop, Trento, Italy, 2005, p. 16.

[6] A. Gangemi, C. Catenacci, M. Ciaramita, J. Lehmann, R. Gil, F. Bolici, O. Strignano, Ontology evaluation and validation. Technical report, Lab. for Applied Ontology, 2005.

[7] A. Gillies, Software Quality: Theory and Management. International Thomson Computer Press, 1997.

[8] A. Gomez-Perez, Some ideas and examples to evaluate ontologies, in: Proceedings the 11th Conference on Artificial Intelligence for Applications, IEEE Comput. Soc. Press, Los Angeles, CA, USA, 1995, pp. 299–305.

[9] A. Gómez-Pérez, Asunción Gómez-Pérez,Towards a framework to verify knowledge sharing technology, Expert Systems with Applications,11 (1996), 519-529.

[10] A. Gomez-Perez, Evaluation of ontologies. Int. J. Intell. Syst., 16 (2001), 391-409.

[11] A. Verma, An abstract framework for ontology evaluation, in: International Conference on Data Science and Engineering (ICDSE), 2016, pp. 1-6.

[12] A.Vizedom, OntologySummit2013_From-Use-Context-to-Evaluation--AmandaVizedom_20130228.pdf, http://ontolog-archives.cim3.net/file/work/OntologySummit2013/2013-02-28_OntologySummit2013_OntologyEvaluation-ExtrinsicAspects-2/OntologySummit2013_From-Use-Context-to-Evaluation--AmandaVizedom_20130228.pdf (accessed June 26, 2021).

[13] A. Zaveri, D. Kontokostas, M.A. Sherif, L. Bühmann, M. Morsey, S. Auer, J. Lehmann, User-Driven Quality Evaluation of DBpedia, in: Proceedings of the 9th International Conference on Semantic Systems, Association for Computing Machinery, New York, NY, USA, 2013, pp. 97–104.

[14] A. Zaveri, A. Rula, A. Maurino, R. Pietrobon, J. Lehmann, S. Auer, Quality assessment for Linked Data: A Survey, in: Semantic Web, IOS Press, 2016, pp. 63–93.

[15] B. Boehm, J. Brown, M. Lipow, Quantitative evaluation of software quality, in: Proceedings of the 2nd International Conference on Software Engineering, San Francisco, USA, 1976, pp. 592–605.

[16] B. Hnatkowska, Verification of SUMO ontology, journal-of-intelligent-and-fuzzy-systems/ifs169118, 2017, 1183 – 1192.

[17] B. Kitchenham, Procedures for Undertaking Systematic Reviews, Joint Technical Report, Computer Science Department, Keele University (TR/SE0401) and National ICT Australia Ltd. (0400011T.1), 2004.

[18] B. Kitchenham, S.L. Pfleeger, N. Fenton, Towards a framework for software measurement validation, IIEEE Trans. Software Eng. 21 (1995) 929–944.

[19] B. Lantow, OntoMetrics: Application of On-line Ontology Metric Calculation., in: BIR Workshops, 2016, pp. 1–12.

[20] B. Lantow, OntoMetrics: Putting Metrics into Use for Ontology Evaluation., in: KEOD, 2016, pp. 186–191.

[21] B. Stvilia, A model for ontology quality evaluation, First Monday. 12 (2007). https://doi.org/10.5210/fm.v12i12.2043.

[22] B.C. Grau, B. Motik, G. Stoilos, I. Horrocks, Completeness Guarantees for Incomplete Ontology Reasoners: Theory and Practice, J. Artif. Int. Res. 43 (2012) 419–476.

[23] C. Batini, M. Scannapieco, Data Quality: Concepts, Methodologies and Techniques. Springer-Verlag New York, Inc., 2006.
[24] C. Fürber, M. Hepp, SWIQA – A SEMANTIC WEB INFORMATION QUALITY ASSESSMENT FRAMEWORK, in CIS 2011 Proceedings, 2011, 76.
[25] credibility, https://dictionary.cambridge.org/dictionary/english/credibility (accessed June 26, 2021).
[26] D. Garvin, What Does 'Product Quality' Really Mean?, in: Sloan Management Review, Fall, 1984, pp. 25–45.
[27] D. Sánchez, Semantic variance: An intuitive measure for ontology accuracy evaluation, Engineering Applications of Artificial Intelligence. (2015) 11.
[28] D. Schober, I. Tudose, V. Svatek, M. Boeker, OntoCheck: Verifying ontology naming conventions and metadata completeness in Protégé 4, Journal of Biomedical Semantics. 3 (2012).
[29] D. Vrandečić, Ontology evaluation, in: Handbook on Ontologies, Springer, 2009: pp. 293–313.
[30] D.D. Kehagias, I. Papadimitriou, J. Hois, D. Tzovaras, J. Bateman, A Methodological Approach for Ontology Evaluation and Refinement, (2008) 12.
[31] F. Neuhaus, A. Vizedom, K. Baclawski, M. Bennett, M. Dean, M. Denny, M. Grüninger, A. Hashemi, T. Longstreth, L. Obrst, S. Ray, R. Sriram, T. Schneider, M. Vegetti, M. West, P. Yim, Towards ontology evaluation across the life cycle: The ontology summit 2013, Applied Ontology. 8 (2013) 179–194.
[32] F. Neuhaus, S. Ray, R.D. Sriram, National Institute of Standards and Technology, U.S. Department of Commerce, (2014). https://nvlpubs.nist.gov/nistpubs/ir/2014/NIST.IR.8008.pdf (accessed July 23, 2021).
[33] F. Radulovic, R. García-Castro, A. Gómez-Pérez, SemQuaRE: An extension of the SQuaRE quality model for the evaluation of semantic technologies, Computer Standards & Interfaces. 38 (2015) 101–112.
[34] H. Tan, A. Adlemo, V. Tarasov, M. E. Johansson, Evaluation of an Application Ontology, in The Joint Ontology Workshops, 2017.
[35] H. Zhu, D. Liu, I. Bayley, A. Aldea, Y. Yang, Y. Chen, Quality model and metrics of ontology for semantic descriptions of web services, Tsinghua Science and Technology. 22 (2017) 254–272.
[36] IEEE 730-2014 - IEEE Standard for Software Quality Assurance Processes. https://standards.ieee.org/standard/730-2014.html (accessed June 26, 2021).
[37] ISO 25012. https://iso25000.com/index.php/en/iso-25000-standards/iso-25012 (accessed June 26, 2021).
[38] ISO 9241-11:2018(en), Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts. https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en (accessed June 26, 2021).
[39] ISO/IEC 25010:2011(en), Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models. https://www.iso.org/obp/ui/#iso:std:iso-iec:25010:ed-1:v1:en (accessed June 26, 2021).
[40] ISO/IEC TR 9126-4:2004(en), Software engineering — Product quality — Part 4: Quality in use metrics. https://www.iso.org/obp/ui/fr/#iso:std:iso-iec:tr:9126:-4:ed-1:v1:en (accessed June 26, 2021).
[41] J. Ashraf, O.K. Hussain, F.K. Hussain, E.J. Chang, Evaluation of u ontology, in: Measuring and Analysing the Use of Ontologies, Springer, 2018, pp. 243–268.
[42] J. Brank, M. Grobelnik, D. Mladeni, A survey of ontology evaluation techniques, in: Proc. Conf. Data Min. Data Warehouses, 2005, pp. 166–170.
[43] J. Cavano, J. McCall, A framework for the measurement of software quality, ACM SIGMETRICS Performance Evaluation Review 7 (1978) 133–139.
[44] J. Evermann, J. Fang, Evaluating ontologies: Towards a cognitive measure of quality, Information Systems. 35 (2010) 391–403.
[45] J. García, J. García-Peñalvo, R. Therón, A Survey on Ontology Metrics. in: Lytras M.D., Ordonez De Pablos P., Ziderman A., Roulstone A., Maurer H., Imber J.B., ed., Knowledge Management, Information Systems, E-Learning, and Sustainability Research. WSKS 2010. Communications in Computer and Information Science, Springer, Berlin, Heidelberg, 2010, pp. 22-27.
[46] J. Hartmann, P. Spyns, A. Giboin, D. Maynard, R. Cuel, M.C. Suárez-Figueroa, D1.2.3 Methods for ontology evaluation. EU-IST Network of Excellence (NoE) IST-2004-507482 KWEB Deliverable D1.2.3 (WP 1.2), 2005.
[47] J. Murdock, C. Buckner, C. Allen, Evaluating dynamic ontologies, in: Communications in Computer and Information Science, Springer Verlag, 2013, pp. 258–275.
[48] J. Pak, L. Zhou, A framework for ontology evaluation. In: Workshop on E-Business. 2009, pp. 10–18."
[49] J. Raad, C. Cruz, A Survey on Ontology Evaluation Methods, in: Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, SCITEPRESS - Science and Technology Publications, Lisbon, Portugal, 2015, pp. 179–186.
[50] J. Yu, J.A. Thom, A. Tam, Evaluating Ontology Criteria for Requirements in a Geographic Travel Domain, in: Meersman R., Tari Z., ed., On the Move to Meaningful Internet Systems 2005: CoopIS, DOA, and ODBASE. OTM 2005. Lecture Notes in Computer Science, Springer, Berlin, 2005. pp. 1517-1534
[51] J. Yu, J.A. Thom, A. Tam, Requirements-oriented methodology for evaluating ontologies, Information Systems 34 (8) (2009) 766–791.
[52] L. Ouyang, B. Zou, M. Qu, C. Zhang, A method of ontology evaluation based on coverage, cohesion and coupling, in: 2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 2011: pp. 2451–2455.
[53] L. Tankeleviciene, R. Damasevicius, Characteristics of Domain Ontologies for Web Based Learning and their Application for Quality Evaluation, Informatics in Education. 8 (2009) 131–152.
[54] M. Amith, F. Manion, C. Liang, M. Harris, D. Wang, Y. He, C. Tao, Architecture and usability of OntoKeeper, an ontology evaluation tool, BMC Med Inform Decis Mak. 19 (2019) 152.
[55] M. Batet, D. Sánchez, A Semantic Approach for Ontology Evaluation, in: IEEE 26th International Conference on Tools with Artificial Intelligence, 2014, pp. 138-145.
[56] M. Fernández, C. Overbeeke, M. Sabou, E. Motta, What makes a good ontology? A case-study in fine-grained knowledge reuse, in: The Semantic Web, 2009.
[57] M. Franco, J.M. Vivo, M. Quesada-Mart\'ınez, A. Duque-Ramos, J.T. Fernández-Breis, Evaluation of ontology structural metrics based on public repository data, Briefings in Bioinformatics. 21 (2020) 473–485.
[58] M. Klein, D. Fensel, A. Kiryakov, D. Ognyanov, Ontology Versioning and Change Detection on the Web, Ontology Versioning and Change Detection on the Web, in: Gómez-Pérez A., Benjamins V.R., ed., Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web.

EKAW 2002. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2002, pp. 197-212
[59] M. McDaniel, V.C. Storey, V. Sugumaran, The Role of Community Acceptance in Assessing Ontology Quality, in: Métais E., Meziane F., Saraee M., Sugumaran V., Vadera S., ed., Natural Language Processing and Information Systems, Lecture Notes in Computer Science, Springer, Cham, 2016, pp. 24-36.
[60] M. McDaniel, V.C. Storey, V. Sugumaran, Assessing the quality of domain ontologies: Metrics and an automated ranking system, Data & Knowledge Engineering. 115 (2018) 32–47.
[61] M. McDaniel, V.C. Storey, Evaluating Domain Ontologies: Clarification, Classification, and Challenges, ACM Comput. Surv. 52 (2019).
[62] M. Poveda-Villalón, A. Gómez-Pérez, M.C. Suárez-Figueroa, Oops!(ontology pitfall scanner!): An on-line tool for ontology evaluation, International Journal on Semantic Web and Information Systems (IJSWIS). 10 (2014), 7–34.
[63] M. Rico, M.L. Caliusco, O. Chiotti, M. R. Galli, OntoQualitas: A framework for ontology quality assessment in information interchanges between heterogeneous systems, Comput. Ind. 65 (2014), 1291-1300.
[64] M. Uschold, M. Gruninger, Ontologies: Principles, methods and applications,The Knowledge Engineering Review, 11 (1996), 93-136.
[65] M.C. Suárez-Figueroa, A. Gómez-Pérez, B. Villazón-Terrazas, How to Write and Use the Ontology Requirements Specification Document, in: Meersman R., Dillon T., Herrero P., ed., On the Move to Meaningful Internet Systems: OTM 2009, Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2009, pp. 966-982.
[66] M.N. Demaidi, M.M. Gaber, TONE: A Method for Terminological Ontology Evaluation, in: Proceedings of the ArabWIC 6th Annual International Conference Research Track, Association for Computing Machinery, New York, NY, USA, 2019, pp. 1-10.
[67] M.R. Khondoker, P. Mueller, Comparing ontology development tools based on an online survey, in: World Congress on Engineering (WCE 10), London, UK, 2010.
[68] M.S. Fox, M. Grüninger, Ontologies for Enterprise Modelling. In:: Kosanke K., Nell J.G., ed., Enterprise Engineering and Integration. Research Reports Esprit. Springer, Berlin, Heidelberg, 1997, pp. 190-200.
[69] N. Bevan, Usability is Quality of Use, in: Advances in Human Factors/Ergonomics, Elsevier, 1995, pp. 349–354.
[70] N. Guarino, C.A. Welty, An Overview of OntoClean, in: S. Staab, R. Studer (Eds.), Handbook on Ontologies, Springer Berlin Heidelberg, Berlin, Heidelberg, 2004: pp. 151–171.
[71] N.F. Noy, D.L. McGuinness, What is an ontology and why we need it, Ontology Development 101: A Guide to Creating Your First Ontology. https://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html (accessed June 26, 2021).
[72] Ontology Summit 2013, http://ontolog.cim3.net/OntologySummit/2013/ (accessed June 26, 2021).
[73] P. Alexopoulos, P. Mylonas, Towards vagueness-oriented quality assessment of ontologies, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Springer Verlag, 2014, pp. 448–453.
[74] P. Brereton, B.A. Kitchenham, D. Budgen, M. Turner, M. Khalil, Lessons from applying the systematic literature review process within the software engineering domain, Journal of Systems and Software, 80 (2007), 571-583.
[75] P. Delir Haghighi, F. Burstein, A. Zaslavsky, P. Arbon, Development and evaluation of ontology for intelligent decision support in medical emergency management for mass gatherings, Decision Support Systems. 54 (2013) 1192–1204.
[76] P.Haase, F.V. Harmelen, Z. Huang, H. Stuckenschmidt, Y. Sure, A Framework for Handling Inconsistency in Changing Ontologies, in: Gil Y., Motta E., Benjamins V.R., Musen M.A. (eds) The Semantic Web – ISWC 2005, Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2005, pp. 353-367.
[77] P. Lambrix, Completing and Debugging Ontologies: State of the Art and Challenges, arXiv:1908.03171, 2019.
[78] R. Djedidi, M.A. Aufaure, ONTO-EVO A L an ontology evolution approach guided by pattern modeling and quality evaluation, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2010: pp. 286–305.
[79] R. Lourdusamy, A. John, A review on metrics for ontology evaluation, in: 2nd International Conference on Inventive Systems and Control (ICISC), 2018, pp. 1415-1421
[80] R. Stamper, K. Liu, M. Hafkamp, Y. Ades, Understanding the role of signs and norms in organizations—a semiotic approach to information systems design, Behaviour & Information Technology 19 (1) (2000) 15–27.
[81] R. Studer, V.R. Benjamins, D. Fensel, Knowledge engineering: Principles and methods, Data & Knowledge Engineering. 25 (1998) 161–197.
[82] R.S.I. Wilson, J.S. Gonatillake, A.I. Walisadeera, A. Ginige, ICCSA 2013, Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2021 (accepted June 8, 2021).
[83] R.Y. Wang, D.M. Strong, Beyond Accuracy: What Data Quality Means to Data Consumers, Journal of Management Information Systems, 12 (1996), 5–33.
[84] S. Kumar, N. Baliyan, Quality Evaluation of Ontologies, in: Semantic Web-Based Systems, Springer Singapore, Singapore, 2018, pp. 19–50.
[85] S. Mc Gurk, C. Abela, J. Debattista, Towards Ontology Quality Assessment., in: MEPDaW/LDQ@ ESWC, 2017, pp. 94–106.
[86] S. Oh, H.Y. Yeom, J. Ahn, Cohesion and coupling metrics for ontology modules, Information Technology and Management. 12 (2011) 81.
[87] S. Staab, R. Studer, eds., Handbook on ontologies, Second Edition, Springer, Berlin Heidelberg, 2009.
[88] S. Tartir, I.B. Arpinar, A.P. Sheth, Ontological Evaluation and Validation, in: Poli R., Healy M., Kameas A., ed., Theory and Applications of Ontology: Computer Applications. Springer, Dordrecht., 2010, pp. 115-130.
[89] S. Tartir, I.B. Arpinar, M. Moore, A.P. Sheth, B. Aleman-Meza, OntoQA: Metric-based ontology quality analysis, IEEE Workshop on Knowledge Acquisition from Distributed, Autonomous, Semantically Heterogeneous Data and Knowledge Sources, 2005.
[90] T. Gavrilova, V. Gorovoy, E. Bolotnikova, New Ergonomic Metrics for Educational Ontology Design and Evaluation, Frontiers in Artificial Intelligence and Applications, 246 (2012), 361 - 378.
[91] T.R. Gruber, A translation approach to portable ontology specifications, Knowledge Acquisition, 5(1993), 199-220.
[92] T. R. Gruber, Toward Principles for the Design of Ontologies Used for Knowl edge Sharing, International Journal of Human Computer Studies, 43(1995), 907–928.

[93] V.R. Basili, G. Caldiera, H.D. Rombach, The goal question metric approach. Encyclopedia of Software Engineering 1 (1994), 528–532.
[94] W. Borst, Construction of Engineering Ontologies. PhD thesis, Institute for Telematica and Information Technology, University of Twente, Enschede, The Netherlands, 1997
[95] Y. Ma, B. Jin, Y. Feng, Semantic oriented ontology cohesion metrics for ontology-based systems, Journal of Systems and Software. 83 (2010) 143–152.

Appendix A

Citation	Year	Title of the paper
Evermann et al. [44]	2010	Evaluating ontologies: Towards a cognitive measure of quality
Ma et al. [95]	2010	Semantic oriented ontology cohesion metrics for ontology-based systems
Djedidi et al. [78]	2010	ONTO-EVO^{A}L an Ontology Evolution Approach Guided by Pattern Modeling and Quality Evaluation
Tartir et al. [88]	2010	Ontological evaluation and validation
Oh et al. [86]	2011	Cohesion and coupling metrics for ontology modules
Duque-Ramos et al. [4]	2011	OQuaRE: A SQuaRE-based Approach for Evaluating the Quality of Ontologies
Murdock et al. [47]	2011	Evaluating Dynamic Ontologies
Bouiadjra et al. [1]	2011	FOEval: Full Ontology Evaluation
Ouyang et al. [52]	2011	A Method of Ontology Evaluation Based on Coverage, Cohesion and Coupling
Gavrilova et al. [90]	2012	New Ergonomic Metrics for Educational Ontology Design and Evaluation
Schober et al. [28]	2012	OntoCheck: verifying ontology naming conventions and metadata completeness in Protégé 4
Haghighi et al. [75]	2013	Development and evaluation of ontology for intelligent decision support in medical emergency management for mass gatherings
Neuhaus et al. [31]	2013	Towards ontology evaluation across the life cycle
Sánchez et al. [27]	2014	Semantic variance: An intuitive measure for ontology accuracy evaluation
Poveda-Villalón et al. [62]	2014	OOPS! (OntOlogy Pitfall Scanner!): An On-line Tool for Ontology Evaluation
Alexopoulos et al [73]	2014	Towards Vagueness-Oriented Quality Assessment of Ontologies
Rico et al. [63]	2014	OntoQualitas: A framework for ontology quality assessment in information interchanges between heterogeneous systems
Batet et al. [55]	2014	A Semantic Approach for Ontology Evaluation
Radulovic et al. [33]	2015	SemQuaRE — An extension of the SQuaRE quality model for the evaluation of semantic technologies
Lantow [20]	2016	OntoMetrics: Putting Metrics into Use for Ontology Evaluation
McDaniel et al. [59]	2016	The Role of Community Acceptance in Assessing Ontology Quality
Hnatkowska et al. [16]	2017	Verification of SUMO ontology
Zhu et al. [35]	2017	Quality Model and Metrics of Ontology for Semantic Descriptions of Web Services
Tan et al. [34]	2017	Evaluation of an Application Ontology
Ashraf et al. [41]	2018	Evaluation of U Ontology
McDaniel et al. [60]	2018	Assessing the Quality of Domain Ontologies: Metrics and an Automated Ranking System
Kumar et al. [84]	2018	Quality Evaluation of Ontologies
Demaidi et al. [66]	2019	TONE: A Method for Terminological Ontology Evaluation
Amith et al. [54]	2019	Architecture and usability of OntoKeeper, an ontology evaluation tool
Franco et al. [57]	2020	Evaluation of ontology structural metrics based on public repository data