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 use-case. 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

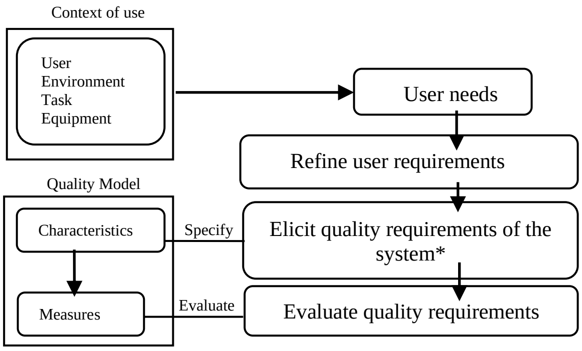
1. Introduction

Ontology is a specified conceptualization of the world that is represented for some purpose [140]. More specifically, an ontology has been defined as "a formal, explicit specification of a shared conceptualization" in [124] merging the two prominent definitions of Gruber [140]: "ontology is an explicit specification of a conceptualization" and Borster [143]: "an ontology is a formal specification of a shared conceptualization". The conceptualization denotes an abstract and simplified view of the world that consists of concepts that are expected in the world being represented, and relationships among them [132]. Thus, ontology has also been defined as a representational vocabulary in [107,140]. A formal specification is machine-interpretability in a way that a machine can understand and process the specified conceptualization. This would enable ontology to be used for several purposes such as to make meaningful communication among their agents/users, facilitate interoperability, knowledge sharing and reuse among subsys-

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

tems and, to distinguish the domain knowledge from operational knowledge. These can be viewed as capabilities of ontologies [97,107] and consequently, ontologies are increasingly incorporated in information systems. However, the above capabilities would not be a reality unless a good quality ontology is produced at the end of the development. To this end, it would be beneficial if a universally accepted methodology or model or an approach for the quality evaluation of ontologies is available. However, such a methodology, model or an approach has not been identified for ontology quality evaluation as the way it is in the other related fields such as system and software engineering. For instance, the availability of widely agreed quality standards in software engineering such as ISO/IEC 9126 [63], ISO/IEC 9241-11 [61], ISO/IEC 25012[60], ISO/IEC 25010 [62], and IEEE software quality standards [59] 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 has been 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 [12, 13, 22, 26, 118]. Thus, it is required to sufficiently evaluate both the quality of the content of the ontology during the development, as well as, the ontology as a whole after the development [12]. This has been viewed as two aspects of ontology evaluation in [52,109,146] 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, quality is considered as a judgment and not a feature. Quality is also referred to as the fitness of a product or service or information for the purpose of the intended users in a specified context [67]. Basically, the quality of a system indicates how far the system caters to user needs with respect to a particular context. From the ontological point of view, a user would be the person who interacts with an ontology or an ontology-enabled IS. The context (i.e., the context of use) is the actual conditions under which a product (i.e., ontology/ontology-enabled information systems) is used by intended users in performing a task (see Figure 1). The context not only considers the technical environment of the system but also user type (i.e., attitude, literacy, experience), the physical and social environment, available equipment and tasks to be performed [48,67]. Accordingly, the quality of an ontology should be evaluated relative to the intended user needs in a particular context [21,51,105]. Thus, the ontology quality requirements to be assessed varies from domain to domain. Not only that, the quality requirements may even vary for different use cases in a single domain.

As inspired by the software and system standards ISO/IEC 9126 [63], ISO/IEC 9241-11 [61], and ISO/IEC 25010 [62], the quality requirements, which can also be viewed as a set of quality characteristics, should be elicited from user needs at the requirement specification phase [21,105]. To this end, a quality model plays a significant role that consists of a proper set of quality characteristics and associated measures. The quality model provides a framework for determining characteristics in relation to the user needs [48,60,62]. Then, the identified quality characteristics can be evaluated at the intrinsic and extrinsic levels of an ontology using their respective measures. This helps to ensure whether the required quality is being achieved across the ontology development. Figure 1 shows the abstract process of deriving the quality requirements and how a quality model supports that process. The significance of a quality model is further elaborated in Section 3 with an example.

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 [3], the quality model of Gangemi et al [7,10], and OQuaRE [6], 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,[96] and the quality model of Zhu et al. [58]. Therefore, these models do not guarantee to work well for other contexts rather than their 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, in turn, the relevant quality measures to achieve the specified quality requirements [109].

(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 [125]. 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 [5].

Thus, there is in fact a requirement to perform a thorough analysis on the domain of ontology quality models and problems. As a result, a systematic review was performed to streamline the findings of the existing quality problems and thereby produce 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. Moreover, in ontology development, ontology can be developed entirely from scratch or can be reused from the existing ontologies or a combination of both. However, ontology quality assessment is performed against the same set of requirements (i.e., characteristics) specified in a particular context regardless of whether the developed ontology is built from scratch or reused [51]. Thus, the conceptual model proposed in this study is helpful for assessing the quality of ontologies irrespective of how they are developed (i.e., from scratch and/or reused). Furthermore, in discussing the conceptual model, we focus on web ontologies which are built upon one of the standard web ontology languages such as RDF(S) [137] and different variants of OWL [138]. However, some quality characteristics such as compliance, consistency, completeness and accuracy can also be associated with other ontology languages. Moreover, we do not distinguish between TBox (i.e., ontology structure) and ABox (i.e., knowledge base) when discussing the quality characteristics. Instead of, the term ontology is used to refer to both TBox and ABox.

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 existing 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, Section 6 discusses the important gaps related to ontology quality models/approaches and Section 7 concludes the survey 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 [20, 69, 72-73,121]. However, we consider that they are not comprehensive surveys as they have not clearly defined: the research gaps and questions that have been addressed and the survey methodology that they have followed. Furthermore, these surveys have provided only a general overview. Moreover, the discussions were limited to a set of characteristics proposed in [7,13,15,46,141]. 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 [46]	**Databases**: Not specified **Scope**: ontologies specified in Web Ontology languages **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 [129]	**Databases**: Three databases: ScienceDirect, IEEE Xplore Digital Library, ACM Digital Library (the selected articles have not been mentioned) **Scope**: ontologies specified in Web Ontology languages

	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 [5]	**Databases**: Semantic Web Journal, Journal of Web Semantics (2003-2017) **Scope**: ontologies specified in Web Ontology languages **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 [94]	**Databases**: Web of Science journals, approximately 170 articles **Scope**: ontologies specified in Web Ontology languages **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 [28,113] 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 reduce the likelihood of publication bias, to select the recently discussed studies and to keep the sample size down to make the review easier to handle. Accordingly, the defined inclusion and exclusion criteria were used 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 eliminating the duplicates using the reference management tool (i.e., Mendeley).

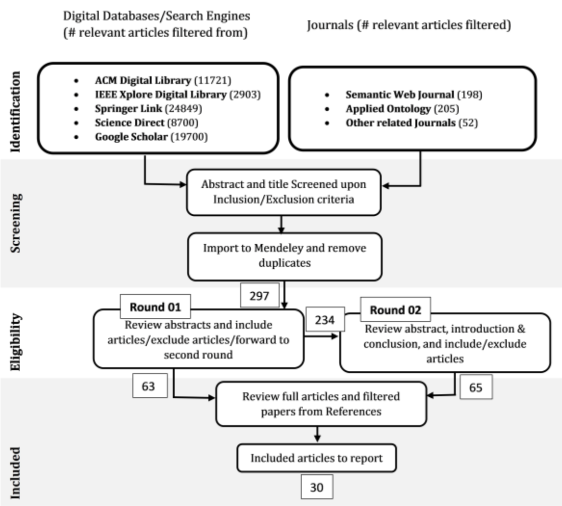


Fig. 2. Main steps of the review process and number of articles retrieved

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 [113]. 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 and they have been considered as the core papers in the survey. In addition to that, the milestone papers, which were captured through the backward citation searching (i.e., before 2010), have also been taken into account for the discussions (see APPENDIX A).

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 [35]. 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 and Table 2 presents a list of terms used in this article.

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) [60,62]. In software engineering, quality models are twofold, relational and hierarchical [11]. 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 [29]. In the ontology quality context, an entity can be a set of concepts, properties, or an ontology [41].

Measure (i.e., metrics): describes an attribute formally and assesses them either quantitively or qualitatively [142].

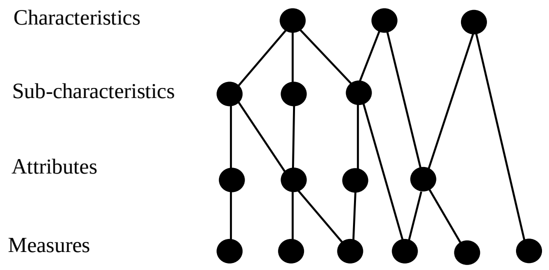


Fig. 3. Association between Characteristics, Attributes, and Measures

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, the same attributes/ measures can be used to assess many characteristics. This has been further illustrated in Figure 4 with the selected ontology characteristics. For instance, *complexity* is one of the ontology characteristics that describes the properties of the ontology structure (i.e., taxonomy, non-taxonomy) of an ontology [7,10,43,46]. Then, it is associated with several attributes such as *size (A1)*, *depth (A2)*, *breadth (A3)*, *fan-outness (A4)* and *density (A5)* of the ontology [8] and each attribute has a set of measures. If the *size* attribute is considered, the related measures are *the number of classes (M1)*, *the number of relations (M2)*, and *the number of instances (M3)*. Thus, the associated levels in the quality models are <characteristics, attributes, measures>.

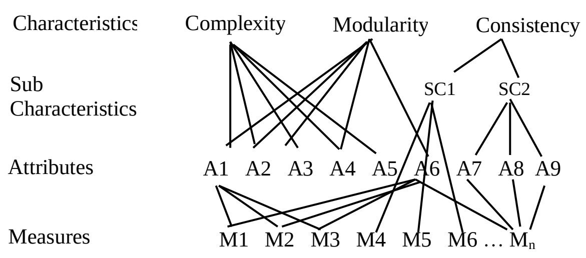


Fig. 4. An example for a simple quality model

Moreover, the measures of the attributes: *size*, *depth*, *breadth*, and *fan-outness* have also been used to assess the *cohesion (A6)* attribute of the *modularity* characteristic [58,135].

If we consider the *consistency* characteristic of an ontology, it can be further derived into two sub-characteristics namely, *internal consistency (SC1)* and *external consistency (SC2)* [15,58]. *Internal consistency* contains direct measures such as *circularity errors (M4)*, *the number of multi-parent classes (M5)*, and *the number of misused disjointness(M6)* [12,15]. Thus, with respect to that scenario, the associated levels in the quality model are <characteristics, sub-characteristics, measures>. However, *external consistency* can be measured through attributes *interpretability (A7), clarity (A8), precision (A9)* [81,96], and each attribute has a set of measures that have been elaborated in Table 8. 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 [25] as the way it is in OQuaRE [6]. 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 classify the characteristics/attributes which are related to similar aspects of the ontology. For instance, the quality model of Gangemi et al. [7,10] contains three dimensions namely structural, functional and usability-related. Similarly, the quality model of Zhu et al. [58] 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 [23,126], 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 [13,46,65,72-73]. Initially, three layers to be focused on have been proposed in [13], namely: content, syntactic & lexicon, and architecture. Later, this was expanded by including the layers: hierarchical and context [46,65]. The syntactic layer considers the properties related to the language that is used to represent the ontology formally [13,65]. The hierarchy or taxonomy layer focuses on the properties related to the taxonomic structure (i.e., *is-a* relationship) of ontologies [65]. A number of measures have been defined related to the taxonomy layer [7,8,43,86,139]. These measures are often useful to observe the dispersion of the ontology structure [43, 86] and to track the evolution of ontologies easily [46]. The authors in [43,86] have shown that the taxonomic measures such as maximum depth/breadth, average depth and depth/breadth variance are good predictors of ontology reliability. The architectural layer is known as the structural or design layer in [65,72]. It considers whether an ontology is modeled based upon the pre-defined design principles and criteria [13,65]. The lexicon layer is also named as vocabulary or data layer [46,65], 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, architecture and semantic layers, 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 [65]. 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 [125] for more detail.

Ontology development methodologies and design patterns. Ontology methodologies and design patterns and principles provide guidelines for knowledge modeling and representation paving the path for accurate knowledge capturing and defining. This enables to achieve the quality characteristics that we emphasize in this study such as compliance, consistency, completeness, and conciseness. Due to this reason, it is necessary to explore the development methodologies, design patterns and principles. However, discussing them in detail is not the main focus of our research.

There are a number of *methodologies* introduced for ontology development such as Cyc [40], Uschold and King's method [98], TOVE [88], METHONTOLOGY [14], SENSUS [34], On-To-Knowledge [133] and DILIGENT [56]. Among them, only Uschold and King's method, TOVE and On-To-Knowledge have discussed ontology evaluation as a phase to be performed after the development of the entire ontology. The METHONTOLOGY mechanism performs the evaluation at every phase of the ontology development. However, none of these methodologies have neither described nor revealed the ontology evaluation methods in detail [108,76]. More details of the comparison of these methodologies can be found in [77,101,108]. In addition to that, agile ontology development methodologies have been introduced by adopting agile principles and practices in software engineering. These include methodologies such as SAMOD [131], XD [49,50], and AMOD [4] and CD-OAM framework [19]. Out of these, only SAMOD and XD discussed the evaluation (i.e., model test, data test, unit test) of ontologies at least in a nutshell. The significance of XD (i.e., eXtreme Design with Content Ontology Design Patterns) is the use of ontology content design patterns to construct ontologies that instinctively support to prevent making common modeling mistakes. This, in turn, directs ontologists to construct a quality ontology particularly assisting the inexperienced ontologists [49].

Ontology Design Patterns (ODPs) are reusable modeling solutions to recurrent ontology design problems [9]. Several groups have been proposed a set of design patterns. For instance, the authors in [85] have proposed seventeen ODPs in relation to the biological knowledge domain. They have classified ODPs into three groups namely (i.) *extension ODPs* which provide solutions to bypass the limitations of OWL such as n-ary relations, (ii.) *good practice ODPs* which support to construct robust and cleaner design (iii.) *domain modeling ODPs* which provide solutions for concrete modeling problems in biology [85]. Moreover, semantic web best practices and deployment working group has introduced ODPs which includes n-ary relations, classes as property values, value partitions/sets, simple part-whole relations. In addition to that, the group ontologydesignpatterns.org suggests more than two hundred patterns and principles which have been categorized into six families namely content ODPs, structural ODPs, correspondence ODPs, reasoning ODPs, presentation ODPs and lexico-syntactic ODPs [18]. The reuse of the patterns and principles can be determined at the ontology design phase with respect to the requirements of the ontology development and it makes ontology modeling faster [51, 117]. To this end, the authors in [49, 117] have shown that pattern-based ontology development avoids making common modeling mistakes through experiments. This in turn improves the quality and usability of ontologies.

Furthermore, OntoClean has proposed a set of principles (i.e., rules) that helps to identify problematic modeling choices associated with taxonomy relationships [106]. Mainly, the rules of OntoClean have been constructed upon a set of metaproperties related to taxonomy namely identity, unity, dependence, essence and rigidity. By embedding these principles, an ontology-driven conceptual modeling language has been proposed namely OntoUML [1, 111, 117]. Moreover, OntoUML provides a set of design patterns that enables to speed up the ontology modeling. In addition to that, it describes a set of anti-patterns to be avoided in modeling ontologies.

Ontology evaluation space. A system usually consists of many components. The quality of each component of a system would influence the overall quality of a system and in turn usability of a system which is often critical for business success [105,144]. As explained under Section 1, the quality requirements which are to be achieved through a system or each system component should be traced from user needs also known as business requirements (see Figure 1). Specification of quality requirements can be viewed as a connected flow starting from user needs to the internal quality of a system. For instance, the *quality requirements* can be elicited from *user needs* and they can also be specified as *a set of quality requirements* to be achieved by a system when it is in use (see Figure 5). This is also defined as *quality in use* [60,105]. Then, this *set of quality requirements* can be used to determine *the external quality requirements* of each component of a system, accordingly, determines *the internal quality requirements*. The quality requirements (i.e., Internal/External) can be expressed as a set of characteristics with associated measures. To this end, a quality model provides a framework for determining characteristics/sub-characteristics in relation to the quality requirements. Then, these characteristics/sub-characteristics can be evaluated using the relevant measures to ensure whether the desired quality in each stage is achieved (see Figure 5).

For an ontology-enable system, ontology quality requirements also need to be distinctly identified from the overall quality requirements of the system [21]. Based on that, the external and internal quality requirements to be achieved from the ontology can be derived. The external and internal quality have been discussed as *extrinsic* and *intrinsic* evaluation aspects of ontologies in [52,109]. Figure 5 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 software and hardware components of the ontology-enabled system have not been elaborated in Figure 5.

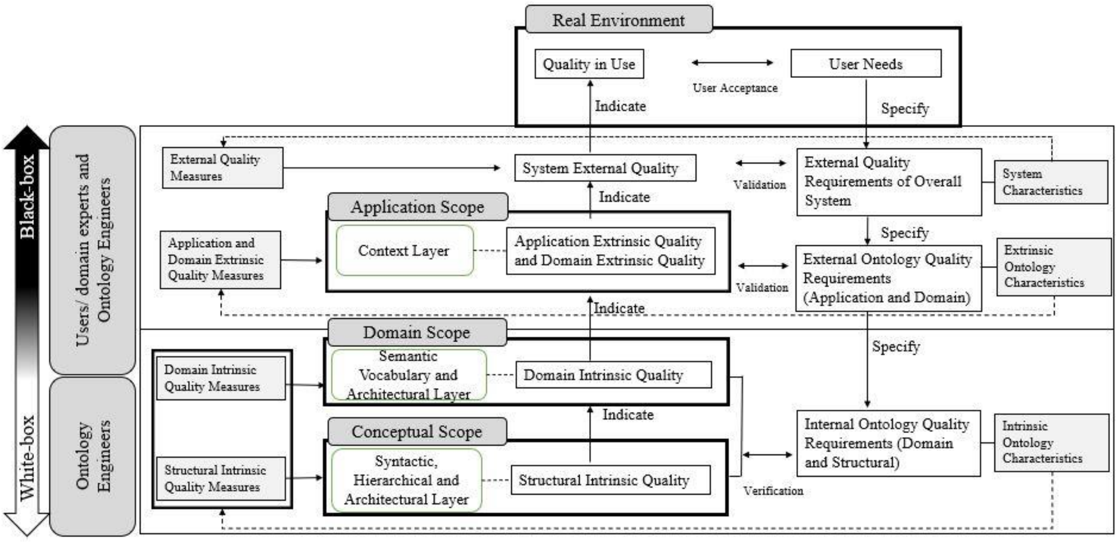


Fig. 5. Conceptualization of ontology quality evaluation

The extrinsic evaluation considers quality under two aspects namely *domain extrinsic* and *application extrinsic*. The ontology quality requirements elicited from the user quality needs can be separated under these aspects [21,100]. For instance, quality requirements associated with the domain knowledge are considered under the domain extrinsic aspect. On the other hand, ontology quality requirements that are specifically needed by the system that the ontology will be deployed are considered under the application extrinsic aspect [52,109]. Moreover, the application extrinsic requirements are independent of the domain knowledge. In performing a quality evaluation under these aspects, an ontology is taken into account as a component of a system. Traditionally this evaluation is called "black-box" or "Task-based" testing [52,70] due to the quality assessment being performed without peering to the ontology structure and content also defined as ontology validation in [132]. Thus, ontology is assessed to confirm whether the right ontology was built [46,132], mainly this is performed with the involvement of users and domain experts (see Figure 5). Moreover, this has also been considered as the context-layer evaluation in [46, 65].

The intrinsic evaluation focuses on the ontology content and structural quality to be achieved in relation to the extrinsic quality requirements. 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* [52] mainly done by ontology engineers. In the structural intrinsic aspect, the focus is given to the syntactic and structural requirements (i.e., syntactic, structural, architectural 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 needs of users [52]. Thus, ontology engineers need to get the assistance of domain experts to evaluate this aspect. Moreover, the semantic, vocabulary and architectural layers of ontologies are evaluated under this aspect. Furthermore, the quality evaluation performed under the structural and domain intrinsic aspects are similar to the paradigm of "white-box" testing in software engineering [70]. Thus, from the intrinsic aspects, verification is being done to ensure whether the ontology is built in the right way [46,132].

The quality requirements identified under each aspect of the ontology evaluation space can be associated with the corresponding set of characteristics together with measures. Then, those characteristics can be measured in order to ensure whether the defined quality requirements are being achieved through ontology development. In the following section, we have illustrated how the quality characteristics can be derived in relation to an ontology requirement using a use case in agriculture.

Use Case. We considered an ontology-enabled decision support system that provides pest and disease management knowledge for farmers. For simplicity, we assumed that ontology quality requirements have already been elicited from user needs and as such consider that the requirement specified below is the key quality requirement of the system.

- The ontology should provide correct pest and disease knowledge for user queries requested through the system.

From the extrinsic aspect of the ontology, the mentioned quality requirement is associated with the domain knowledge in agriculture. Thus, it is a domain extrinsic quality requirement and discusses the accuracy of knowledge to be provided through the ontology. At this stage, the *accuracy* of the knowledge can be observed in terms of the accuracy of answers given for the competency questions [75]. Moreover, *accuracy* should be evaluated with the assistance of domain experts and users. With respect to the identified extrinsic quality requirement, it is required to define what set of characteristics should be met from the intrinsic aspect. To this end, we need to utilize techniques such as ROMEO [75] and GQM [142] that support to derive the intrinsic quality requirements as a set of questions from the higher-level (i.e., extrinsic) quality requirements. By adopting ROMEO methodology [75], the following questions were formulated to derive the intrinsic quality requirements. For instance, to provide accurate pest and disease knowledge for farmers, ontology representation should be correctly modeled. For that, all required definitions (i.e., axioms) with regards to pest and disease management should be correctly defined in the ontology. This is where the adoption of ontology design patterns will become significant. Accordingly, the question: *Q1* can be derived. In addition to that, the ontology should be free from internal contradiction to reason and produce the correct knowledge. Thus, the question: *Q2* also can be derived with respect to the defined extrinsic ontology requirements. It is noteworthy that we have not elaborated all the steps followed under the ROMEO methodology and only the derived questions were highlighted. However, an interesting reader may refer to the article [75].

- Q1: Does the ontology capture the definitions (i.e., axioms) required for representing knowledge in pest and disease management correctly?
- Q2: Is the ontology free from internal contradiction?

Based on the derived questions in relation to the extrinsic quality requirement, it is required to identify the associated characteristics and the relevant measures to be evaluated. To this end, ontology quality models play a key role that supports identifying the corresponding characteristic/sub-characteristic in relation to the derived questions. However, there is no such a well-formed quality model for ontologies as highlighted in Section 1. Thus, to this end, we determined the characteristics based on the literature survey. For instance, Q1 discusses the *external consistency* as it considers whether the definitions (i.e., axioms) are consistent with the domain knowledge. Q2 discusses the *internal consistency* as it considers whether there are any internal contradictions within the ontology definitions. After identifying the characteristics, the relevant measures to be evaluated through ontology development should be identified. This process would be easier if a well-formed quality model is available. For our context, external consistency can be evaluated using the measure of precision which provides a ratio between the correctly defined definitions (i.e., axioms) and the total definitions defined in the ontology. The correctness/consistency of definitions can be compared against a frame of reference such as valid corpus or the standard ontology [75]. To this end, if the value of precision is one then it implies that the particular ontology has correctly captured all axioms in the considered context. Moreover, internal consistency can be assessed by observing the number of logical inconsistencies with the support of a reasoner. In addition to that, the measures such as the number of circularity errors, the number of subclass partitions with common classes and the number of partitions with common instances can be observed (see Table 7). These measures are considered as a set of pitfalls which could lead to wrong inference [15, 95,96] and they can be detected using the tool OOPS! [95].

In this way, for the rest of the ontology quality requirements, the relevant ontology characteristics can be determined with the support of a quality model. These characteristics can then be measured through ontology development to ensure that the ontology being modeled is of good quality.

Table 2

Related Terms and Definitions

Term	Definition
Quality model	A set of characteristics and the relationships between them that provide the basis for specifying quality requirements and evaluating the quality of an entity [60,62]

Context of use	users, tasks, equipment (hardware, software and materials), and the physical and social environments in which a product is used [60,62]
Quality in use	The degree to which a product or system can be used by specific users to meet their needs to achieve specific goals with effectiveness, efficiency, freedom from risk and satisfaction in specific contexts of use [60,62]
Semiotics	The field of study of signs and their representations [80]
Design patterns	Reusable modeling solutions to recurrent ontology design problems [9]
Anti-pattern	Domain-related ontological misrepresentations [117]
Pitfalls	Potential errors, modeling flaws, and missing good practices in ontology development [95]
Lawfulness	Correctness of syntactic rules of the ontology profile [3, 83]
Language Richness	The amount of ontology related syntax features used [3]
Interpretability	Meaningfulness of terms [3]
Clarity	Comprehensibility of the term label [3,83]
Cohesion	The degree to which the elements of a module belong together [94].
Coupling	The degree of relatedness between ontology module [130]
Tangledness	The multi-hierarchical nodes of a graph (i.e., a class with several parent-classes [8]

4. Ontology quality models

4.1. Quality models in software engineering

It is worthwhile to analyze the related contribution in neighboring research fields that support identifying established theories and practices relevant to our study. Moreover, the related theories and practices can be reused or adopted rather than developing the theories and practices from scratch. Therefore, the quality models in the software engineering field were explored. Many well-accepted quality models are available in software engineering such as McCall's quality model [66]; Boehm's quality model [25]; Garvin's quality model [42], the data quality model: ISO/IEC 25012[60], and the ISO 9126 quality model [63] (see Table 3).

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 (see Table 3). 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 that 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 [60]. 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 which 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: ISO/IEC 25010 (**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. The reason for redesigning the ISO 9126 model is that it contains issues as a result of the advancement of information technologies and the changes in its environment [63]. SQuaRE comprises of the same set of characteristics that were defined in ISO 9126 with several amendments as described in [63].

Table 3

Comparison of software quality model

Model	Levels and Characteristics
McCall's quality model	**Levels:** factor, characteristics, measures **Characteristics**: correctness, reliability, efficiency, integrity, usability, maintainability, testability, flexibility, portability, reusability and interoperability
Boehm's quality model	**Levels:** primary, intermediate, primitive **Characteristics (intermediate)**: portability, reliability Efficiency, usability (Human Engineering), testability, understandability and modifiability
Garvin's quality model	**Level:** product quality characteristics **Characteristics:** performance, features, reliability, conformance, durability, serviceability, aesthetics, and perceived quality
ISO/IEC 25012	**Levels:** inherent data and system-dependent data quality characteristics **Characteristics:** accuracy, completeness, consistency, credibility, currentness, accessibility, compliance, confidentiality, efficiency, precision, traceability, understandability, availability, portability and recoverability
ISO/IEC 25010	**Levels:** Factors (Quality in use, external quality, internal quality), characteristics, sub-characteristics and measures **Characteristics:** functional suitability, performance efficiency, compatibility, usability, security, maintainability, portability, effectiveness in use, efficiency in use, satisfaction, free from risk, context coverage

From the ontological perspective, there are no agreed quality models as the way it is in software engineering. However, significant contributions have been made to the field and some quality models have been developed by adopting theories in software engineering. For instance, OQuaRE [6] and SemQuaRE [53] are quality models related to ontologies constructed upon the theories described in system and software standards: SQuaRE-ISO/IEC 25010 [63]. Additionally, we realized that the findings of our survey can be classified as the way it has been done for data quality in ISO/IEC 25012. Moreover, it has been identified that some of the characteristics of the data quality model such as credibility, timeliness, recoverability and availability can be adopted for the ontology quality. This has been further discussed in Section 5. The subsequent section discusses the existing quality models for ontologies.

4.2. Quality models for ontologies

Initially, a framework has been proposed in [13] to verify that developers are building a correct ontology. The proposed framework consists of a set of characteristics namely *soundness, correctness, consistency, completeness, conciseness, expandability*, and *sensitiveness*, which have been adopted in later research works and developments [46,75,94,96]. Thereafter, significant methods [97,106-107] and tools such as OntoTrack [103], OntoClean [106], OntoQA [135], OntoMetric [30], SWOOP [103] 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 [3], the quality model of Gangemi [7,10], and OQuaRE [6].

Semiotic metric suite [3]: Semiotic theory [123] has taken as the foundation to this model as the ontology has a semiotic nature. The semiotic theory has been defined in [123] as "the study of the interpretations of signs and norms" with six levels namely syntactic, semantic, pragmatic, social, physical, and empiric. In the semiotic metric suite, a set of attributes have been classified related to the ontology quality under the aforementioned levels as follows excluding physical and empiric levels. These aspects can also be considered as quality dimensions.

- **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 on an extrinsic nature. To the reason that, under the pragmatic and social levels, the ontology is assessed as a whole and measures are evaluated with reference to the domain experts' knowledge [83], and external documents: usage logs, page ranking, details of external links with other ontologies [94].

The quality model of Gangemi et al [7,10]: 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 [8]. 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 [6]: By adopting the standard SQuaRE (ISO/IEC 25000:2005) [62], a quality model has been proposed to evaluate ontology quality. OQuaRE 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, it has been recognized that OQuaRE [6] has not focused on the semantic nature of the ontology that is an essential component when processing meaningful interpretations. In addition to that, the authors in [46, 58, 96] have highlighted the issues related to OQuaRE. They have stated that the proposed sub-characteristics are subjective and are difficult to be applied in practice. Consequently, OQuaRE [6] has not appeared in the later research. However, noticeable studies are available related to the other two models: semiotic metric suit [3] and the quality model of Gangemi et al [7,10].

OntoKeeper [83] and DoORS [93] are semiotic-metric-suite-driven initiatives. Of 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 [3] with a set of additional characteristics: *structure*, *precision* (i.e., instead of clarity), *adaptability*, *ease of use*, and *recognition*.

The quality model of Gangemi et al. [7,10] has been adopted in multiple studies [46,95,119,139]. OOPS! [95] has used the model to classify the proposed common pitfalls which can occur during ontology development. Furthermore, OOPS! has extended the dimensions to another level as presented in Table 4. The authors in [139] have used the structural measures defined in the model of Gangemi et al. [7,10] to evaluate the ontology cognitive ergonomic. Moreover, a theoretical model proposed in [46] 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 4) [119].

In addition to that, few efforts have been made in designing quality models for a specific purpose such as OntoQualitas [96], the quality model for semantic descriptions of web services [58], and SemQuaRE [53]. 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 [75] 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 [58] have presented a quality model which comprises the four levels namely: *aspects*, *attribute*, *factors*, and *metrics* (see Table 4). This can be viewed as dimensions, characteristics, sub-characteristics/attributes, and measures as per our definitions in section 3.

SemQuaRE [53] 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, there are two strategies that have been followed in constructing models for software quality [120] namely the top-down approach and the bottom-up approach. The top-down approach starts to construct the model from the characteristics and then decomposes them into other levels which contain sub-characteristics and attributes respectively. Finally, the corresponding measures for each characteristic/attribute are identified. In contrast to that, the bottom-up approach observes the related measures which have been derived in the existing studies and classify them until reaching the top-level characteristics. [120]. When considering the discussed models related to ontologies, only SemQuaRE [53] model has been constructed following the bottom-up approach. All the other models have followed the top-

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

down approach, in which the characteristics are determined related to ontologies, from that, the possible sub-characteristics/attributes and measures are derived. The quality models for ontologies that we have explored through the survey are summarized in Table 4. 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 5, the characteristics/attributes proposed in the ontology quality models in [58,93,96,119] have been mapped with the respective aspects (i.e., *structural intrinsic, domain intrinsic, extrinsic (domain/application)*, and *quality in use*).

Table 4

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 [119], 2010	**Structure**: Hierarchical model **Base model**: No specific model is used **Purpose**: To assess the quality of evolving ontology **Approach**: Top-down approach: the characteristics have been derived from [8,46,135] **Evaluation**: Empirical evaluation has not been presented.	Content	Complexity Cohesion Conceptualization Semantic Richness Attribute Richness Inheritance Richness Abstraction
		Usage	Completeness Precision Recall Comprehension
OQuaRE [6], 2011	**Structure**: Hierarchical model **Base model**: The SQuaRE standard **Purpose**: To rank, select, compare and assess the ontologies. **Approach**: Top-down approach **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 [96], 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 [75], then the criteria have been derived from [3,15,106] **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! [95], 2014	**Structure**: Hierarchical model **Base model**: the model of Gangemi et al. [7,10] **Purpose**: To classify the identified pitfalls **Approach**: Top-down approach **Evaluation:** the model of Gangemi et al. [7,10] has been extended to classify the pitfalls, thus, no evaluation has been provided.	Structural	Correctness Modeling Completeness Ontology Language conformance
		Functional	Requirement Completeness Content Adequacy
		Usability-related	Ontology Understanding Ontology Clarity
Quality model of Zhu et al. [58], 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 five ontologies which underlying of five web ser-	Content	Correctness Internal Consistency External Consistency Compatibility Completeness Syntactic Completeness Semantic Completeness

[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

	vices in applications of reporting weather forecasts. The evaluation has been performed against the standard[5] weather ontology.	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. [93], 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 [6] and the quality model in OOPS! [95] has not been included in Table 5. To the reason that the OQuaRE [6] characteristics have been defined by considering the ontology as a software artifact, thus it is difficult to distinctly map with the ontological evaluation aspects and OOPS has adopted 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 characteristics related to the intrinsic extent. Of which, the characteristics: *completeness* (coverage), *conciseness* (Precision), and *representational correctness*, are attached with the domain that the ontology is considered (see Table 5). Additionally, OntoQualitas has taken quality in use of ontology into account defining the characteristics *usefulness*, i.e., *usefulness* of the ontology for the heterogeneous information interchange.

The quality model of Zhu et al. model [58] 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. [93] 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.

[5] https://www.auto.tuwien.ac.at/downloads/thinkhome/

Table 5

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 [119], 2010	Complexity Cohesion Conceptualization Abstraction Comprehension	Completeness (Precision and recall)		
OntoQualitas, [96], 2014	Language conformance Completeness: *is-a/non-isa* Conciseness: *is-a/non-isa* Syntactic Correctness Semantic Correctness: *is-a/non-isa*	Completeness: *Coverage* Conciseness: *Precision* Semantic Correctness: *interpretability, clarity* Representation correctness		Usefulness
Quality model of Zhu et al. [58], 2017	Internal consistency Well-formedness Structural Complexity Modularity	External Consistency Compatibility Completeness Conciseness	Applicability Adaptability Efficiency Comprehensibility	
Quality model of McDaniel et al. [93], 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 [142], the study [75] 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, fourteen (14) significant characteristics: *compliance, complexity, internal consistency, modularity, conciseness, coverage, external consistency, comprehensibility* *(i.e., intrinsic point of view),* *accuracy, relevancy, functional completeness,* *understandability (i.e., extrinsic point of view)*, *adaptability, efficiency*, were identified. Then, the identified characteristics were 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 [60] which consists of only two categories namely *inherent data quality*, and *system-dependent data quality*. However, we identified and defined three main categories for ontology quality viz: *inherent ontology quality*, *domain-dependent ontology quality*, and *application-dependent ontology quality* (see Table 6). The *structural intrinsic aspect* attaches with the *inherent ontology quality. The domain intrinsic and domain extrinsic aspects* are grouped under the *domain-dependent ontology quality*. Finally, the *application extrinsic* aspect was mapped with the *application-dependent ontology quality*.

Table 6

Ontology quality model (the gray-colored characteristics were adopted from the data quality standard: ISO/IEC 25012)

Dimensions	Structural Intrinsic	Domain Intrinsic	Domain Extrinsic	Application Extrinsic
Characteristics	Compliance Complexity Internal Consistency Modularity	Conciseness Coverage External Consistency Comprehensibility	Accuracy Relevancy Functional Completeness Understandability Timeliness *Volatility* *Currentness* Credibility	Efficiency Accessibility Availability Recoverability Adaptability
	Inherent Ontology Quality	Domain-Dependent Ontology Quality		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* [60]. Moreover, other two time-related characteristics: *timeliness* and *volatility* were identified with the characteristic: *currentness*, in which, timeliness depends on both currentness and volatility characteristics [23]. Altogether, twenty-one (21) 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.

5.1. Structural intrinsic characteristics (i.e., 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. i.e., ontology structural properties and internal consistency. 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 have been introduced for structural intrinsic evaluation such as OntoQA [135], OOPS! [95], OntoMetrics [30], XD analyzer [50], ontologyAnalyzer [136,148] and Delta [55] (see Table 13).

5.1.1. Compliance

In this study, the term compliance is used to denote language conformity (i.e., syntactic correctness) and adherence to the guidelines and specifications provided by the ontology language. Language conformity refers to “how the syntax of the ontology representation conforms to an ontology language” [15]. To this end, Rico et al. [96] defined the term syntactic correctness as “the quality of the ontology according to the way it is written” by adopting the definition provided in [3]. Moreover, they have used the measures of attributes: *lawfulness* and *richness* proposed by Burton-Jones et al. [3] to gauge the syntactic correctness (see Table 7). Similarly, Zhu et al. [58] stated the term: well-formedness as “syntactic correctness with respect to the rules of the language in which it is written”. Furthermore, Neuhaus et al. [51] have stated that syntactic well-formedness is a criterion for a well-built ontology. They have discussed syntactic correctness as a part of the *craftsmanship* dimension which considers “whether an ontology is well-built in a way that adheres to established best practices”. In addition to that, Poveda-Villalón et al. [95] have discussed ontology compliance, not only by considering the syntactic correctness but also by taking into account the standards, i.e., specification and guidelines/styles, introduced for the ontology language (i.e., OWL [32, 111]). For instance, in addition to syntax correctness, the authors in [95] have considered the aspect of whether developers use the primitives provided by the ontology implementation languages in a correct way. For example, if we consider Web Ontology Language (OWL), there are a few pitfalls (i.e., style issues [32,111]) related to the ontology primitives such as defining "*is*" relationship instead of using "*rdfs:subClassOf*", defining a relationship itself as an inverse relationship instead of using symmetric property and defining a chain property with one relationship[95]. To this end, OOPS! tool has been introduced to detect such pitfalls [95]. Moreover, many ontology reasoners facilitate detection of syntactic invalidity and style issues (i.e., guidelines) considering them as a set of ontology inconsistencies [23,79]. Accordingly, compliance can be viewed as a core feature which is a prerequisite for other quality characteristics [3,51,95]. Based on the provided definitions in works of literature, we defined compliance as in definition 1;

Definition 1 (Compliance). *Compliance refers to the degree to which the ontology being constructed is in accordance with the rules, specifications and guidelines defined in the ontology representation language.*

5.1.2. Complexity

Complexity describes the topological properties of an ontology [7,10]. There is no definition found in the selected papers, nevertheless, several attributes have been described such as depth, breadth, fanoutness, category size, semantic variance, to measure the complexity of ontologies (see Table 7). Moreover, complexity is also referred to as cognitive complexity in [33, 149] since complexity influences humans to understand the ontology and to interpret the knowledge. For instance, it is difficult for humans to understand an ontology with a thousand terms [149] which thus gets high depth, breadth and tangledness [139, 8]. To this end, authors in [8] have defined depth, breadth and tangledness as a set of parameters (i.e., attributes) that adversely affect ontology cognitive ergonomics. Cognitive ergonomics has been defined as a principle related to the quality of an ontology that considers whether "an ontology can be easily understood, manipulated, and exploited by end-users" [8,10]. The same principle and structural attributes (i.e., depth. breadth, fan-out, circularity) have been adopted to assess the structure of educational ontologies with the intention of observing the balance of ontology structure and its perception by users [139]. Moreover, complexity is an important characteristic that provides pieces of evidence of redundancy, reliability, efficiency of an ontology [43,58,139,149]. For instance, Sánchez et al. [43] 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 [68,149]. For instance, Evermann et al. [68] have shown 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 in definition 2;

Definition 2: (Complexity). *Complexity refers to the extent of how complicated the ontology is.*

5.1.3. Internal consistency

There are different types of consistencies discussed in the previous works [12,15,79,116,117]. For instance, *structural consistency, logical consistency* and *external consistency*. *Structural consistency* refers to the syntactic validity of ontologies with respect to the language used to model the ontology [116]. This notion is considered as a part of ontology compliance under our study that has been explained in section 5.1.1. *External consistency* refers to whether knowledge represented in the ontology is consistent with the relative domain knowledge [58]. Accordingly, external consistency evaluation is associated with the domain knowledge. Thus, this notion was discussed under the domain intrinsic aspect. Under the internal consistency, we considered the *logical consistency* of ontologies and viewed the related definitions. Zhu et al. [58] defined internal consistency as "whether there is no self-contradiction within the ontology". The authors in [79,116] have defined internal consistency with respect to description Logic as, "an ontology *O* is consistent if there is an interpretation *I* of *O* that satisfies every axiom in *O*". Hnatkowska et al. [27] in their study on ontology verification have adopted the definition provided in [13,15] which stated that "consistency refers to whether it is possible to obtain contradictory conclusions from valid input data". To this end, the authors in [12,15] have explained a set of errors such as the *circularity* errors and *partition* errors (i.e., subc1ass partition with common instances, subclass partition with common classes, exhaustive subclass partition with common classes) which could lead to ontology inconsistencies. Moreover, the authors in [39] have highlighted a set of anti-patterns such as *AntiPattern AndIsOr, AntiPattern OnlynessIsLoneliness, AntiPattern UniversalExistence* and *AntiPattern EquivalenceIsDifference* which could also lead to logical inconsistencies. To this end, the authors in [49,117] have shown that the use of design patterns (i.e., best practices) enables to mitigate of such anti-patterns. Poveda-Villalón et al. [95], have described the pitfalls that affect the logical consistency of an ontology and grouped them as the critical pitfalls which should be fixed. Some examples for these pitfalls are defining circles in the hierarchy, misusing "owl:allValuesFrom", specifying the domain or the range excessively, and defining wrong equivalent/transitive/symmetric relationships. The OOPS tool is employed to detect such pitfalls. In addition to that, reasoners can infer the logical consequences underlying knowledge representation and detect logical inconsistencies [8,13,15,139]. We adopted the definition given in [23, 116] to define logical consistency for ontologies as in definition 3;

Definition 3: (Internal consistency/Logical consistency). *Internal consistency refers to the extent to which the ontology is free of logical contradictions with respect to particular knowledge representation, i.e., logical contradiction is the assertion of some statement S, and its denial not-S is true at the same time.*

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*" [8]. In the selected papers, the articles [130,36] defined a module as "a subset or a part of an ontology that is partitioned or extracted from an original ontology related to a certain topic". Similarly, the authors in [45] have defined a module as "a subset of an ontology that captures all the knowledge the ontology contains about a given set of terms". However, all the definitions mentioned so far describe the term *module*. To this end, Khan et al. [149] have provided a comprehensive definition for modularity considering the process of modularization. Mainly, the definition emphasizes five dimensions namely *use-case, techniques, type, annotation features* and *evaluation criteria* that are required to take into account in the modularization process. For this study, we considered only the definitions given for modularization from the evaluation perspective. From that perspective, in the selected papers, only two papers have given definitions for modularity. Duque-Ramos et al. [6] 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. [58] stated: "how well the ontology is decomposed into smaller parts, to make it easier to understand, use, and maintain". Other articles have described the attributes related to modularity mainly cohesion, and coupling [58,80,128,130,134] (see Table 7). Many of the measures of these attributes are defined upon the structural notion. Some examples for the measures of modularity given from the structural notions are 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 [134], proposed behavioral measures for modularity by relating the characteristics: *knowledge encapsulation* and *coverage*. Importantly, Ma et al. [147] have 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. [130] have 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. In the light of this, we defined modularity from the evaluation perspective by adopting the definition given in [6], which is as given in definition 4. Furthermore, as highlighted just above, a number of previous studies have assessed modularity using the structural properties of an ontology [149] (see Table 7). Nevertheless, modularity assessment may also depend on the domain knowledge as explained in [147,149]. Accordingly, modularity can be placed in both structural and domain intrinsic aspects (see Table 6).

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 the other components.*

Table 7

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 [3, 83, 93, 96].
	Richness	The ratio of the total syntactical features used in this ontology divided by the total number of possible features in the ontology [3, 83, 93, 96].
	Pitfalls	The number of pitfalls related to the ontology languages (i.e., as explained in OOPS!) [95]
Complexity	Size	The number of classes, number of attributes, number of binary relationships, number of instances [58]. 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. [8,139].
	Depth	Absolute depth, average depth, minimal depth, maximal depth; dispersion of depth; dispersion of depth divided by the average depth [8,139].
	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 [8,139].
	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 [139].
	Tangledness	The 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, [8,139].
	Cycles	The number of cycles in an ontology, the number of nodes that are members of any of the cycles divided by a number of all nodes of an ontology graph [8,139]
	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)) [2,92,134]. 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 ($P(C_i,C_j)$) [134]
	Inheritance Richness	The average number of subclasses per class [134]
	Attribute Richness	The average number of attributes (slots) per class [2,134]
	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) [134]
	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 [43]
Internal Consistency	-	The number of subclass partitions with common classes, the number of subclass partitions with common instances, the number of exhaustive subclass partitions with common classes, the number of exhaustive subclass partitions with common instances. [15,96] The number of logical consistencies (i.e., using reasoners) [23,116]
Modularity	Cohesion	The number of ontology partitions, Number of minimally inconsistent subsets, Average value of axiom inconsistencies [147]. 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 [58], Average number of connected components (classes and instances) [134,119]
	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 [130], the total number of relationships instances of the class have with instances of other classes [134], the number of class in external ontologies which referenced by the discussed ontology [80]

5.2. Domain intrinsic characteristics

The domain intrinsic aspect mainly considers whether an ontology for a certain application domain is modeled according to the relative domain knowledge. 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 [51]. Moreover, a frame of reference can be used to make a judgment about the quality characteristics [13,15,75,88]. A frame of reference can be a standard ontology, corpus given by experts, requirement specification document, and/or a set of competency questions.

5.2.1. Conciseness

The articles [27,58,96] have adopted the definition for conciseness from [15], 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". Moreover, three types of redundancies have been explained in [15] (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: defining the *is-a* relation twice in between the same source and target classes, defining the *instances-of* relation twice between the same instance and class. For an example of indirect repetitions, consider the three definitions defined in an ontology: A is a subclass of B, B is a subclass of C and A is a subclass of C. To this end, A is a subclass of C is a definition that can be inferred from the first two definitions. Thus, an explicit definition of this (i.e., A is a subclass of C) in an ontology will create an indirect repetition. Similarly, this can occur in the instances level. These types of errors (i.e., *Grammatical redundancy errors*) can be eliminated by adhering to the best practices and design principles (i.e., design patterns [18,85], principles [106,111]) defined for ontology modeling [49,117], (ii). *Identical formal definition of some classes*: this occurs when two or more classes exist in an ontology with the same formal definition, however, with different class names. To resolve this, ontologists should identify differences between the particular classes so that they can be distinguished. Otherwise, the duplicate definitions should be removed in order for ontology to contain only the unique formal definitions. (iii) *Identical formal definition of some instances:* this also occurs when the same formal definition has been defined for two or more instances only differentiating them with names. To solve this error, as explained in (ii) above, ontologists should identify the different attributes of the instances in order to distinguish them. Otherwise, the duplicated instances should be removed in a way that an ontology consists of instances only with unique formal definitions.

Rico et al [96] have adopted the theories described under points (i), (ii) and (iii) and have defined measures to evaluate the conciseness of ontologies (see Table 8). In addition to that, conciseness has been assessed by measuring the precision of an ontology with respect to the standard ontology (i.e., a frame of reference). In this way, conciseness ensures that the ontology does not consist of any unnecessary or useless classes, relations/object properties, attributes/data properties and instances/individuals with respect to the considered domain knowledge. The definition frequently used in the previous studies was adopted in our study as in definition 5;

Definition 5: (Conciseness). *Conciseness refers to the fact that all the information included in the ontology is useful and precise. Thus, in an ontology, neither explicit redundancies exist between definitions nor they can be inferred using other definitions and axioms.*

5.2.2. Coverage (i.e., Completeness)

Completeness from a real-world (i.e., domain knowledge) perspective is considered as coverage in [2,80,102]. Authors have used the terms completeness and coverage interchangeably to describe the coverage of domain knowledge in an ontology [96]. The authors in [27,96] have adopted the definition given by Gomez-Perez [15] 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. [58] stated completeness is "the number of elements in the standard (i.e., a frame of reference) that are covered by the candidate ontology". As well, Ouyang et al. [80] stated that "the coverage is a number of concepts and relations with regards to the ontology set (i.e., a frame of reference)". Thus, in the domain intrinsic aspect, completeness is the coverage of structure, content and design (i.e., concepts, instances, relations, and constraints) that can be determined concerning the domain knowledge being modeled. Accordingly, the definition was derived for completeness as in definition 6. Moreover, incompleteness may occur due to missing disjointness, missing domains and ranges, missing necessary and sufficient conditions, missing existential and universal restrictions [15,95,96]. These are some pitfalls that can be detected under the characteristic: *compliance.* Thus, compliance is a prerequisite to completeness in the domain intrinsic aspect.

Definition 6: (Coverage). *Coverage refers to the degree to which an ontology covers the definitions (i.e., axioms) which have been specified (i.e., requirement specifications, standard ontologies, stand-*

ard corpus) with respect to the domain knowledge that the ontology was developed to represent.

5.2.3. External consistency

Zhu et al. [58] defined external consistency as "whether the ontology is consistent with the subject domain knowledge". Vrandečić [46] stated that external consistency can also be named as domain coherence. Hnatkowska [27] and Rico et al. [96] adopted the consistency definitions provided in [15]. the author has also discussed the different notions of consistencies as we highlighted in section 5.1.3. With respect to the external consistency, the author has defined ontology consistency as "interpretation of definitions (formal/informal) should be consistent with respect to the real-world". For instance, if the term *Monday* is defined as a *month* in an ontology, then, that definition is inconsistent in relation to the domain knowledge. Moreover, the author in [15] has discussed these types of inconsistencies as *incorrect semantics*. A few more examples are (i) *cat* class is defined as a subclass of *house*, (ii) *tom* who is a *cat* in the real-world defined as an instance-of *house*, (iii) the relationship *eats* is defined between *cat* and *house*. To this end, the use of development methodologies, adhering to design patterns and principles would prevent modeling errors and mistakes (i.e., as we highlighted in section 3). Thus, in turn, pave the way for building a well-designed ontology.

Neuhaus et al. [51] 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 [58] and semantic correctness in [13,15,114]. 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 frame of reference to assess the ontology [12,15,75]. Based on the provided definitions, we derived external consistency/semantic accuracy as in definition 7;

Definition 7: (External consistency/Semantic correctness). *External consistency refers to the degree to which an ontology (i.e., ontology definitions) is coherent with the specified domain knowledge (i.e., requirement specifications, standard ontologies, standard corpus) that the ontology was developed to represent.*

5.2.4. Comprehensibility

The authors of the articles [2,51,119] have described comprehensibility 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 (see Table 8). In the linked data quality assessment, Zaveri et al. [23] show that comprehensibility has been interchangeably used with understandability. Same for ontologies, Poveda-Villalón et al. [95] and McDaniel et al. [93] evaluate ontology comprehensibility in terms of understandability using the same measures similar to [2] 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 [7,8,10] which are related to the communication context of an ontology. Also, it has been named as ease of use in [93]. To this end, Poveda-Villalón et al. [95] have revealed a set of pitfalls related to the ontology usability-profiling (i.e., ontology understandability and clarity) such as missing annotations (i.e., metadata), misuse of annotations, and use of different naming criteria. The authors recommend correcting these pitfalls in order to enable users (i.e., ontology developers and consumers) to easily recognize and understand the ontology elements. Moreover, the proper use of annotations (i.e., metadata) and the same naming criteria in ontology modeling have been identified as best practices that ensure comprehensibility of ontologies enabling clean and consistent knowledge representations [10,18]. To this end, the presentation ontology design patterns (ODPs) have been introduced in [18] that consist of two ODPs namely Annotation ODPs and Naming ODPs. Furthermore, as highlighted under section 5.1.2, the increased complexity makes an ontology difficult to comprehend and ontology modularization is one of the solutions for it [58,149]. Based on these facts, we derived comprehensibility as in definition 8;

Definition 8: (Comprehensibility). *Comprehensibility refers to the degree of annotations (i.e., metadata) of an ontology and how its elements enable users (i.e., ontology developers and consumers) to understand the appropriateness of the ontology for a specified context of use.*

Table 8

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 [96]
	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) [96]
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 [58,96,102]
	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) [57,102]
	F-measure	The harmonic means between the Recall and Precision metrics [102]
External Consistency		Whether ontology users disagree on the validity of the (potential) instances of the ontology elements [112]
	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 [83,93,96]
	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 [83,93]
	Semantic Richness	The ratio of correct concepts, Average ratio of correct instances, Average ratio of correct attributes, Average ratio of correct relations [58]
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 [2,8,93,119].
	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 divided by the total number of vague elements [102]
	Clarity	The average number of word senses per unique word, and divide that value with the total number of unique words [83], the number of word senses for the term/s (i.e., classes or properties) in WordNet [3]
	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 [83,93,96]

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. It considers whether the ontology meets the domain requirements that are specifically needed for the particular use case (i.e., the context of use) which is also defined as the fitness of the ontology in [51]. To this end, Neuhaus et al. [51] 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”. Moreover, the characteristics associated with this aspect are functional and subjective as the evaluation is performed with respect to the specified tasks and domain requirements of users (i.e., ontology consumers’ views, domain experts, application users, agents of intelligent systems). Hence, some domain knowledge is required for evaluating these characteristics.

5.3.1. Accuracy (Functional correctness)

The selected papers have not provided definitions for accuracy from the domain extrinsic aspect. However, the authors in [83,93] have adopted measures from the semiotic metric suit of Burton-Jones et al.

[3]. According to them, accuracy is "whether the claims an ontology makes are true". Similar to this definition, Duque-Ramos et al. [6] 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 [57] have used answers provided to the competency questions to assess functional correctness (see Table 9). 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" [62]. This definition can be adopted for ontology under the domain extrinsic aspect because the ontology is considered as a whole and a part of an information system under the domain extrinsic aspect [51]. Accordingly, we defined the accuracy of ontologies as in definition 9.

Definition 9: (Accuracy/Functional correctness). *Accuracy refers to the degree to which an ontology provides the correct results (i.e., information and knowledge) with the needed degree of precision*.

5.3.2. Relevancy

From the selected articles, only Amith et al. [83] have provided a definition for relevance as "fulfillment of a specific use case". McDaniel et al. [93] adopted measures for relevance from [3], 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". From the ontological point of view, we define relevancy as in definition 10. It can be measured by assessing the percentage that the ontology is successfully answering (i.e., the correct answers with respect to the specified context of use) from the competency questions [57,83].

Definition 10: Relevancy. *Relevancy refers to the degree to which an ontology provides appropriate information and knowledge with respect to the specified context of use.*

5.3.3. Functional completeness

Only Poveda-Villalón et al. [95] have discussed the completeness in the functional aspect/end-user perspective 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 [104] 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 (i.e., use case). 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" [126]. 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 these facts, we derived functional completeness of ontologies as in definition 11.

Definition 11: (Functional completeness). *Functional completeness refers to the degree to which an ontology provides necessary and sufficient information and knowledge with respect to the specified context of use.*

5.3.4. Understandability

From the ontological point of view, many authors such as in [2,10,51,93,95,51,119] have discussed understandability as an ontology intrinsic characteristic. To this end, understandability has been assessed by observing internal attributes of ontologies such as annotations and naming conventions. We discussed this notion of understandability under the domain intrinsic aspect by defining the term comprehensibility (see Section 5.2.4). Nonetheless, understandability can also be viewed as a domain extrinsic characteristic in terms of how easily users (i.e., ontology consumers, end-users) can understand the knowledge (i.e., answers to competency questions) provided through the ontology. To this end, Zhu et al. [58] defined the understandability of an ontology as "whether human readers can easily understand the semantic description" given in the ontology. The data quality standard: ISO/IEC 25012 also describes understandability from the extrinsic aspect. It defines understandability as "the degree to which data has attributes that enable it to be read and interpreted by users, and are expressed in appropriate languages, symbols and units in a specific context of use" [60]. In addition to that, understandability can be viewed as a characteristic associated with the application extrinsic aspect when an application interprets the knowledge retrieved from the ontology. For instance, in an ontology-enabled information system, end-users access ontologies through software applications rather than directly accessing the ontology. To this end, information and knowledge provided through the ontology are required to interpret by the software application. Thus, from the application extrinsic aspect, understandability can be defined as how easily an application can interpret the knowledge that is retrieved from ontologies in order to express them in appropriate languages/format that the end-users/agents can understand. Consequently, understandability has been mapped associating both domain and application extrinsic aspects in Table 6. Moreover, we defined understandability from the extrinsic aspect as in definition 12 by adopting the definition given in the data quality standard: ISO/IEC 25012.

Definition 12: (Understandability). *Understandability refers to the degree to which the information and knowledge provided through the ontology can be comprehended, without ambiguity and is expressed in appropriate languages, symbols and units in a specific context of use.*

Table 9

Measures of characteristics related to the domain extrinsic aspects

Characteristic	**Measures**
Accuracy	The ratio of the number of false/true statements in the ontology is divided by the number of statements in the ontology [3,83,93]. The number of competency questions that are correctly answered [57].
Relevance	The ratio of the type of syntax relevant to the user is divided by the number of statements in the ontology [3,93], the percentage of adherence for the competency questions [83].
Functional Completeness	There is no significant measure that has been defined for completeness in the functional aspect. Thus, based on the definitions, the measure "the number of competency questions that have provided sufficient answers concerning the context of use" was derived.
Understandability	There is no significant measure that has been defined for understandability in the functional aspect. Based on the definitions, the measure was derived as follows: "The number of competency questions that received answers in the human-readable language (or the language that can be interpreted to other languages with appropriate context terms, symbols and units in a specified context is divided by the total number of competency questions.

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. It considers the requirements of an ontology that are needed by a particular application that the ontology is integrated. Thus, application extrinsic quality attaches with capabilities of reasoning tools, computer systems' components: hardware/software, the technical environment in which the ontology is used, and the tasks to be performed. The characteristics relevant to this aspect are evaluated from the application perspective and this evaluation is independent of domain knowledge.

5.4.1. Adaptability

Adaptability has been frequently discussed under ontology evolution. Ontology evolution has been defined as "timely adaptation of an ontology with respect to certain requirements by maintaining the consistency" [54,115,116]. Similarly, the authors in [15,58] have defined adaptability as how easily ontology can be changed with certain requirements. The requirements for ontology changes may occur due to (i.) changes in user needs, (ii) changes in application needs that the ontology is integrated, and (iii) changes in ontology conceptualization [54]. These changes could lead to add, remove or modify axioms in the ontology. Sometimes, when changes are performed to the ontology that may cause inconsistencies (i.e., structural, logical, domain/external). Thus, to avoid such inconsistencies, some other additional changes to the ontology are required to be performed. To this end, the authors in [12,15] stated that the changes should be performed without altering the axioms already guaranteed. In the selected survey papers, Zhu et al. [58] defined the 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 [93] adopted the definition from Vrandečić [46] 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 [15] referring to *expandability* and *sensitiveness*. In addition to that, Gangemi et al. [8,10] defined adaptability in terms 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 in definition 13. When considering the evaluation of adaptability, the authors in [8,10,93] have been assessed adaptability using quantitative measures which are independent of domain knowledge. However, the authors in [54] have highlighted that adaptability can be assessed by analyzing the answers given to competency questions after modifying the ontology with respect to the changes (see Table 10). Moreover, ontology changes may occur due to the domain and/or application requirements as pointed in (i) and (ii). Therefore, adaptability can be considered as a characteristic that can come under both domain and application extrinsic aspects (see Table 6).

Definition 13: (Adaptability). *Adaptability refers to the effort required to change (i.e., add, remove, modify) the ontology definitions (i.e., axioms) without altering the definitions that are already guaranteed.*

5.4.2. Efficiency

In this study, the term efficiency refers to ontology computational or performance efficiency. The articles [2,31,64] adopted the definition of computational efficiency given by Gangemi et al. [7,8,10] 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. [8,10] proposed measures related to computational efficiency are disjointness ratio, tangledness, circularity, and restrictions. Additionally, Evermann et al. [68] 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. [2] 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. [6] 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 these facts, the definition for efficiency of an ontology was derived as in definition 14.

Definition 14: (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 10

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 [93]. The number of competency questions was correctly answered after the changes [54].
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 [2].
		Response time, the number of resources utilized [6]. disjointness ratio [8], 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 [8,139], 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 [8]

5.5. Time-related characteristics adopted from ISO/IEC 25012

Batini et al. [35] proposed three types of time-related dimensions such as currentness, timeliness, and volatility which are interchangeably discussed in the literature. Zaveri et al. [23] have adopted these characteristics for linked data quality assessments and have shown timeliness depends on characteristics: currentness and volatility. Each of these characteristics was grouped under the dataset dynamicity aspect. Similarly, these three characteristics would become applicable for ontologies as the knowledge that is considered for the representation can constantly change and expand in the real domain [54,115,134]. 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 11.

5.5.1. Currentness

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. [35] defined currentness as "how promptly data are updated". For linked data, Zaveri et al. [23] adopted the same definition given by Batini et al. [35]. There is no definition specifically defined for ontology. However, ontology evolution is required over time while domain information is changed as mentioned in section 5.4.1. We adopted the definition provided by Batini et al. [35] for ontologies as in definition 15.

Definition 15: (Currentness). *Currentness refers to how promptly the ontology information is updated.*

5.5.2. Volatility

Batini et al. [35] stated volatility is "the frequency with which data vary in time". Similarly, for ontologies, Murdock et al. [71] defined volatility as "a measure of the amount of change between two or more different versions of a populated ontology". Stvilia [33] 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 has zero volatility. Information such as stock exchange prices have a high degree of volatility, thus it is valid for a very short time.

Definition 16: (Volatility). *Volatility refers to the frequency of which the content of the ontology remains valid.*

5.5.3. Timeliness

It is not only enough ontology is updated, but also the information and knowledge of an ontology should be made available timely for a specified task [35]. For example, assume a timetable of a particular course unit has been updated. It may not be timely and will become useless if the timetable can only be made available for students after starting the classes although it already contains updated data. For this aspect, there is no definition found in the literature of ontologies. Thus, we adopted the definition given for data quality by Batini et al. [35] for ontologies as in definition 17.

Definition 17: (Timeliness). *Timeliness refers to how up-to-date information and knowledge of an ontology are for a specified task.*

5.6. Other characteristics adopted from ISO/IEC 25012

5.6.1. Credibility

Credibility is the quality of being trusted and believe in [38]. 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" [60]. The term believability has been used interchangeably to describe credibility. Wang et al. [126] 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. [92] have discussed ontology quality attributes such as authority and history that determine community acceptance of ontologies. These are also measures of social quality of ontologies that initially were defined in the semiotic metric suite answering the question "Can ontology be trusted" [3]. Moreover, the semiotic metric suite has defined authority as "the extent to which other ontologies/ ontology consumers' rely on it". History is another attribute that is evaluated upon: the number of times that an ontology has been used, the number of years ontology in the public library, and the number of revisions made. By adopting the definition provided in ISO/IEC 25012 [60], we derived the definition for credibility from the ontological point of view as in definition 18.

Definition 18: (Credibility). *Credibility refers to the extent to which an ontology is accepted or the information and knowledge provided through ontologies regards as true and believable with respect to the specified context of use.*

5.6.2. Accessibility

Ontology is not directly accessed by the users and usually, it is accessed through applications. To this point, it is required to consider how easy an ontology can be used for the specified purpose. The articles that we have selected for this survey have not provided a definition for accessibility. Gangemi et al. [8,10]

have defined accessibility as "an ontology that can be easily accessed for effective application". Vrandečić [46] has adopted the same definition provided by Gangemi et al. [8,10]. Moreover, the author has shown that the term *organizational fitness* is also used to refer to accessibility. The supported sub characteristics for accessibility are modularity, logical complexity, annotation, and accuracy with respect to the specified task. However, none of the authors have provided the measures to those sub characteristics and have not defined how those sub characteristics are associated with accessibility. Based on the definition given in [8,10] for accessibility from the ontological perspective and based on the ISO/IEC 25012 definition, we derived the definition for accessibility as in definition 19.

Definition 19: (Accessibility). *Accessibility refers to the extent to which an ontology can be easily accessed through its application.*

5.6.3. Availability

Duque-Ramos et al. [6] 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 [95] stated that "ontology not availability" is a critical pitfall in a situation where a system entirely depends on the ontology. We adopted the definition given in [23] for ontology availability as in definition 20.

Definition 20: (Availability). *Availability refers to the extent to which the knowledge represented in the ontology (or some portion of it) is present, obtainable and ready for use.*

5.6.4. Recoverability

Duque-Ramos et al. [6] 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 although it is an important characteristic. This is due to the fact that, ontologies are evolving because of the changes in the domain (i.e., user needs), changes in the application needs, changes in conceptualization, and changes in the explicit specification [91,134]. Thus, keeping different versions of an ontology is useful to track changes, detect invalid modifications, detect inconsistencies, and re-establish the specified level of quality in the case of a failure [115,134]. Thus, the definition for recoverability was derived as given in definition 21 by adopting the definition provided in ISO/IEC 25012.

Definition 21: (Recoverability). *Recoverability is the degree to which the ontology maintains consistency and preserves a specified level of quality, even in the event of failure with respect to a specific context of use.*

Table 11

Measures of characteristics adopted 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 [23,37]
Currentness	-	Average class currency [33] (This measure can also be extended for other entities of an ontology) Average update rate [33]
Volatility	-	The amount of change between two or more different versions of a populated ontology [71] Average update rate [33]
Credibility	Authority	The number of other ontologies that link to the target ontology, the number of shared terms there are within those linked ontologies [92]
	History	the number of times ontology has been used, the number of years the ontology is in the open resource, and the number of revisions made [92]
		The number of positive user feedback given on the ontology (i.e, information and knowledge)
Accessibility		The measures of characteristics: complexity and modularly (see Table 7), comprehensibility (see Table 8), and semantic accuracy (see Table 8) can be adopted [8].
Availability		This has been measured w.r.t. YES/NO scenarios. Thus, the following questions can be used to check availability. is an ontology made available to the application that the ontology is integrated (i.e., as RDF file/as HTML) [95]?, is documentation of an ontology made available [95]?, is an ontology consists of URI without any supporting RDF metadata [23]?, is an ontology consists dead links [23]?
Recoverability		There are no specified measures that have been provided. This can be qualitatively evaluated by considering an ontology as a software artifact [6].

Table 12

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	Comprehensibility	Accuracy	Relevancy	Functional Completeness	Timeliness	Volatility	Currentness	Credibility	Understandability	Adaptability	Efficiency	Accessibility	Availability	Recoverability
Evermann et al. [68],2010		x																			
Ma et al. [147],2010				x																	
Djedidi et al. [119], 2010		x		x		x		x													
Tartir et al. [134], 2010		x		x																	
Oh et al. [130], 2011				x																	
Duque-Ramos et al. [6], 2011	x	x	x	x	x	x	x		x	x	x						x	x	x	x	x
Murdock et al. [71], 2011							x						x								
Bouiadjra et al. [2], 2011		x				x		x										x			
Ouyang et al. [80], 2011				x		x															
Gavrilova et al. [139], 2012		x																			
Schober et al. [44], 2012								x													
Haghighi et al. [114], 2013			x		x	x	x				x						x				
Neuhaus et al. [51], 2013	x						x	x	x	x	x						x				
Sánchez et al. [43], 2014		x																			
Poveda-Villalón et al. [95], 2014	x		x			x	x	x												x	
Alexopoulos et al [112], 2014							x														
Rico et al. [96], 2014	x		x		x	x	x														
Batet et al. [84], 2014		x																			
Lantow [31], 2016		x		x																	
McDaniel et al. [92], 2016															x						
Hnatkowska et al. [27], 2017			x		x	x															
Zhu et al. [58], 2017	x	x	x	x	x	x	x									x	x	x			
Tan et al. [57], 2017	x		x						x	x											
Ashraf et al. [64], 2018		x	x			x					x										
McDaniel et al. [93], 2018	x	x				x	x	x	x	x					x		x				
Kumar et al. [128], 2018		x		x		x															
Demaidi et al. [102], 2019						x															
Amith et al. [83], 2019	x	x					x		x	x					x						
Franco et al. [87], 2020		x																			

5.7. Comparison with the selected approach in the survey

In our survey, we selected 30 core papers for review. Of which, sixteen journals, eleven conference papers, and three chapters were included. The distribution of articles across the years is presented in Figure 6.

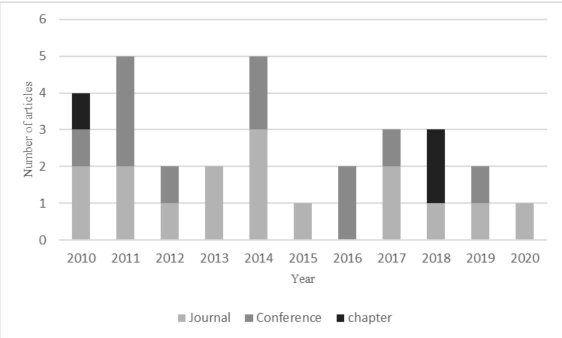


Fig. 6. Distribution of the core survey papers across the years

In Table 12, the number of characteristics considered in each approach (i.e., articles) was mapped along with the characteristics defined under the evaluation aspects in Table 6. The authors in [51] 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 [64] has evaluated ontology named U ontology (i.e., representation of the ontology usage analysis domain) by adopting eight characteristics proposed in [46]. Although it is unclear how the characteristics: 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 have considered timeliness and currentness for ontology quality assessments. There were frequent discussions on the characteristics compared to others in descending order: complexity, coverage, external consistency, internal consistency, modularity, compliance, and comprehensibility. Most of them come under the structural intrinsic aspect. Accordingly, it can be observed that less attention has been provided to the extrinsic aspects of the ontology particularly after the ontology is deployed in a system.

6. Discussion

As it has been highlighted in Section 1, quality is the fitness of a product for the intended use (i.e., requirements/needs) [67]. From the ontological point of view, ontology quality is the fitness of the representation of the domain knowledge for its intended use [51,96]. Therefore, the quality characteristics which are to be achieved by ontologies should be specified with respect to the intended needs of the specified context of use. Based on this fact, all the characteristics presented in Table 6 are not equally important for all contexts of use. The characteristics such as compliance, internal consistency, external consistency, coverage, accuracy, functional completeness and timeliness can be viewed as core characteristics that are required for ontologies in any context of use. In addition to these, there is a set of characteristics that are necessary to achieve through the ontology with respect to the intended needs. For instance, the authors in [149] have exemplified that modularity depends on the requirements of users and applications. Moreover, the authors in [17] have presented that relevancy is a crucial factor for the ontology-enabled decision support system in agriculture. In addition to that, authors in [75,96] have stated that conciseness of ontologies is not a required characteristic when an ontology is modeled for extensive coverage of the domain. Accordingly, to model a good quality ontology, it is required to identify the necessary and sufficient characteristics in relation to the requirements of the considered use case (i.e., the context of use) [75,96]. These identified characteristics should be assessed across the ontology development life cycle [51].

In the light of the survey conducted by the authors, the important research gaps further to be investigated have been identified such as (i) the absence of relational quality models for ontologies, (ii) the lack of systematic approaches for ontology quality specification and evaluation (iii) the lack of methods and tools for assessing the extrinsic quality.

6.1. Relational quality models for ontologies

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. [68] 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 [43,84] 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 [86]. Franco et al. [87] 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. [58] 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. [43] 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 has been declared that tangledness negatively affects computational efficiency in [8], however, Yu et al. [75] 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 [46]. Furthermore, empirically validated relational models for ontologies are essential that can be used as a foundation for future researches rather than performing experiments from the beginning and over again.

6.2. Lack of systematic approaches for ontology quality specification and evaluation

Although quality evaluation is considered as a phase of ontology development methodologies [34,88,98,133], it is an iterative process that should be started from requirement analysis and specification. For example, quality assessment consists of several activities such as; *identification of intended needs through stakeholder discussions (i.e., requirement specification), elicitation of quality requirements from the identified needs (i.e., quality specification), prioritizing the quality requirements, specifying quality characteristics (i.e., intrinsic and extrinsic aspect) and performing quality assessment across the development [21,48,96]*. When carrying out these activities, having a systematic approach is beneficial that can assist ontology developers to identify the quality characteristics and measures relevant to the intended needs. Then, to select the appropriate evaluation methods and tools for assessing characteristics. Based on our survey, a few contributions that cover a part of the ontology quality assessment process were identified. For instance, ROMEO methodology provides a set of guidelines for specifying intrinsic quality characteristics and measures after identifying the ontology extrinsic requirements [75]. The authors in [24] have discussed the specification of measures associated with the selected ontology characteristics such as accuracy, completeness, consistency and uniqueness (i.e., conciseness). Moreover, there are several evaluation approaches (i.e., application-based, data-driven, golden-standards and expert-based) as presented in survey papers [65,73,82,146]. These approaches discuss the techniques of ontology evaluation that usually can be performed after ontology development. In addition to that, ontology methodologies and design patterns provide guidelines for modeling well-designed ontologies as discussed under section 3. However, none of the methodologies discuss quality specifications in relation to intended needs and their evaluation in detail.

Due to the lack of a systematic approach that covers ontology quality specification and evaluation, it is difficult for inexperienced developers to identify a proper set of characteristics related to the considered use cases and to select appropriate methods and tools for evaluation [49,82,96,125]. Thus, in turn, the following pitfalls can occur (i) quality evaluation is limited to the frequently discussed characteristics such as functional completeness (i.e., expressiveness), consistency, and practical usefulness [5], and in turn (ii) essential quality characteristics which are required for a considered context of use in the domain may get ignored [5,96,125] and (iii) inappropriate evaluation approaches/methods and tools for a considered context can be selected. To this end, the derived conceptual quality model through this survey helps to provide an understanding of the possible characteristics associated with ontology evaluation aspects (i.e., intrinsic and extrinsic). Thereby, it supports preventing pitfalls (i) and (ii). Even though, it is worth having a set of formal guidelines that supports ontologists to systematically specify the extrinsic and intrinsic characteristics of an ontology in relation to the intended needs of the system that the ontology is integrated [51]. Thereafter, to identify the appropriate evaluation methods and tools for a considered use case.

6.3. Lack of methods and tools for assessing the extrinsic quality

The use of methods and tools makes the ontology evaluation process easy and reduces the cost of manual evaluation [108,146]. Based on the analysis of the related survey papers and literature reviews [16,95,108,109,136,146], it has been observed that several methods and tools have been introduced to support ontology quality assessment. Table 13 summarizes a set of tools/methods and the evaluation aspects that each tool/method has covered. Based on that, it can be understood that limited work has been carried out related to the extrinsic aspect. Moreover, the survey performed in [109] has analyzed thirteen tools. The results of its also have shown that a few tools such as OntologyTest [127], COLORE [89], and Open Link Virtuoso support [110] assess the domain extrinsic characteristics/attributes. None of the selected tools in that survey has focused on the application extrinsic aspect. Specifically, it has been highlighted that the functions: assessing query time performance (i.e., efficiency), validating the application requirements of the software that the ontology is integrated and assessing user experience with ontology (i.e., quality in use) have not been considered. Accordingly, it is evident that introducing tools and methods for extrinsic quality evaluation are still open for future research.

Table 13

Preliminary analysis of methods and tools

Tool/Method	Topic	Characteristics/Attributes	Aspects
RDF Validation Service [145]	Tool: syntax validator	Language compliance	Structural Intrinsic
OWL Validator [90]	Tool: syntax validator	Language compliance	Structural Intrinsic
OntoAnalyser [15,148]	Plug-in (i.e., OntoEdit): for language conformity and consistency analysis	Compliance and Internal Consistency	Structural Intrinsic
OntoKick [148]	Plug-in (i.e., OntoEdit): requirements specification and evaluation	Accuracy and functional completeness	Domain Extrinsic
OntoClean [106]	Method: for validating the ontological adequacy of taxonomic relationships	External Consistency	Domain Intrinsic
OntoQA [135]	Tool: for metric-based ontology quality analysis	Complexity: relationship richness, attribute richness, inheritance richness, class richness, Average population, fullness. Modularity: cohesion, importance, connectivity, instance relationship richness, etc.	Structural Intrinsic
Unit Test [47]	Method: for analyzing unwanted changes and side effects during the maintenance	External Consistency and Accuracy	Domain Intrinsic and Domain Extrinsic
S-OntoEval tool [122]	Tool: ontology quality analysis	Complexity, Modularity, Internal Consistency, External Consistency, Comprehensibility	Structural Intrinsic and Domain Intrinsic
OntologyTest [127]	Tool: for checking ontology functional requirements	Accuracy and Functional completeness	Domain Extrinsic
OntoCheck [44]	Plug-in (i.e., protégé): for verifying annotations and naming conventions	Comprehensibility	Structural intrinsic attributes that are complementary to the domain intrinsic aspect have been automated.
XD analyzer [50]	Plug-in (i.e., NeOn Toolkit): for verifying annotations and naming conventions	Coverage: isolated entities, missing types, missing domain or range in properties, missing inverse, Comprehensibility: instance missing labels and comments, unused imported ontologies	Structural intrinsic attributes that are complementary to the domain intrinsic aspect have been automated.
Copeland et al method [99]	Method: for ontology regression test	Internal consistency and External Consistency	Structural Intrinsic and Domain Intrinsic

RepOSE[6] [118]	Tool: for detecting and repairing defects in ontologies and alignments	Compliance and External Consistency	Structural intrinsic and Domain Intrinsic
OOPs[7] [95]	Tool: for common pitfalls detection (i.e., pitfalls have been numbered as P01, P02, … P041 in [95])	Compliance: e.g., P34, P35 and P38. Consistency: e.g., P05, P06, P07, P19 and P24. Coverage: e.g., P04, P10, P11, P12 and P13. Conciseness: e.g., P02, P03, P21 and P32. Comprehensibility: e.g., P08, P20 and P22. Availability: e.g., P36 and P37.	Structural Intrinsic, Domain Intrinsic/Extrinsic and Application Extrinsic (i.e., only automated structural intrinsic metrics)
OntoMetric[8] [30,31]	Tool: for metric-based ontology quality analysis	Complexity: basic metric, knowledge base metric Modularity: graph metrics, class metrics Comprehensibility: annotation metrics	Structural Intrinsic and Domain Intrinsic (i.e., only automated structural intrinsic metrics)
OntoDebug [78]	Plug-in (i.e., protégé): for ontology inconsistency debugging	Internal Consistency	Structural Intrinsic
DoORS[9] [93]	Tool: metric-based ontology quality analysis	Compliance, External Consistency, Conciseness, Comprehensibility, Accuracy, Relevancy, and Credibility	Structural Intrinsic, Domain Intrinsic and Domain Extrinsic
OntoKeeper [83]	Tool: for metric-based ontology quality analysis	Compliance: lawfulness, richness, Conciseness, Comprehensibility, Accuracy, Relevancy, and Credibility	Structural Intrinsic, Domain Intrinsic, and Domain Extrinsic
Delta [55]	Tool: for ontology quality analysis (i.e., adopted OntoQA metrics and used OOPs as an external plug-in)	Complexity, Modularity	Structural Intrinsic

7. 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 6) with the formalized definitions of the characteristics and related measures. There are nineteen 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-one definitions were derived. The proposed quality model (see Table 6) 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 adopted 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 [6]. 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, to propose a systematic approach for ontology quality specification and evaluation concerning a use-case in a selected domain.

Acknowledgment

The authors wish to thank the Editors-in-Chief of the journal, the responsible editor: Aldo Gangemi, the reviewer: Stefano de Giorgis and anonymous reviewers for their work in processing our article and for their insightful comments and suggestions.

[6] RepOSE: https://www.ida.liu.se/~patla00/research/RepOSE/
[7] OOPS!:
[8] OntoMetric: https://ontometrics.informatik.uni-rostock.de/ontologymetrics/
[9] DoORS: https://owlparser.herokuapp.com/

References

[1] A.B. Benevides, G. Guizzardi, A Model-Based Tool for Conceptual Modeling and Domain Ontology Engineering in OntoUML, in: J. Filipe, J. Cordeiro (Eds.), Enterprise Information Systems, Springer, Berlin, Heidelberg, 2009: pp. 528–538.

[2] 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.

[3] 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.

[4] A., Darwish, N.R. and Hefni, H.A., 2019. An agile methodology for ontology development. International Journal of Intelligent Engineering and Systems, 12(2), pp.170-181.

[5] 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.

[6] 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.

[7] 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.

[8] A. Gangemi, C. Catenacci, M. Ciaramita, J. Lehmann, R. Gil, F. Bolici, O. Strignano, Ontology evaluation and validation: An integrated formal model for the quality diagnostic task. Technical report, Lab. for Applied Ontology, 2005.

[9] A. Gangemi, Ontology Design Patterns for Semantic Web Content, in: Y. Gil, E. Motta, V.R. Benjamins, M.A. Musen (Eds.), The Semantic Web – ISWC 2005, Springer, Berlin, Heidelberg, 2005: pp. 262–276. https://doi.org/10.1007/11574620_21.

[10] A. Gangemi, C. Catenacci, M. Ciaramita, J. Lehmann, Modelling Ontology Evaluation and Validation, in: Y. Sure, J. Domingue (Eds.), The Semantic Web: Research and Applications, Springer Berlin Heidelberg, Berlin, Heidelberg, 2006: pp. 140–154.

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

[12] 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.

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

[14] A. Gomez-Perez, M. Fernandez-Lopez, A. de Vicente, Towards a Method to Conceptualize Domain Ontologies, in: ECAI96 Workshop on Ontological Engineering, Budapest, 1996, pp. 41–51.

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

[16] A. Gyrard, G. Atemezing, M. Serrano, PerfectO: An Online Toolkit for Improving Quality, Accessibility, and Classification of Domain-based Ontologies. Semantic IoT: Theory and Applications: Interoperability, Provenance and Beyond, (2021), p.161.

[17] A.I. Walisadeera, A. Ginige, G.N. Wikramanayake, User centered ontology for Sri Lankan farmers, Ecological Informatics. 26 (2015) 140–150.

[18] Association for Ontology Design & Patterns (ODPA), OP-Types - Odp, Ontology Design Patterns . Org (ODP). (2021). http://ontologydesignpatterns.org/wiki/OPTypes (accessed December 29, 2021).

[19] A. Takhom, S. Usanavasin, T. Supnithi, P. Boonkwan, A Collaborative Framework Supporting Ontology Development Based on Agile and Scrum Model, IEICE Trans. Inf. & Syst. E103.D (2020) 2568–2577.

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

[21] 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).

[22] 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.

[23] 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.

[24] B. Behkamal, M. Kahani, E. Bagheri, Z. Jeremic, A Metrics-Driven Approach for Quality Assessment of Linked Open Data, Journal of Theoretical and Applied Electronic Commerce Research. 9 (2014) 16.

[25] 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.

[26] 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.

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

[28] 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.

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

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

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

[32] B. Motik, P.F. Patel, B. Parsia, OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax (Second Edition), (2012). https://www.w3.org/TR/owl2-syntax/ (accessed December 29, 2021).

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

[34] B. Swartout, P. Ramesh, K. Knight, T. Russ, Toward Distributed Use of Large-Scale Ontologies, AAAI, Symposium on Ontological Engineering, Stanford, 1997.

[35] C. Batini, M. Scannapieco, Data Quality: Concepts, Methodologies and Techniques. Springer-Verlag New York, Inc., 2006.

[36] C. Del Vescovo, P. Klinov, B. Parsia, U. Sattler, T. Schneider, D. Tsarkov, Empirical study of logic-based modules: Cheap is cheerful, in: Proceedings of the 26th International Workshop on Description Logics (DL'13), CEUR Workshop Proceedings, Ulm, Germany, (2013): pp. 144–155.
[37] C. Fürber, M. Hepp, SWIQA – A semantic web information quality assessment framework, in CIS 2011 Proceedings, 2011, 76.
[38] credibility, https://dictionary.cambridge.org/dictionary/english/credibility (accessed June 26, 2021).
[39] C. Roussey, O. Corcho, L.M. Vilches-Blázquez, A catalogue of OWL ontology antipatterns, in: Proceedings of the Fifth International Conference on Knowledge Capture, Association for Computing Machinery, New York, NY, USA, 2009: pp. 205–206.
[40] D.B. Lenat, R.V. Guha, Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project, Addison-Wesley, Boston, 1990.
[41] D.D. Kehagias, I. Papadimitriou, J. Hois, D. Tzovaras, J. Bateman, A Methodological Approach for Ontology Evaluation and Refinement, (2008) 12.
[42] D. Garvin, What Does 'Product Quality' Really Mean?, in: Sloan Management Review, Fall, 1984, pp. 25–45.
[43] D. Sánchez, Semantic variance: An intuitive measure for ontology accuracy evaluation, Engineering Applications of Artificial Intelligence. (2015) 11.
[44] 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).
[45] D. Tsarkov, Improved algorithms for module extraction and atomic decomposition, in: The International Workshop on Description Logics (DL'12), Rome, Italy, 2012: pp. 7–10.
[46] D. Vrandečić, Ontology evaluation, in: Handbook on Ontologies, Springer, 2009: pp. 293–313.
[47] D. Vrandečić, A. Gangemi, Unit Tests for Ontologies, in: R. Meersman, Z. Tari, P. Herrero (Eds.), On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops, Springer, Berlin, Heidelberg, 2006: pp. 1012–1020.
[48] D.W. Surry, Bringing the User into the Project Development Process, in: Handbook of Research on Technology Project Management, Planning, and Operations. edited by Kidd, Terry T., 105-113. Hershey, PA: IGI Global, 2009.
[49] E. Blomqvist, A. Gangemi, V. Presutti, Experiments on Pattern-Based Ontology Design, in: Proceedings of the Fifth International Conference on Knowledge Capture, Association for Computing Machinery, New York, NY, USA, 2009: pp. 41–48.
[50] E. Daga, XDTools - NeOn Wiki, (2012). http://neon-toolkit.org/wiki/XDTools.html (accessed December 29, 2021).
[51] 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.
[52] 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).
[53] 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.
[54] G. Flouris, D. Manakanatas, H. Kondylakis, D. Plexousakis, G. Antoniou, Ontology change: classification and survey, Knowledge Eng. Review. 23 (2008) 117–152.
[55] H. Kondylakis, A. Nikolaos, P. Dimitra, K. Anastasios, K. Emmanouel, K. Kyriakos, S. Iraklis, K. Stylianos, N. Papadakis, Delta: A Modular Ontology Evaluation System, Information, 12 (2021), 301.
[56] H. S. Pinto, S. Staab, C. Tempich,, DILIGENT: Towards a fine-grained methodology for DIstributed, Loosely-controlled and evolvInG Engineering of oNTologies, 16 (2004), p. 393.
[57] H. Tan, A. Adlemo, V. Tarasov, M. E. Johansson, Evaluation of an Application Ontology, in The Joint Ontology Workshops, 2017.
[58] 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.
[59] IEEE 730-2014 - IEEE Standard for Software Quality Assurance Processes. https://standards.ieee.org/standard/730-2014.html (accessed June 26, 2021).
[60] ISO 25012. https://iso25000.com/index.php/en/iso-25000-standards/iso-25012 (accessed June 26, 2021).
[61] 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).
[62] 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).
[63] 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).
[64] 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.
[65] J. Brank, M. Grobelnik, D. Mladeni, A survey of ontology evaluation techniques, in: Proc. Conf. Data Min. Data Warehouses, 2005, pp. 166–170.
[66] J. Cavano, J. McCall, A framework for the measurement of software quality, ACM SIGMETRICS Performance Evaluation Review 7 (1978) 133–139.
[67] J. Defeo, Juran's Quality Handbook: The Complete Guide to Performance Excellence, Seventh Edition, 7th edition, McGraw Hill, New York, 2016.
[68] J. Evermann, J. Fang, Evaluating ontologies: Towards a cognitive measure of quality, Information Systems. 35 (2010) 391–403.
[69] 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.
[70] 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.
[71] J. Murdock, C. Buckner, C. Allen, Evaluating dynamic ontologies, in: Communications in Computer and Information Science, Springer Verlag, 2013, pp. 258–275.
[72] J. Pak, L. Zhou, A framework for ontology evaluation. In: Workshop on E-Business. 2009, pp. 10–18."

[73] 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 and Technology Publications, Lisbon, Portugal, 2015, pp. 179–186.
[74] 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
[75] J. Yu, J.A. Thom, A. Tam, Requirements-oriented methodology for evaluating ontologies, Information Systems 34 (8) (2009) 766–791.
[76] K. Hammar, Content Ontology Design Patterns: Qualities, Methods, and Tools, Linköping University Electronic Press, 2017.
[77] K.I. Kotis, G.A. Vouros, D. Spiliotopoulos, Ontology engineering methodologies for the evolution of living and reused ontologies: status, trends, findings and recommendations, The Knowledge Engineering Review. 35 (2020) e4.
[78] K. Schekotihin, P. Rodler, W. Schmid, Ontodebug: Interactive ontology debugging plug-in for protégé, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Springer Verlag, 2018: pp. 340–359.
[79] L. Bayoudhi, N. Sassi, W. Jaziri, How to Repair Inconsistency in OWL 2 DL Ontology Versions?, Data & Knowledge Engineering. 116 (2018) 138–158.
[80] 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.
[81] 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.
[82] M. Amith, Z. He, J. Bian, J.A. Lossio-Ventura, C. Tao, Assessing the practice of biomedical ontology evaluation: Gaps and opportunities, Journal of Biomedical Informatics. 80 (2018) 1–13.
[83] 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.
[84] M. Batet, D. Sánchez, A Semantic Approach for Ontology Evaluation, in: IEEE 26th International Conference on Tools with Artificial Intelligence, 2014, pp. 138-145.
[85] M.E. Aranguren, E. Antezana, M. Kuiper, R. Stevens, Ontology Design Patterns for bio-ontologies: a case study on the Cell Cycle Ontology, BMC Bioinformatics. 9 (2008) S1.
[86] 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.
[87] 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.
[88] M. Gruninger, M.S. Fox, Methodology for Design and Evaluation of ontologies, in: Workshop on Basic Ontological Issues in Knowlege Sharing, 1995. http://www.eil.utoronto.ca/wp-content/uploads/enterprise-modelling/papers/gruninger-ijcai95.pdf (accessed September 21, 2021).
[89] Michael Grüninger. 2013. Ontology Evaluation Workflow in COLORE. http://ontolog.cim3.net/file/work/ OntologySummit2013/2013-02-14_OntologySummit2013_OntologyEvaluation-SoftwareEnvironment/wip/ Gruninger_colore-summit_20130214a.pdf.
[90] M. Horridge, OWL 2 Validator, (2009). http://mowl-power.cs.man.ac.uk:8080/validator/ (accessed December 29, 2021).
[91] 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
[92] 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.
[93] 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.
[94] M. McDaniel, V.C. Storey, Evaluating Domain Ontologies: Clarification, Classification, and Challenges, ACM Comput. Surv. 52 (2019).
[95] 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.
[96] 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.
[97] M. Uschold, M. Gruninger, Ontologies: Principles, methods and applications,The Knowledge Engineering Review, 11 (1996), 93-136.
[98] M. Uschold, M. King, Towards a Methodology for Building Ontologies, in: IJCAI95 Workshop on Basic Ontological Issues in Knowledge Sharing, Montreal, 1995.
[99] M. Copeland, R.S. Gonçalves, B. Parsia, U. Sattler, R. Stevens, Finding Fault: Detecting Issues in a Versioned Ontology, in: P. Cimiano, M. Fernández, V. Lopez, S. Schlobach, J. Völker (Eds.), The Semantic Web: ESWC 2013 Satellite Events, Springer, Berlin, Heidelberg, 2013: pp. 113–124.
[100] 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.
[101] M. Cristani, R. Cuel, R., A survey on ontology creation methodologies. International Journal on Semantic Web and Information Systems (IJSWIS), 2 (2005), pp.49-69.
[102] 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.
[103] M.R. Khondoker, P. Mueller, Comparing ontology development tools based on an online survey, in: World Congress on Engineering (WCE 10), London, UK, 2010.
[104] 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.

[105] N. Bevan, Usability is Quality of Use, in: Advances in Human Factors/Ergonomics, Elsevier, 1995, pp. 349–354.
[106] 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.
[107] 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).
[108] O. Corcho, M. Fernández-López, A. Gómez-Pérez, Methodologies, tools and languages for building ontologies. Where is their meeting point?, Data & Knowledge Engineering. 46 (2003) 41–64.
[109] Ontology Summit 2013, http://ontolog.cim3.net/OntologySummit/2013/ (accessed June 26, 2021).
[110] OpenLink Software: Home, Virtuoso Universal Server. (n.d.). https://www.openlinksw.com/ (accessed December 29, 2021).
[111] OntoUML — OntoUML specification documentation, (2018). https://ontouml.readthedocs.io/en/latest/intro/ontouml.html (accessed December 29, 2021).
[112] 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.
[113] 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.
[114] 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.
[115] 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.
[116] P. Haase, L. Stojanovic, Consistent Evolution of OWL Ontologies, in: A. Gómez-Pérez, J. Euzenat (Eds.), The Semantic Web: Research and Applications, Springer, Berlin, Heidelberg, 2005: pp. 182–197.
[117] Hitzler, P., 2016. Anti-patterns in ontology-driven conceptual modeling: the case of role modeling in OntoUML. Ontology Engineering with Ontology Design Patterns: Foundations and Applications, 25, p.161.
[118] P. Lambrix, Q. Liu, Debugging the missing is-a structure within taxonomies networked by partial reference alignments, Data & Knowledge Engineering. 86 (2013) 179–205.
[119] 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.
[120] R. Dromey, Software Product Quality: Theory, Model, and Practice, Software Quality Institute, Brisbane, Australia, 1998.
[121] 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
[122] R.Q. Dividino, M. Romanelli, D. Sonntag, Semiotic-based Ontology Evaluation Tool (S-OntoEval), in: LREC, 2008.
[123] 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.
[124] R. Studer, V.R. Benjamins, D. Fensel, Knowledge engineering: Principles and methods, Data & Knowledge Engineering. 25 (1998) 161–197.
[125] R.S.I. Wilson, J.S. Goonetillake, W.A. Indika, A. Ginige, Analysis of Ontology Quality Dimensions, Criteria and Metrics, in: O. Gervasi, B. Murgante, S. Misra, C. Garau, I. Blečić, D. Taniar, B.O. Apduhan, A.M.A.C. Rocha, E. Tarantino, C.M. Torre (Eds.), Computational Science and Its Applications – ICCSA 2021, Springer International Publishing, Cham, 2021: pp. 320–337.
[126] R.Y. Wang, D.M. Strong, Beyond Accuracy: What Data Quality Means to Data Consumers, Journal of Management Information Systems, 12 (1996), 5–33.
[127] S. García-Ramos, A. Otero, M. Fernández-López, OntologyTest: A Tool to Evaluate Ontologies through Tests Defined by the User, in: S. Omatu, M.P. Rocha, J. Bravo, F. Fernández, E. Corchado, A. Bustillo, J.M. Corchado (Eds.), Distributed Computing, Artificial Intelligence, Bioinformatics, Soft Computing, and Ambient Assisted Living, Springer, Berlin, Heidelberg, 2009: pp. 91–98.
[128] S. Kumar, N. Baliyan, Quality Evaluation of Ontologies, in: Semantic Web-Based Systems, Springer Singapore, Singapore, 2018, pp. 19–50.
[129] S. Mc Gurk, C. Abela, J. Debattista, Towards Ontology Quality Assessment., in: MEPDaW/LDQ@ ESWC, 2017, pp. 94–106.
[130] S. Oh, H.Y. Yeom, J. Ahn, Cohesion and coupling metrics for ontology modules, Information Technology and Management. 12 (2011) 81.
[131] S. Peroni, A simplified agile methodology for ontology development, in: OWL: Experiences and Directions–Reasoner Evaluation, Springer, Cham, 2016, pp. 55-69.
[132] S. Staab, R. Studer, eds., Handbook on ontologies, Second Edition, Springer, Berlin Heidelberg, 2009.
[133] S. Staab, H.P. Schnurr, R. Studer, Y. Sure, Knowledge processes and ontologies, IEEE Intelligent Systems 16 (1) (2001) 26–34.
[134] 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.
[135] 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.
[136] T. Aruna, K. Saranya, C. Bhandari, A Survey on Ontology Evaluation Tools, in: 2011 International Conference on Process Automation, Control and Computing, 2011: pp. 1–5.
[137] The RDF Working Group, Resource Description Framework (RDF): Concepts and Abstract Syntax, (2004). https://www.w3.org/TR/rdf-concepts/ (accessed December 29, 2021).
[138] The OWL Working Group, OWL Web Ontology Language Guide, (2004). https://www.w3.org/TR/owl-guide/ (accessed December 29, 2021).
[139] 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.
[140] T.R. Gruber, A translation approach to portable ontology specifications, Knowledge Acquisition, 5(1993), 199-220.

[141] T. R. Gruber, Toward Principles for the Design of Ontologies Used for Knowledge Sharing, International Journal of Human Computer Studies, 43(1995), 907–928.
[142] V.R. Basili, G. Caldiera, H.D. Rombach, The goal question metric approach. Encyclopedia of Software Engineering 1 (1994), 528–532.
[143] W. Borst, Construction of Engineering Ontologies. PhD thesis, Institute for Telematica and Information Technology, University of Twente, Enschede, The Netherlands, 1997
[144] W.H. DeLone and E. R. McLean, The DeLone and McLean Model of Information Systems Success: A Ten-Year Update, J. Manag. Inf. Syst. 19 (2003) 9–30.
[145] W3C RDF Validation Service, (2006). https://www.w3.org/RDF/Validator/ (accessed December 29, 2021).
[146] X. Zhu, J.W. Fan, D.M. Baorto, C. Weng, J.J. Cimino, A review of auditing methods applied to the content of controlled biomedical terminologies, J. Biomed. Inform. 42 (3) (2009) 413–425.
[147] Y. Ma, B. Jin, Y. Feng, Semantic oriented ontology cohesion metrics for ontology-based systems, Journal of Systems and Software. 83 (2010) 143–152.
[148] Y. Sure, M. Erdmann, J. Angele, S. Staab, R. Studer, D. Wenke, OntoEdit: Collaborative Ontology Development for the Semantic Web, in: I. Horrocks, J. Hendler (Eds.), The Semantic Web — ISWC 2002, Springer, Berlin, Heidelberg, 2002: pp. 221–235.
[149] Z.C. Khan, C.M. Keet, An empirically-based framework for ontology modularisation, Applied Ontology. 10 (2015) 171–195.

Appendix A

Citation	**Year**	**Title of the paper**
The papers selected upon the inclusion and exclusion criteria defined in the systematic review (Core Survey Papers)		
Evermann et al. [68]	2010	Evaluating ontologies: Towards a cognitive measure of quality
Ma et al. [147]	2010	Semantic oriented ontology cohesion metrics for ontology-based systems
Djedidi et al. [119]	2010	ONTO-EVO^{A}L an Ontology Evolution Approach Guided by Pattern Modeling and Quality Evaluation
Tartir et al. [134]	2010	Ontological evaluation and validation
Oh et al. [130]	2011	Cohesion and coupling metrics for ontology modules
Duque-Ramos et al. [6]	2011	OQuaRE: A SQuaRE-based Approach for Evaluating the Quality of Ontologies
Murdock et al. [71]	2011	Evaluating Dynamic Ontologies
Bouiadjra et al. [2]	2011	FOEval: Full Ontology Evaluation
Ouyang et al. [80]	2011	A Method of Ontology Evaluation Based on Coverage, Cohesion and Coupling
Gavrilova et al. [139]	2012	New Ergonomic Metrics for Educational Ontology Design and Evaluation
Schober et al. [44]	2012	OntoCheck: verifying ontology naming conventions and metadata completeness in Protégé 4
Haghighi et al. [114]	2013	Development and evaluation of ontology for intelligent decision support in medical emergency management for mass gatherings
Neuhaus et al. [51]	2013	Towards ontology evaluation across the life cycle
Sánchez et al. [43]	2014	Semantic variance: An intuitive measure for ontology accuracy evaluation
Poveda-Villalón et al. [95]	2014	OOPS! (OntOlogy Pitfall Scanner!): An On-line Tool for Ontology Evaluation
Alexopoulos et al [112]	2014	Towards Vagueness-Oriented Quality Assessment of Ontologies
Rico et al. [96]	2014	OntoQualitas: A framework for ontology quality assessment in information interchanges between heterogeneous systems
Batet et al. [84]	2014	A Semantic Approach for Ontology Evaluation
Radulovic et al. [53]	2015	SemQuaRE — An extension of the SQuaRE quality model for the evaluation of semantic technologies
Lantow [31]	2016	OntoMetrics: Putting Metrics into Use for Ontology Evaluation
McDaniel et al. [92]	2016	The Role of Community Acceptance in Assessing Ontology Quality
Hnatkowska et al. [27]	2017	Verification of SUMO ontology
Zhu et al. [58]	2017	Quality Model and Metrics of Ontology for Semantic Descriptions of Web Services
Tan et al. [57]	2017	Evaluation of an Application Ontology
Ashraf et al. [64]	2018	Evaluation of U Ontology
McDaniel et al. [93]	2018	Assessing the Quality of Domain Ontologies: Metrics and an Automated Ranking System
Kumar et al. [128]	2018	Quality Evaluation of Ontologies
Demaidi et al. [102]	2019	TONE: A Method for Terminological Ontology Evaluation
Amith et al. [83]	2019	Architecture and usability of OntoKeeper, an ontology evaluation tool
Franco et al. [87]	2020	Evaluation of ontology structural metrics based on public repository data
The papers that are relevant to specific characteristics (do not satisfy the inclusion and exclusion) / Milestone papers		
Gruninger et al. [88]	1995	Methodology for Design and Evaluation of ontologies
Gomez-Perez [15]	2001	Evaluation of ontologies
Burton-Jones et al. [3]	2005	A semiotic metrics suite for assessing the quality of ontologies
Haase et al. [116]	2005	Consistent Evolution of OWL Ontologies
Gangemi et al. [10]	2006	Modelling Ontology Evaluation and Validation
Stvilia [33]	2007	A model for ontology quality evaluation
Tsarkov [45]	2012	Improved algorithms for module extraction and atomic decomposition
Vescovo et al. [36]	2013	Empirical study of logic-based modules: Cheap is cheerful
Khan et al. [149]	2015	An empirically-based framework for ontology modularization
Bayoudhi et al [79]	2018	How to Repair Inconsistency in OWL 2 DL Ontology Versions?