publications

MDPI

Article

Design and Construction of an Ontology Model for Semantic Representation of Scientometric Indicators

Hasan Mahmoudi Topkanlo [1,2], Mehrdad CheshmehSohrabi [1,*] and Akram Fathian Dastgerdi [3]

[1] Department of Knowledge and Information Science, Faculty of Education and Psychology, University of Isfahan, Isfahan 81746-73441, Iran; hmahmoudi@semnan.ac.ir
[2] Department of Knowledge and Information Science, Faculty of Psychology and Educational Sciences, Semnan University, Semnan 35131-19111, Iran
[3] Knowledge and Information Science, Islamic World Science & Technology Monitoring and Citation Institute (ISC), Shiraz 71946-94171, Iran; fathian@isc.ac
* Correspondence: mo.sohrabi@edu.ui.ac.ir; Tel.: +98-9134269390

Abstract

This paper proposes a novel ontology, called SciOnt, for representing scientometric indicators. Scientometrics is the field that studies the quantitative aspects of science and scientific phenomena. Scientometric indicators are used to evaluate scientific publications and research. However, many indicators and their diversity make them challenging to understand and use. SciOnt addresses this challenge by providing a formal and structured representation of scientometric indicator knowledge. It defines classes, relationships, and instances of these indicators. The ontology can be used for various purposes, including improving the analysis and comparison of scientometric indicators, facilitating the development of new and more efficient indicators, and enhancing the understanding and criticism of existing indicators. The paper details the methodology used to design SciOnt, including identifying and classifying scientometric indicators, determining competency questions, and constructing the ontology with Protégé 5.5 software. Finally, the paper evaluates the ontology and visualizes the SciOnt ontology structure using various plugins.

Keywords: scientometric indicators; ontology; SciOnt; knowledge representation; semantic web

1. Introduction

Scientometrics is a new field that examines the quantitative aspects of science, scientific subjects and phenomena, scientific communication, research methods and development of science, and science and technology policies (Vinkler, 2010; Mingers & Leydesdorff, 2015). Critical functions of scientometrics include:

- Quantitative evaluation of scientific resources and publications;
- Assessment of the quality and impact of scientific publications;
- Examination of the inputs, outputs, and effectiveness of scientific publications;
- Quantitative and qualitative evaluation of scientific publication owners, including authors, researchers, universities, etc.;
- Study of various aspects of scientific communication between researchers and scientific works;
- Development of science and research policies and guidelines (Vinkler, 2010; Mingers & Leydesdorff, 2015; Waltman, 2016; Glänzel et al., 2019).

Check for updates

Academic Editor: Houqiang Yu

Received: 10 February 2025
Revised: 19 March 2025
Accepted: 7 September 2026
Published: 14 September 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Scientometrics studies the quantitative aspects of the process of science, science policy, and communication in science by focusing on the author, article, journal, and institution metrics, as well as understanding the related citations (Zakka et al., 2021). In this context, numerous indicators have been proposed in scientometric studies, each with its function and objectives.

According to Vinkler (2010), indicators are scientometric measures describing one or more aspects of science and scientific research. The diversity of indicators has led researchers to categorize them in previous studies. For instance, Mahmoudi Topkanlo and CheshmehSohrabi (2023) provided a comprehensive overview of their study's different categorizations proposed for scientometric indicators. In their view, classifying indicators makes understanding, studying, critiquing, and using them more accessible.

However, despite the categorizations and the presentation of these indicators in many resources and databases, it is still difficult to access all indicators in a single, organized, and structured manner. As Bonacin et al. (2016) point out, in today's world, accessing information about a concept or phenomenon is not a challenge; instead, it is the retrieval of relevant information in the right place and time in an easier way that is a challenge due to the unstructured nature of information repositories and heterogeneous data forms. This challenge also exists in accessing and retrieving information related to scientometric indicators.

One way to improve indicators' access and precision is to design an ontology of indicators. Ontologies are part of semantic web tools and play a significant role in knowledge representation, generation, and sharing (Aminu et al., 2022). Ontology is a clear and explicit model in the field of knowledge, including a set of concepts and relationships between them (Fathian Dastgerdi, 2023). An ontology is a modeling tool that organizes concepts and relationships between them at the knowledge and semantic level (Li et al., 2013). Ontology is a formal representation of a set of concepts of a subject area and the relationships between them, which is used to gain knowledge of the subject area and avoid the ambiguity of words, and is created by ontology representation languages like OWL and RDF (Fathian Dastgerdi, 2016). By using the concepts and terms used in a domain and identifying the relationships between these concepts, an ontology provides a shared and common understanding of the knowledge of a specific domain (Ceccaroni et al., 2004). Through the shared understanding created by ontologies, communication between people and organizations and collaboration between different systems can be improved (Uschold & Gruninger, 1996). Noy and McGuinness (2001) identify the following as the most important goals of ontology:

- Ability to analyze knowledge in a domain;
- Establishing terms and relationships between them;
- Separating knowledge from subject domains;
- Rearranging knowledge;
- Sharing and disseminating a common understanding of the knowledge structure among people and intelligent software.

Sugumaran (2016) also mentions that accurate information processing, information application, sharing, reuse of knowledge, and discovery of new knowledge are some of the goals of ontology creation. Ontology is "a formal and explicit specification of a shared conceptualization" (Gruber, 1993). An ontology provides terminology, central concepts, classification of concepts, and relationships between them in a domain (Gaševic et al., 2009). The following paragraphs enumerate the main ontology components (Slimani, 2015a):

a. Classes or concepts: Also called types or universals, these are a group of individuals that share common characteristics used in a broad sense. In the SciOnt ontology, the "Measurement Method", "Measurement Unit", "Measurement Tool", and "Measurement Environment" are some of the main classes.

b. Relations: In ontologies, relations describe how individuals (instances or particulars) are related. For example, in the SciOnt ontology, we related the two individuals "co-citation" and "co-authorship" with the "is related to" relationship (co-citation is related to co-authorship).
c. Instances: Instances are individuals that model concrete objects (people, proteins, machines). In the SciOnt ontology, the class "Journal Indicators" includes "Eigen Factors Score", "Journal Impact Factor", and "SJR Index" as individuals.

Noy and McGuinness (2001) explain some of the reasons for developing an ontology, including "to share a common understanding of the structure of information among people or software agents", "to enable reuse of domain knowledge", "to make domain assumptions explicit", "to separate domain knowledge from operational knowledge", and "to analyze domain knowledge". By designing an ontology for scientometric indicators, it is possible to introduce and present all existing indicators while outlining their semantic relationships, identifying their similarities and differences, strengths and weaknesses, and applications. Additionally, by providing general knowledge about them, the groundwork can be laid for evaluating and updating them and creating new, more efficient indicators. Moreover, researchers, scientometricians, librarians, universities, and research institutions can use the ontology of scientometric indicators. Therefore, this paper aims to provide an ontology-based approach to the semantic representation of the indicators and relationships between them in scientometrics. The SciOnt Ontology is the result of employing such an approach.

The rest of this paper is structured as follows. First, we reviewed the related literature. Then, we discussed the methodologies used in this study to identify and classify scientometric indicators and the SciOnt ontology construction method. The Results and Discussion Section 4 is divided into three parts, including "ontology evaluation", "comparison of SciOnt with existing ontologies", and "ontology visualization", in which we present the SciOnt evaluation methods and the visual display of concepts and relationships (RDF graphs) designed with popular Protégé visualization plugins. Finally, the Conclusions Section 5 is presented.

2. Literature Review

The significance and applicability of ontologies in knowledge modeling, organization, and representation have prompted researchers in the field of scientometrics to pay attention to the design of ontologies. In this regard, Archambault et al. (2011) designed a new journal ontology to simplify the production of bibliometric data. They proposed a new journal classification consisting of a hierarchical, three-level classification tree based on taxonomies. Categories were created based on journal classifications (ISI, CHI, ERA). Journals were assigned to particular categories via a hybrid approach, including algorithmic methods and expert judgment. This ontology is now freely available at https://www.science-metrix.com.

Mohammadi Ostani et al. (2018) conducted a study to design an ontology in scientometrics called ScientometricsOnt with the domain analysis approach. In this study, they first extracted concepts related to scientometrics from Persian books, articles, dictionaries, thesauruses, research projects, and student dissertations. They identified the relationships and properties in the next step. Finally, they designed the scientometricsOnt ontology with the Protégé 5 editor. The designed ontology in this study comprises 185 concepts in scientometrics classified into 11 main classes, 20 relationships, and 100 individuals. ScientometricsOnt can improve knowledge representation in scientometrics. This study does not provide a classification of indicators and only deals with the terms and concepts of the scientometrics field.

Furthermore, semantic web technologies are used in citation analysis as a branch of scientometrics. For example, Xiao et al. (2018) suggested a new citation analysis method based on semantic technologies (e.g., ontology and linked data) in three steps: citation ontology construction, publishing linked data of citation information, and citation analysis

method implementation. They examined the conditions and requirements of each step in the context of citation analysis. Finally, they presented a framework of an ontology-based citation knowledge service system with four core layers: (1) a data resource layer, (2) an ontology layer, (3) a semantic association layer, and (4) a functional layer.

Artificial Intelligence (AI) and Natural Language Processing (NLP) are essential in classifying citations into different classes. For example, Ihsan and Qadir (2021) developed the Citations' Context and Reasons Ontology (CCRO) and used an ontology-based approach to extract citation reasons. They defined ontology classes and properties based on two collections of citation sentences: (1) a publicly available corpus and (2) a manually curated collection. Finally, they measured the accuracy of both manual and automatic mapping.

In some studies, the need to use new tools in scientometric systems and indicators has also been addressed. For example, Lopez-Rodriguez and Ceballos (2022) conducted a study to model scientometric indicators using data mining techniques semantically. For this purpose, they used the CRISP-DM methodology (CRISP-DM), a data mining process model. This methodology consists of six stages: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment. From their point of view, their proposed model can be used for representation, discovery, search, and combination of scientometric indicators in different organizations, including universities. Kozitsyn (2023) proposed using Thematic Search Algorithms in scientometric systems. From this point of view, the large amounts of data in scientometric systems, the differences between different subject areas, and the ambiguities and shortcomings associated with scientometric indicators make using intelligent methods, including search algorithms, essential.

Motohashi et al. (2024) studied the reciprocal relationships between science and technology using text mining and clustering techniques. The study population was data on articles and patents from Japan. The findings suggest an increasing connection trend between science and technology, although the speed of growth of the science–technology link varies across different subject areas.

Nishikawa and Koshiba (2024) investigated the application of Large Language Models (LLMs), such as GPT, in the analysis of citation contexts. In contrast to traditional citation analysis, which regards all citations as equivalent, citation context analysis takes into account the semantic and contextual information that surrounds each citation. The findings indicate that, although LLMs exhibit commendable consistency, they fall short in predictive performance. The authors propose that LLMs may serve as a valuable complement to human annotators, particularly in situations where human resources are constrained.

Jiang (2025) presents an ensemble learning framework designed to enhance two key aspects of citation analysis: the classification of citation functions and identifying significant citations. This methodology effectively differentiates among various citation functions, including background references, methodological support, or contrasting roles, while concurrently filtering for citations that exert substantial impact on the scholarly narrative. The experimental results of this study demonstrate that the ensemble methods offer a more nuanced and robust analysis of citation behavior.

As can be seen, ontology has been used to represent knowledge in scientometrics. Using ontology to introduce and present various scientometric indicators can be very helpful for universities, research centers, librarians, and researchers. This is especially true since research has yet to be done in scientometrics indicator ontology. This research can be the basis for further research in this field.

3. Materials and Methods

The current research method consists of three parts as follows:

A. Identifying and classifying Scientometric indicators

Initially, documentary analysis was employed to identify scientometric indicators. To achieve this, various indicators used to assess scientific and technical publications were determined by examining diverse texts, articles, and resources in the Scopus and Web of Science databases. While some indicators have already been classified in the existing literature (Braun et al., 2006; Vinkler, 2009; Vîiu, 2016), we also identified additional categories by studying them more closely and considering their similarities, differences, or functions. In total, 31 categories were ultimately recognized. In the subsequent phase, the Delphi technique was employed to gather the expert opinions of professionals in the scientometrics fields concerning the identification and classification of indicators for scientific and technical publications. The Delphi panel comprised 25 distinguished faculty members and researchers, each with a minimum of five published papers in the fields of bibliometrics, scientometrics, informetrics, webometrics, and altmetrics. These experts were selected through purposive sampling to ensure a high degree of expertise and relevance. The diverse and extensive experience of the panel members provided invaluable insights, thereby enhancing the robustness and validity of our classification. The pivotal aspect of research conducted using the Delphi technique is achieving consensus among experts. In this regard, it is widely recommended to conduct data collection in three rounds to ensure thoroughness and reliability (Brady, 2015). In the present study, the Delphi technique was repeated in three rounds. By the third round, panel members had reached agreement on the indicators and their classifications. To assess the level of agreement among the panel members, Kendall's coefficient of concordance and one-sample t-test were utilized, both indicating a high degree of consensus. Consequently, the Delphi study results categorized the scientific indicators into 34 subclasses and eight principal classes: (1) measurement method, (2) measurement unit, (3) measurement tool, (4) measurement purpose, (5) measurement development, (6) measurement source, (7) measurability, and (8) measurement environment. Based on these findings, the initial model for the design of a scientometrics ontology was established (Mahmoudi Topkanlo & CheshmehSohrabi, 2023).

B. Competency Questions for Ontology Design and Validation and Linked Data Integration

To ensure the ontology meets its intended objectives and properly represents bibliometric indicators, a set of competency questions was formulated. These questions help define the ontology's scope, structure, and usability. The questions are categorized into four key areas:

1. Class and Subclass Identification
    - What are the "Scientometrics" indicators (to identify ontology classes)?
    - To which class and group of indicators does the subclass "Altmetrics_indicators" belong (to determine subclasses)?
2. Instances and Relationships
    - What are the Journal Indicators (to identify instances of classes)?
    - Find instances that are linked via the "isRelatedTo" object property (to discover relationships between instances).
    - Retrieve the data properties defined for the instance "journal_impact_factor" along with their values (to determine data properties).
3. Entity Definition and Metadata
    - What does "Measurement_unit" represent in this ontology (to define the application of entities)?
    - Which classes are associated with related references (to find references related to entities such as books, articles, theses, . . .)?
    - Which classes have labels defined in other languages using skos:prefLabel (to identify multilingual labels)?

4. Commonalities and Similarities
    - Which classes share common instances (to identify commonalities between classes)?
    - Find similar instances within the ontology (to detect synonyms and equivalent instances).
5. Linked Data Integration
    - To which datasets in the linked data environment is"Bradford's Law" linked (to identify the entities that have external links)?
    - Which datasets are linked by the entities of the SciOnt ontology (to identify the datasets that entities are linked to)?

These competency questions provide a structured approach to designing and validating the ontology, ensuring it aligns with real-world requirements, as well as the broader linked data environment.

C. The SciOnt Ontology construction method

The SciOnt ontology model provides a semantic representation of the indicators and relationships between them in scientometrics. Other ontologies in scientometrics and related domains can easily reuse this ontology. There are different ontology development tools such as WebOnto, OntoEdit, Protégé, OilED, Ontolingua, WebODE, etc.; The SciOnt ontology was constructed using Protégé desktop version 5.5.0, developed by Stanford University (Stanford, CA, USA).

Protégé is an ontology editor created at Stanford University and is very popular in Semantic Web and computer science research. Protégé is free, developed in Java, and its source code is released under a free license (the Mozilla Public License). Various formats, including RDF, RDFS, OWL, etc., can be created and edited with Protégé, and it is suitable for large-scale ontologies (Slimani, 2015b). It is open-source, easy to learn and use, standalone, and user-friendly, and it fully supports OWL language. It offers plugins for graph visualization of ontology and has a large user community (Mukabunani, 2017). Protégé has become the most widely used software for building and maintaining ontologies (Musen, 2015).

This paper employed the Noy and McGuinness (2001) methodology as a general approach to designing SciOnt ontologies. However, after creating instances, we added a step (matching individuals with BiDO) to this approach. Figure 1 indicates the Noy and McGuinness methodology for designing ontology.

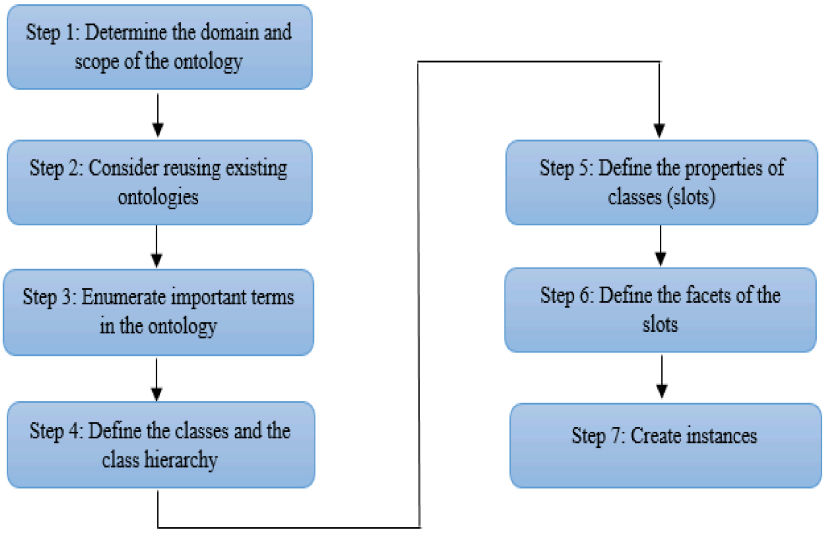


Figure 1. Noy and McGuinness methodology (Noy & McGuinness, 2001).

The following steps explain the SciOnt ontology designing method.

1. Defining classes and class hierarchical structure

We organized the classes into a hierarchical taxonomy, including classes and subclasses, according to the main concepts (indicators) identified in the Delphi panel. This ontology has 43 classes and subclasses that are organized into eight main classes including "Measurement Method", "Measurement Unit", "Measurement Tool", "Measurement Purpose", "Measurement Development", "Measurement Source", "Measurability", and "Measurement Environment". The main classes and the class hierarchical structure of the SciOnt ontology, which have been designed in Protégé, are shown in Figures 2 and 3, respectively.

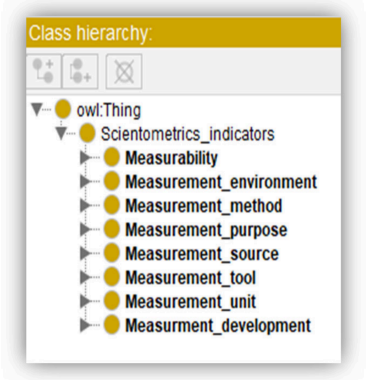


Figure 2. The main classes of SciOnt.

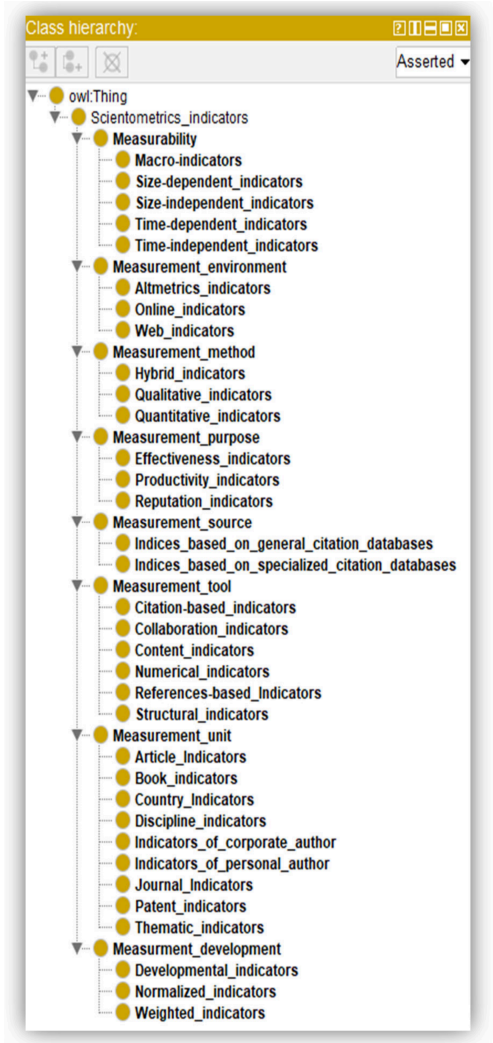


Figure 3. The class hierarchy of SciOnt.

According to Figure 3, the "Measurement Unit" has nine subclasses, including "Article Indicators", "Book Indicators", "Country Indicators", "Discipline Indicators", "Indicators of the Corporate Author", "Indicators of the Personal Author", "Journal Indicators", "Patent Indicators", and "Thematic Indicators". In ontologies, each class inherits the characteristics and properties of its higher class.

2. Defining the properties of classes

After designing the classes, the properties and relationships between them were described. In the hierarchical structure (Figure 3), the classes and subclasses are related by an "is-a" relationship. The "is-a" relationship, for example, indicates that the "Journal Indicators" class is a type of the "Measurement Unit" class. Furthermore, this ontology defines several "object properties," "data properties," and "annotation properties" to represent different relationships between instances. We defined the "is-related-to" object property to represent, for example, the relationships between three individuals, "degree centrality", "closeness centrality", and "betweenness centrality" (degree centrality is related to closeness centrality).

Furthermore, we defined seven data properties, including "alternateTitle", "dc:references", "dcterms:created", "dcterms:description", "doi", "hasURL", and "proposedBy", to relate the individuals to data types (e.g., integer, string, literal, dataTime, anyURI, language, etc.). For example, the data property "dcterms:description" relates the "Scientometrics Indicators" class to the "string" data type (Scientometrics Indicators dcterms:description xsd:string).

The SciOnt ontology also has some annotation properties such as "rdfs:comment", "skos:prefLabel", "schema:image", etc. Figure 4 shows some annotation properties used to describe the "Altmetrics Indicators" class.

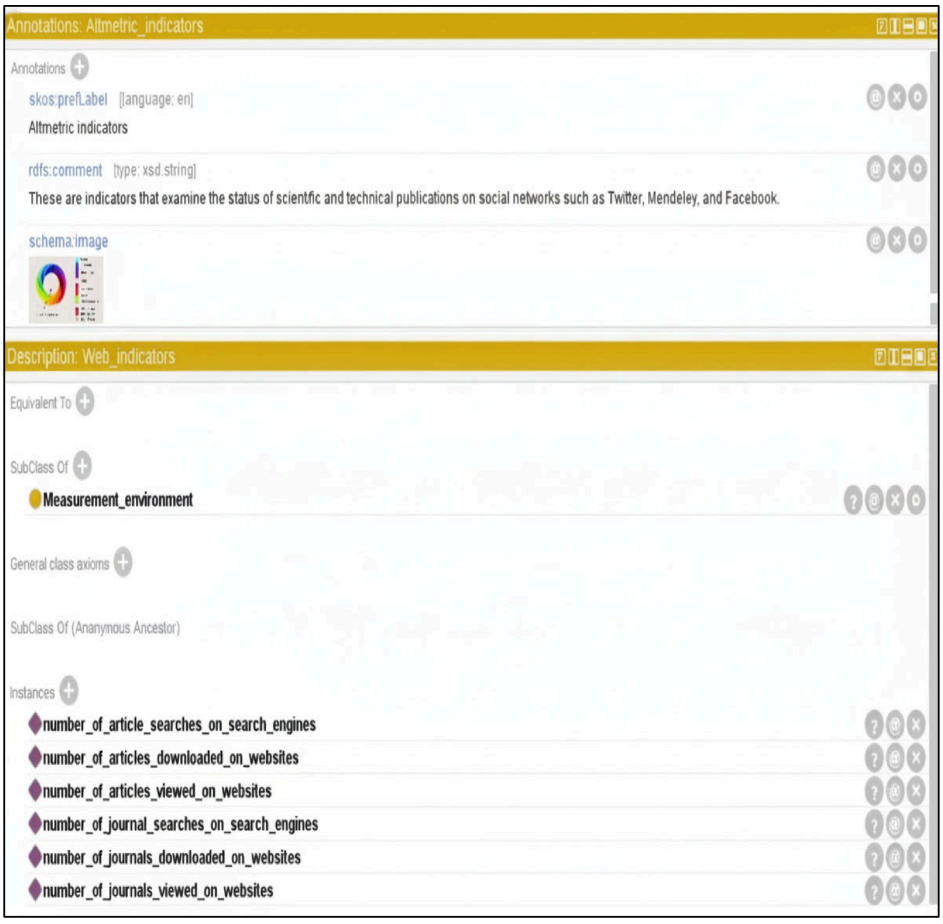


Figure 4. Annotation properties for the "Altmetrics Indicators" class.

3. Creating instances

Another part of ontology designing is adding instances and members to the defined classes. Each class contains several individuals that have common characteristics. The SciOnt ontology has 97 individuals assigned to the classes.

Figure 4 indicates the individuals for the "Altmetrics Indicators" class (e.g., "number of articles downloaded on social networks") at the bottom of the image. Another example is the "Journal Indicators" class which includes the individuals "Eigen Factors Score", "Journal Impact Factor", and "SJR Index". Figure 5 indicates how we specified object properties, data properties, annotation properties, and the type of "Journal Impact Factor" individual (which belongs to the "Journal Indicators" class).

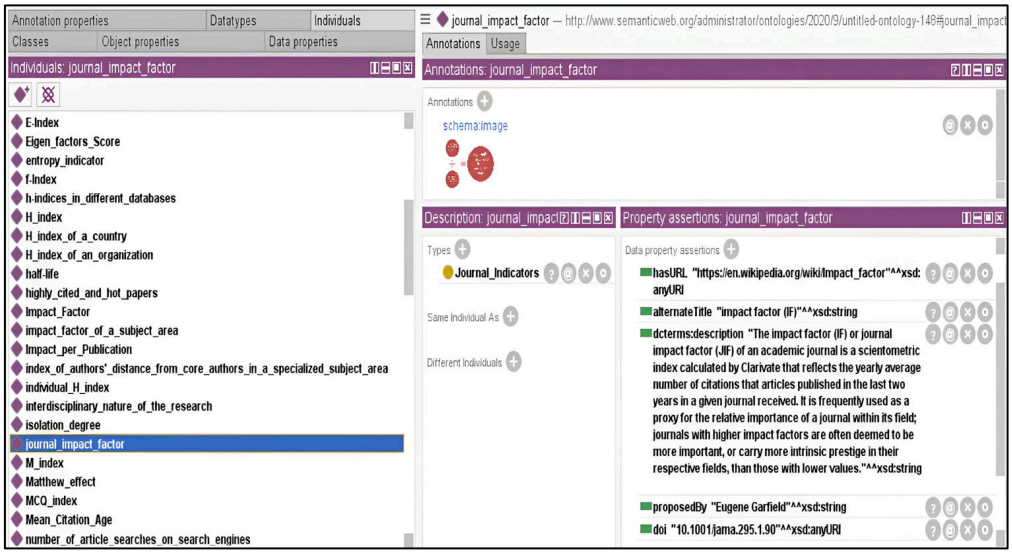


Figure 5. Specified properties for the "Journal Impact Factor" individual.

According to Figure 5, the type of individual is specified in the "Description: journal impact factor" window, which determines that this individual is a subclass of the "Journal Indicators" class.

Figure 5 provides an image of the "Journal Impact Factor" using the "schema:image" annotation property. Data and object property assertions can be defined in the "property assertions" window. So, the Wikipedia URL to the "impact factor" page was provided using the "URL" property. Another name for the individual (IF) was specified by the "alternateTitle" property, and the "dcterms:description" property was used to describe the individual clearly. Furthermore, we used "proposedBy" to identify the person who first introduced or proposed the "Journal Impact Factor" term and used the "created" property to indicate when this term was created or introduced (xsd:dateTime). As shown in Figure 5, the "Journal Impact Factor" was introduced by Eugene Garfield and discussed in various articles, blogs, etc. So, we represented the bibliographic information of the related references by the "dc:references" property and then defined the references DOIs by the "doi" property.

Another example of an individual represented in the SciOnt ontology is "Bradford's Law". Just like the previous individual (Journal Impact Factor), we defined various properties such as "owl:sameAs", "schema:image", "dcterms:description", "hasURL", "alternateTitle", "proposedBy", etc., for the "Bradford's Law" individual (Figure 6).

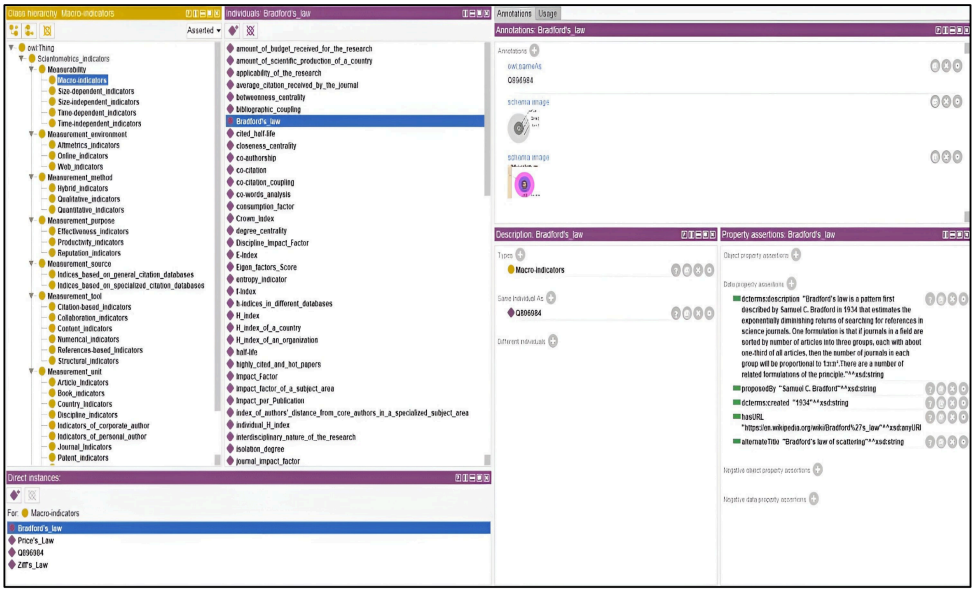


Figure 6. Specified properties for the "Bradford's Law" individual.

4. Matching individuals with BiDO (the Bibliometric Data Ontology[1])

The Bibliometric Data Ontology (BiDO) is a modular ontology that allows the description of numerical and categorical bibliometric data (e.g., journal impact factor, author h-index, E-index, etc.) in RDF. The BiDO is part of SPAR (Semantic Publishing and Referencing) ontologies that concern the specification of qualitative and quantitative evaluations of a bibliographic resource or an agent (Peroni & Shotton, 2018). The SPAR Ontology Network is a suite of orthogonal and complementary ontology modules to describe the scholarly publishing domain[2]. BiDO includes three distinct ontological modules:

- BiDO Core Module;
- BiDO Standard Bibliometric Measures module;
- BiDO Review Measures module.

We used the "BiDO Standard Bibliometric Measures module" to match the SciOnt individuals. This module describes a list of possible bibliometric measures that can be used to define bibliometric data of publication-related entities[3]. It includes four individuals:

"author citation counts", http://purl.org/spar/bido/author-citation-count (accessed on 18 November 2024)
"e-index", http://purl.org/spar/bido/e-index (accessed on 18 November 2024)
"h-index", http://purl.org/spar/bido/e-index (accessed on 18 November 2024)
"journal impact factor", http://purl.org/spar/bido/journal-impact-factor (accessed on 18 November 2024)

In this section, we matched the BiDO and SciOnt individuals to improve and enrich our ontology knowledge. Since all the BiDo individuals existed in the SciOnt ontology, we assigned the BiDO individuals' URIs to the same individuals in the SciOnt ontology.

5. Linking SciOnt to External Datasets

The way linked data enables connections between datasets facilitates the organization and the access of data on the web leading to enhanced web crawling and data retrieval. Integration of knowledge graphs and ontology with other datasets is made possible by Linked Data, which aids in enhanced representation of data and a higher degree of information retrieval across various fields.

While constructing SciOnt, we attempted to utilize linked data in order to improve ontology's interoperability. For example, we have established direct links to DBpedia and

Wikidata for aligning with vast accepted knowledge schemas. We also provided persistent identifiers, controlled vocabulary, and metadata descriptions for the entities using PURL, SKOS, and dcterms to achieve accurate and reliable access in the future.

We also applied the owl:sameAs property to relate certain instances of SciOnt to external datasets. This enables semantic entities in SciOnt to be aligned to their equivalents in existing knowledge bases thereby facilitating more integration and interoperability.

Linking SciOnt to external datasets makes it possible to add relevant contextual information that users can easily access. Furthermore, it facilitates federated queries, allowing researchers to seamlessly integrate SciOnt with other linked datasets for more comprehensive analysis. Ultimately, these connections enhance the ontology's effectiveness in representing and retrieving scientific knowledge.

4. Results and Discussion

This section is divided into two parts, including "ontology evaluation" and "ontology visualization", in which we presented the criteria-based ontology evaluation and the visual display of concepts and relationships (RDF graphs) designed with popular Protégé visualization plugins.

4.1. Ontology Evaluation

Ontology evaluation is the final stage of ontology design. The SciOnt ontology was evaluated using three key approaches: Conceptual Evaluation, Structural Evaluation, and Usability Evaluation (Raoufnezhad et al., 2016). These methods ensure the ontology's correctness, consistency, and practical effectiveness in managing and retrieving scientific data.

A. Conceptual Evaluation: Expert Review of SciOnt Ontology

The conceptual evaluation was performed by expert review, ensuring that the ontology accurately represents the intended domain. The evaluation may include clarity, consistency, extensibility, accuracy, completeness, comprehensibility, and the like (Haghighi et al., 2013). In this study, the criteria proposed by Xing et al. (2019) were used to evaluate the ontology, which includes five criteria—(1) Completeness, (2) Correctness, (3) Extensibility, (4) Conciseness, and (5) Clarity (transparency)—which evaluate the ontology in terms of content. To evaluate these criteria, we employed the focus group method. A focus group, as defined by Nyumba et al. (2018), is a qualitative research method that involves a small group of people discussing a specific topic under the guidance of a moderator. This method is used to gather in-depth insights and understand participants' attitudes, beliefs, and opinions. The focus group consisted of ten faculty members from the Information Science department who were experts in scientometrics and ontology. These experts provided their opinions after thoroughly observing and examining the ontology. Their feedback was incorporated iteratively throughout the ontology development process to refine class definitions, relationships, and hierarchy structures. Table 1 shows the ontology evaluation criteria.

B. Structural Evaluation

The technical aspects of ontology structure receive assessment in structural evaluation through systematic checks for consistency and logical coherence and data verification. The tasks required several evaluation tools to accomplish their purpose. Three primary reasoners including HermiT, Pellet, and FaCT++ served to confirm logical consistency. The ontology received an evaluation through OntoMetrics[4] to determine several structural properties. The RDF Validator[5] tool helped establish the right format in the RDF structure. The logical consistency was verified through the use of reasoners including HermiT, Pellet, and FaCT++. A series of issues arose which led to multiple problem solutions during this undertaking. The ontology contained several errors which involved selecting inappropriate

data types and defining improper domain and range values for properties along with violating class disjointness constraints. The ontology underwent multiple revisions using findings discovered through OntoMetrics and RDF Validator tools which helped verify logical coherence as well as structural soundness.

Table 1. Criteria-based ontology evaluation (Xing et al., 2019).

Row	Components	Statements
1	Completeness	- The classes are complete. - The class hierarchies are complete. - The domains and ranges of the relations defined in the class hierarchy are complete.
2	Correctness	The real-world concepts represented in the ontology are correct.
3	Extendibility	- The proposed ontology has the possibility of being extended to support risk identification applications involving other risks. - The proposed ontology has the possibility of being extended to be used in other specific application domains.
4	Conciseness	The proposed ontology does not include unnecessary concepts or redundancies.
5	Clarity	- Concepts are defined without subjectivity. - Concepts have been documented using natural language. - Concepts can convey their intended meanings that conform to the knowledge comprehension of users and computers.

C. *Usability Evaluation*

The SPARQL queries used competency questions to evaluate SciOnt both in terms of usefulness and effectiveness. The ontology achieved accurate data linking and retrieval results through querying and testing in the Protégé platform. The testing showed that SciOnt provides effective solutions to predefined competency questions, therefore validating its utility for scientific information retrieving and linking tasks. The successful three-step evaluation establishes SciOnt as a dependable system for scientific knowledge representation and retrieval because it has a sound concept and robust structure and functional effectiveness. Examples of the queries are provided in Figures 7–10.

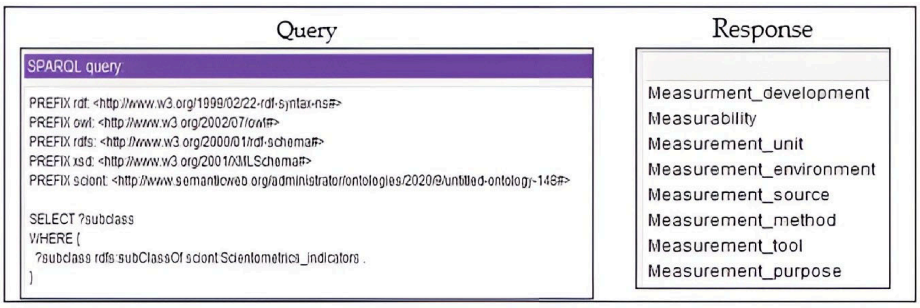


Figure 7. SPARQL query to identify ontology classes.

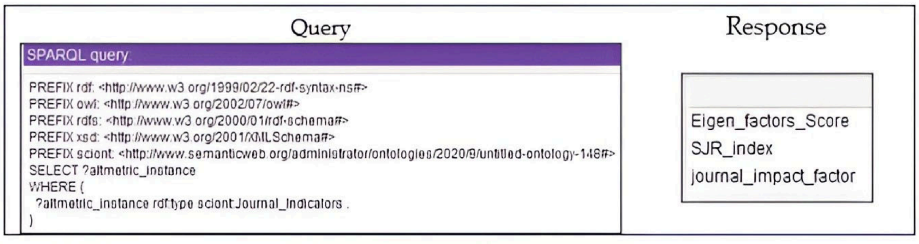


Figure 8. SPARQL query to identify instances of classes.

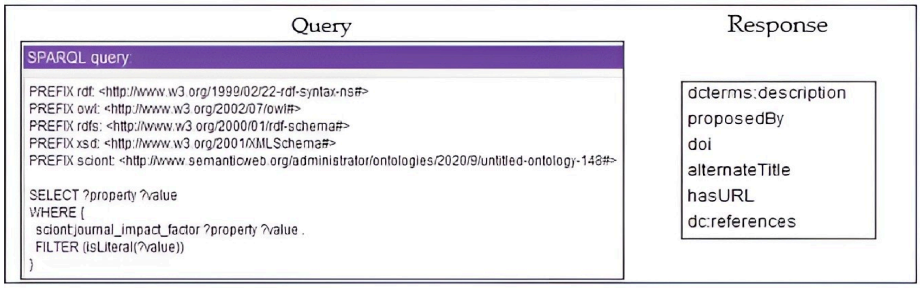


Figure 9. SPARQL query to determine data properties.



Figure 10. SPARQL query to identify instances that have the owl:sameAs property for linking to external resources.

4.2. Comparison of SciOnt with Existing Ontologies

The framework established by SciOnt adds more value to previous ontologies, like BiDO and ScientometricsOnt, through its extensive structured framework that supports interoperable scientometric analysis. The following section highlights the distinctions between SciOnt and previous ontologies.

A. Coverage of Scientometric Indicators

SciOnt provides a more comprehensive framework than BiDO and ScientometricsOnt by integrating both standard bibliometric measures and modern Altmetrics indicators, derived from resources in the Scopus and Web of Science databases. SciOnt also utilizes expert opinions for the development and improvement of indicators. This ensures a complete set of scientometric measurement capabilities, making it superior in covering diverse metrics, whereas the BiDO Standard Bibliometric Measures module is limited to four individuals focused solely on bibliometric indicators, all of which are included in SciOnt. Additionally, ScientometricsOnt derives its concepts exclusively from Persian resources (e.g., books and articles), limiting its applicability and comprehensiveness in a broader scientometric context. It is important to note that SciOnt specifically focuses on scientometric indicators, unlike some existing ontologies, such as ScientometricsOnt, which emphasize general concepts in the field of scientometrics. This approach makes SciOnt a more precise and comprehensive tool for modeling and analyzing scientometric indicators.

B. Validation Approach

SciOnt's classification system is validated through the Delphi method, which involves expert input and extensive documentary research. This validation approach improves accuracy and reliability, setting SciOnt apart from ontologies that lack such systematic validation. This is while BiDO and ScientometricsOnt have not utilized the Delphi method for validation.

C. Use of Linked Data

One of SciOnt's strongest advantages, compared to previous ontologies like ScientometricsOnt, is its adherence to linked data principles, which enhance interoperability. The use of properties like owl:sameAs allows SciOnt to connect with external knowledge

bases, such as DBpedia and Wikidata. Additionally, the incorporation of PURL, dcterms, SKOS, and Schema.org enhances interoperability and enables integration of linked data from different sources for scientometric purposes.

D. Logical Consistency

SciOnt undergoes rigorous structural and conceptual evaluations, ensuring logical consistency. Unlike other ontologies, which may lack structured validation methods, SciOnt guarantees correctness through a well-defined evaluation process. The ontology's logical consistency was validated using reasoners such as HermiT, Pellet, and FaCT++. Multiple revisions were conducted based on insights from OntoMetrics and RDF Validator tools, ensuring both logical coherence and structural integrity.

E. Practical Usability

SciOnt has been comprehensively assessed across multiple dimensions, including conceptual, structural, and usability evaluations. Additionally, SciOnt was evaluated through SPARQL queries, using competency questions. This multi-layered validation ensures that SciOnt is both theoretically sound and practically applicable, making it more effective in real-world scientometric analysis compared to other ontologies.

4.3. Ontology Visualization

This section indicates the SciOnt ontology structure of concepts and relationships with some of the Protégé visualization plugins, providing an easy-to-interpret visual display. Protégé provides various plugins for ontology visualization, such as OntoGraph, OWLPropViz, OWLViz, NavigOWL, etc. Figures 11–17 indicate the SciOnt graphs designed with the OntoGraph and OWLViz plugins and the WebVOWL Editor web application. OntoGraph and OWLViz Plugins are installed by default in version 5.5 of Protégé. WebVOWL is a web application for the interactive visualization of ontologies. It implements the "Visual Notation for OWL Ontologies (VOWL)" by providing graphical depictions for elements of the Web Ontology Language (OWL) that are combined to a force-directed graph layout representing the ontology[6]. Figure 11 indicates the graph view of the SciOnt ontology with the OntoGraph plugin.

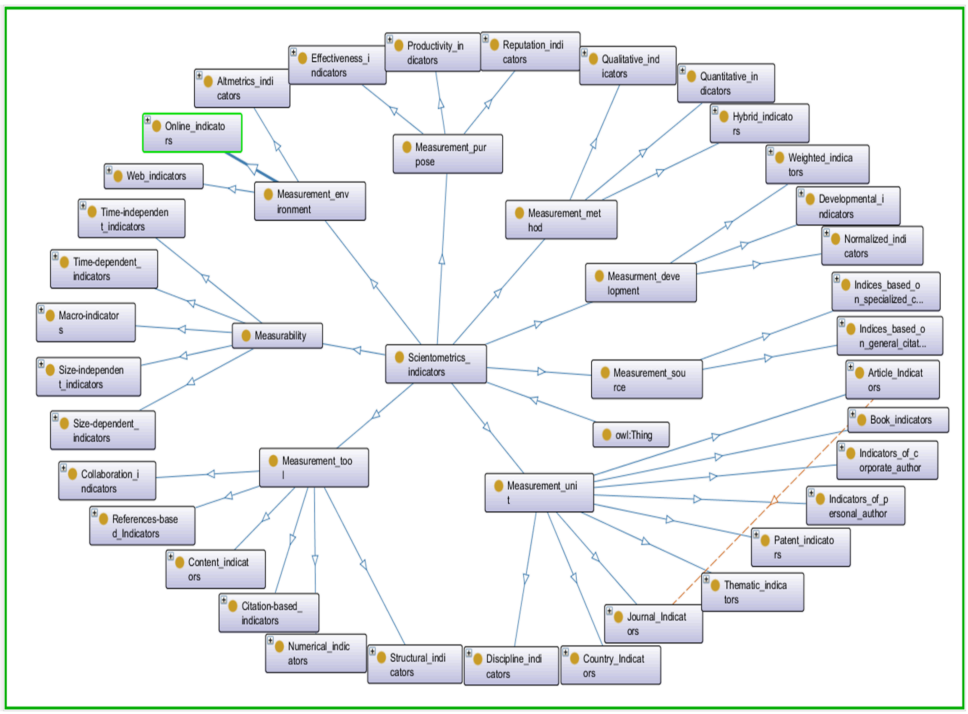


Figure 11. The graph view of the SciOnt ontology with OntoGraph.

Figure 12 indicates the graph view for the "Journal Indicators" class and its individuals designed by the OntoGraph plugin.

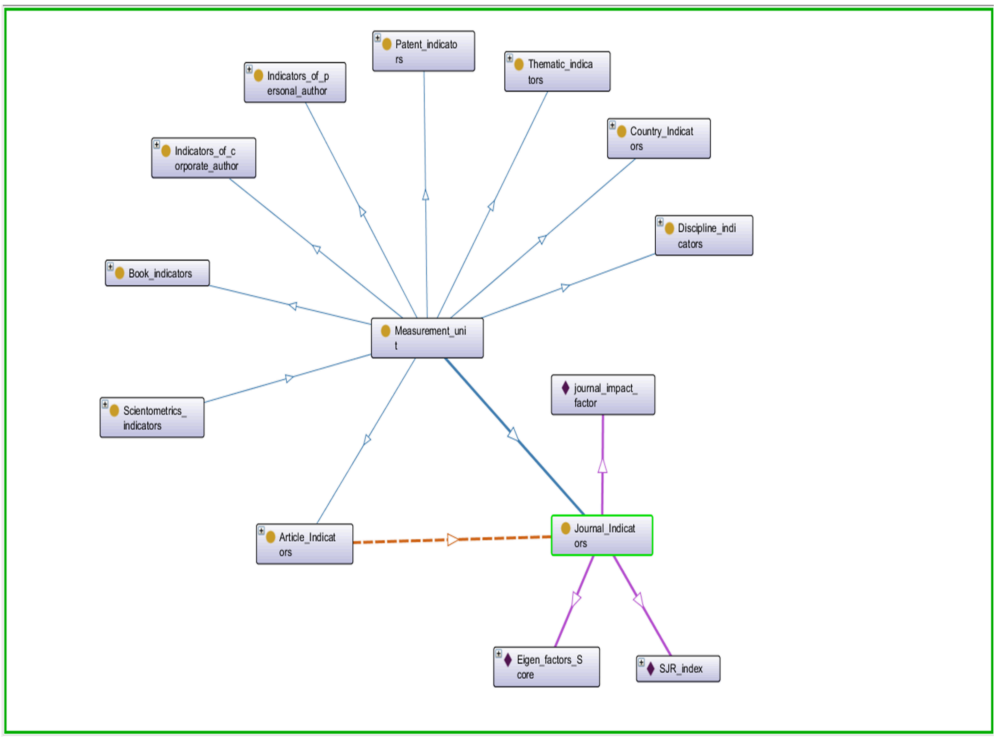


Figure 12. The graph view of the "Journal Indicators" class with OntoGraph—The yellow circular symbols represent classes, while the purple diamond-shaped symbols represent instances of the classes.

Figure 13 indicates the defined properties of the "Journal Impact Factor" individual.

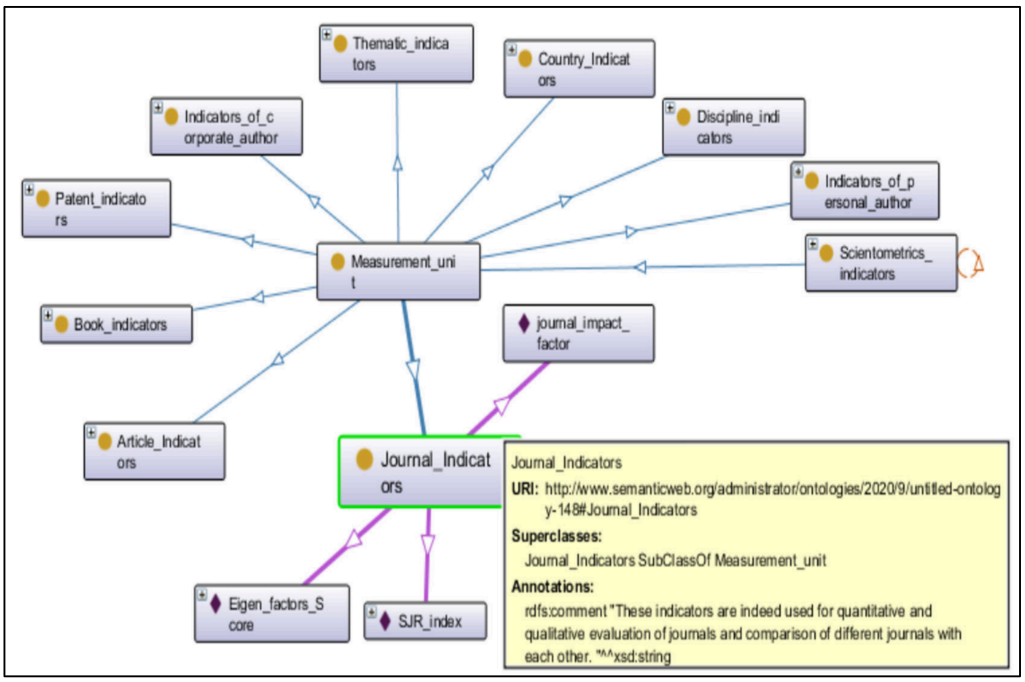


Figure 13. The graph view of the defined properties for the "Journal Impact Factor"—The yellow circular symbols represent classes, while the purple diamond-shaped symbols represent instances of the classes.

Figure 14 indicates the graph view of the SciOnt designed with the OWLViz plugin.



Figure 14. The graph view of the SciOnt ontology with the OWLViz—The different colors of the arrows in the figure represent different levels of the class–subclass hierarchy generated by the OWLViz plug-in.

We designed the graph view of the SciOnt ontology with the WebVOWL Editor (Figures 15–17). To develop the graphs with WebVOWL, we uploaded the RDF/XML file of the SciOnt ontology in WebVOWL and created the ontology graphs.

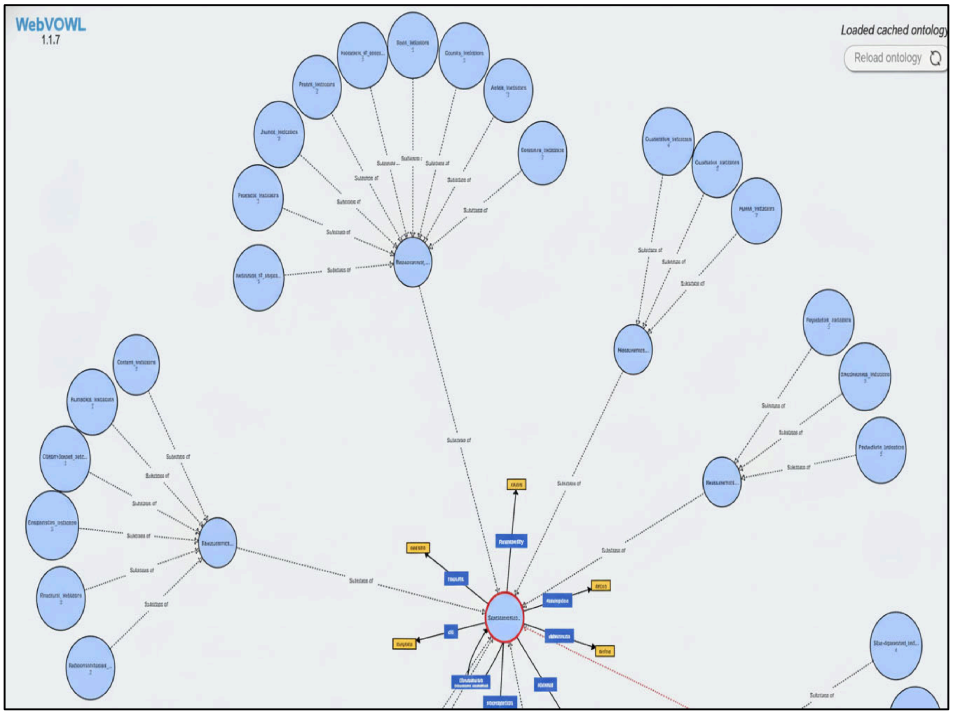


Figure 15. The graph view of the SciOnt ontology with the WebVOWL Editor.

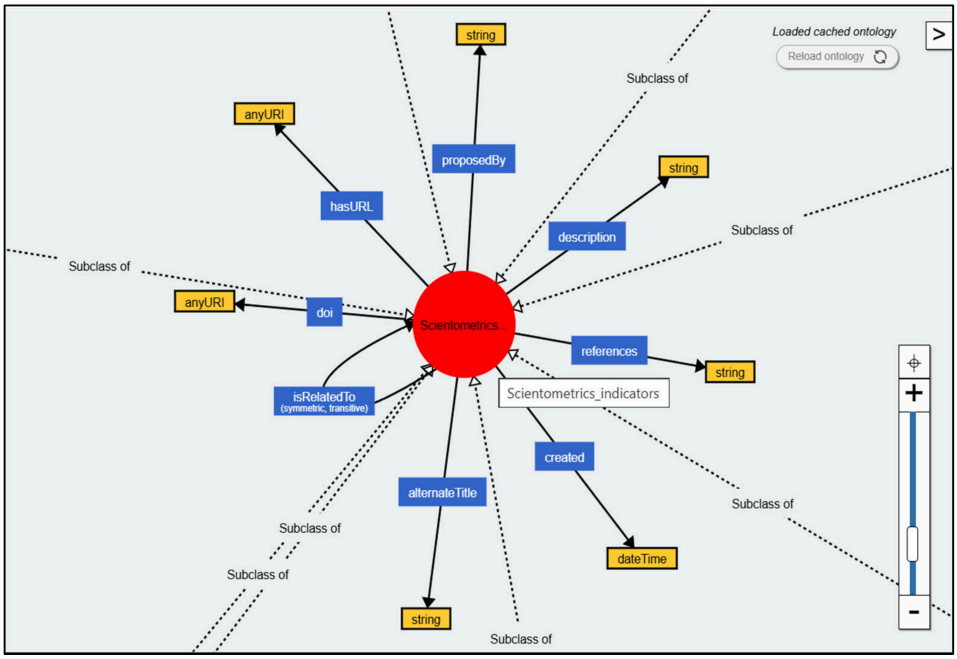


Figure 16. The graph view of the SciOnt ontology properties with WebVOWL.

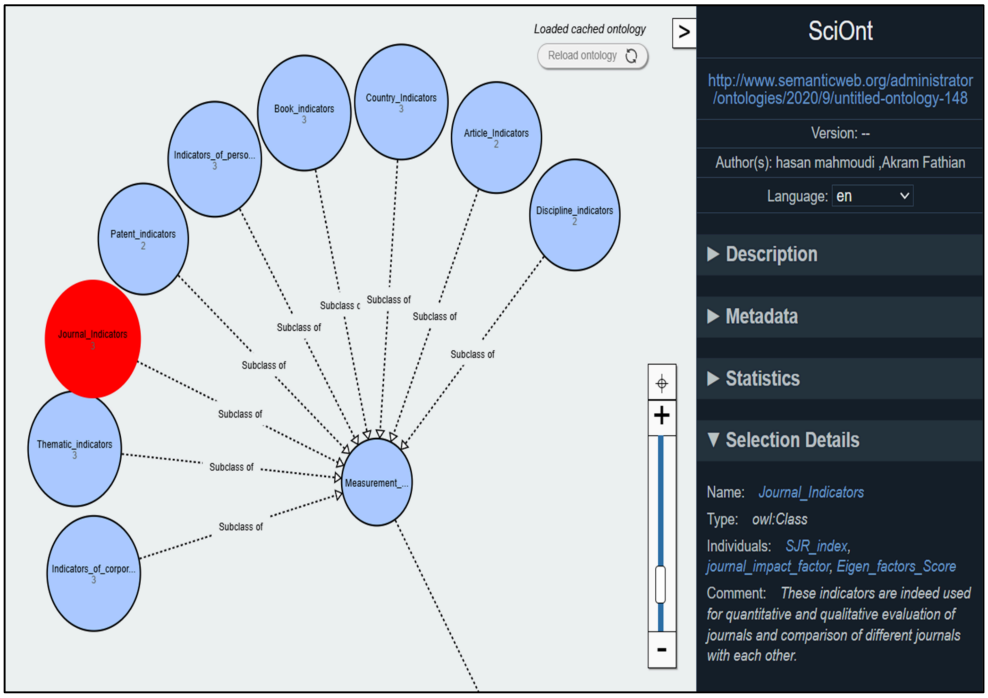


Figure 17. The graph view of the "Journal Indicators" class with WebVOWL.

5. Conclusions

Scientometric indicators are essential for analysis and decision-making related to science and scientific publication. The diversity of scientometric indicators makes it difficult to understand and study them. Semantic tools such as ontologies can help better understand scientometric indicators, analysis, criticism, and the development and creation of new indicators. Ontologies can be used to create a shared understanding of scientometric indicators, making it easier to analyze, compare, and criticize them. Ontologies can also be used to develop new scientometric indicators that are more accurate and relevant to the needs of researchers and policymakers. On this basis, the design of a scientometric indicator ontology was considered in this study. We designed the SciOnt ontology and concluded that using ontologies can

help overcome the challenges of understanding and using scientometric indicators. We proposed a scientometric indicator ontology that can be used to represent the knowledge about scientometric indicators in a formal and structured way. This ontology could then be used to improve scientometric indicators' analysis, criticism, and development.

The SciOnt ontology differs from the previous ontologies mentioned in this study in several ways and can be considered superior to them. First, the conceptual model of the ontology has been developed based on the perspectives of scientometric experts, and its validity and efficiency have been validated by both scientometric and ontology experts. Second, this model is fully comprehensive, allowing for the inclusion of all existing scientometric indicators. Third, the classification within the conceptual model has provided this ontology with a high degree of structure, enhancing its efficiency and capability. Fourth, in addition to the evaluation based on expert opinions, the applicability of the ontology has also been tested and validated in technical terms. Therefore, it can be stated that this ontology possesses high validity and efficiency, making it useful for improving the analysis, criticism, and development of scientometric indicators. The use of SciOnt can lead to the creation of new insights in the evaluation of research.

As a final point, recent studies (e.g., Fatemi et al., 2023; Reif et al., 2024; Doumanas et al., 2025; Val-Calvo et al., 2025) indicate that using LLMs and other deep learning models in ontology design can significantly improve their scope, performance, and efficiency. Therefore, it is recommended that future research on scientometric indicator ontology focus on this matter.

Author Contributions: Conceptualization, M.C.; methodology, M.C., H.M.T. and A.F.D.; validation, M.C.; formal analysis, H.M.T. and A.F.D.; investigation, H.M.T. and A.F.D.; resources, H.M.T. and A.F.D.; data curation, H.M.T.; writing—original draft preparation, H.M.T. and A.F.D.; writing—review and editing, M.C.; visualization, A.F.D.; supervision, M.C.; project administration, M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the University of Isfahan through a postdoctoral research agreement.

Data Availability Statement: The data that support the findings of this study are available on request from the corresponding author.

Acknowledgments: This work was supported by the University of Isfahan.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

OWL	Web Ontology Language
RDF	Resource Description Framework
SciOnt	Scientometric Indicator Ontology
LLMs	Large Language Models

Notes

1 http://purl.org/spar/bido (accessed on 18 November 2024).
2 http://www.sparontologies.net/ontologies (accessed on 16 November 2024).
3 https://sparontologies.github.io/bido-standard-bibliometric-measures/current/bido-standard-bibliometric-measures.html (accessed on 18 November 2024).
4 https://ontometrics.informatik.uni-rostock.de/ontologymetrics/index.jsp (accessed on 5 November 2024).
5 https://www.w3.org/RDF/Validator/ (accessed on 30 December 2024.
6 http://vowl.visualdataweb.org/webvowl.html (accessed on 25 December 2024).

References

Aminu, E. F., Oyefolahan, I. O., Abdullahi, M. B., & Salaudeen, M. T. (2022). MaCOnto: A robust maize crop ontology based on soils, fertilizers and irrigation knowledge. *Intelligent Systems with Applications*, *16*, 200125. [CrossRef]

Archambault, É., Beauchesne, O. H., & Caruso, J. (2011). Towards a multilingual, comprehensive and open scientific journal ontology. In B. Noyons, P. Ngulube, & J. Leta (Eds.), *Proceedings of ISSI 2011—13th international conference of the international society for scientometrics and informetrics (ISSI), Durban, South Africa, 4–7 July 2011* (Vol. 1, pp. 66–77). Leiden University and University of Zululand.

Bonacin, R., Nabuco, O. F., & Junior, I. P. (2016). Ontology models of the impacts of agriculture and climate changes on water resources: Scenarios on interoperability and information recovery. *Future Generation Computer Systems*, *54*, 423–434. [CrossRef]

Brady, S. R. (2015). Utilizing and adapting the Delphi method for use in qualitative research. *International Journal of Qualitative Methods*, *14*, 1609406915621381. [CrossRef]

Braun, T., Glänzel, W., & Schubert, A. A. (2006). Hirsch-type index for journals. *Scientometrics*, *69*, 169–173. [CrossRef]

Ceccaroni, L., Cortes, U., & Sanchez-Marre, M. (2004). OntoWEDSS: Augmenting environmental decision-support systems with ontologies. *Environmental Modelling & Software*, *19*, 785–797. [CrossRef]

Doumanas, D., Soularidis, A., Spiliotopoulos, D., Vassilakis, C., & Kotis, K. (2025). Fine-tuning large language models for ontology engineering: A comparative analysis of GPT-4 and mistral. *Applied Sciences*, *15*(4), 2146. [CrossRef]

Fatemi, B., Rabbi, F., & Opdahl, A. L. (2023). Evaluating the effectiveness of GPT large language model for news classification in the IPTC news ontology. *IEEE Access*, *11*, 145386–145394. [CrossRef]

Fathian Dastgerdi, A. (2016). *Designing the metadata ontology model of Iran's national library catalogue based on the linked data method* [Ph.D. dissertation, Ferdowsi University of Mashhad].

Fathian Dastgerdi, A. (2023). Creating the ontology of islamic medicine based on semantic web: A review study. *Journal of Modern Medical Information Sciences*, *9*, 296–311.

Gaševic, D., Djuric, D., & Devedžic, V. (2009). *Model driven engineering and ontology development* (pp. 46–57). Springer Science & Business Media.

Glänzel, W., Moed, H. F., Schmoch, U., & Thelwall, M. (2019). *Springer handbook of science and technology indicators* (pp. 17–40). Springer.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*, 199–221. [CrossRef]

Haghighi, P. D., Burstein, F., Zaslavsky, A., & Arbon, P. (2013). Development and evaluation of ontology for intelligent decision support in medical emergency management for mass gatherings. *Decision Support Systems*, *54*, 1192–1204. [CrossRef]

Ihsan, I., & Qadir, M. A. (2021). An NLP-based citation reason analysis using CCRO. *Scientometrics*, *126*, 4769–4791. [CrossRef]

Jiang, X. (2025). Ensembling approaches to citation function classification and important citation screening. *Scientometrics*, *130*(3), 1371–1419. [CrossRef]

Kozitsyn, A. S. (2023). Thematic search algorithms in scientometric systems. *Programming and Computer Software*, *49*, 325–332. [CrossRef]

Li, D., Kang, L., Cheng, X., Li, D., Ji, L., Wang, K., & Chen, Y. (2013). An ontology-based knowledge representation and implement method for crop cultivation standard. *Mathematical and Computer Modelling*, *58*, 466–473. [CrossRef]

Lopez-Rodriguez, V., & Ceballos, H. G. (2022). Modeling scientometric indicators using a statistical data ontology. *Journal of Big Data*, *9*, 9. [CrossRef]

Mahmoudi Topkanlo, H., & CheshmehSohrabi, M. (2023). Identification and classification of evaluation indicators for scientific and technical publications and related factors. *Information Research an International Electronic Journal*, *28*(1), 78–105. [CrossRef]

Mingers, J., & Leydesdorff, L. (2015). A review of theory and practice in scientometrics. *European Journal of Operational Research*, *246*, 1–19. [CrossRef]

Mohammadi Ostani, M., Azargoon, M., & Cheshmesohrabi, M. (2018). Methodology of construction and design of ontologies: A case study of scientometrics field. *Iranian Journal of Information Processing and Management*, *33*, 1761–1788. [CrossRef]

Motohashi, K., Koshiba, H., & Ikeuchi, K. (2024). Measuring science and innovation linkage using text mining of research papers and patent information. *Scientometrics*, *129*, 2159–2179. [CrossRef]

Mukabunani, A. (2017). *Ontology-based clinical decision support system applied on diabetes* [Master's thesis, University of Agder].

Musen, M. A. (2015). The protégé project: A look back and a look forward. *AI Matters*, *1*, 4–12. [CrossRef]

Nishikawa, K., & Koshiba, H. (2024). Exploring the applicability of large language models to citation context analysis. *Scientometrics 129*, 6751–6777. [CrossRef]

Noy, N. F., & McGuinness, D. L. (2001). *Ontology development 101: A guide to creating your first ontology*. Stanford University. Available online: http://protege.stanford.edu/publications/ontology_development/ontology101.pdf (accessed on 16 December 2024).

Nyumba, T. O., Wilson, K., Derrick, C. J., & Mukherjee, N. (2018). The use of focus group discussion methodology: Insights from two decades of application in conservation. *Methods in Ecology and Evolution*, *9*(1), 20–32. [CrossRef]

Peroni, S., & Shotton, D. (2018). The SPAR ontologies. In *The semantic web–ISWC 2018: 17th international semantic web conference, Monterey, CA, USA, 8–12 October, proceedings, part II 17* (pp. 119–136). Springer International Publishing.

Raoufnezhad, M., Kahani, M., & Maharati, Y. (2016, September). An ontology-based data model for Iranian research information. In *2016 eighth international conference on information and knowledge technology (IKT)* (pp. 18–23). IEEE.

Reif, J., Jeleniewski, T., Gill, M. S., Gehlhoff, F., & Fay, A. (2024, September). Chatbot-based ontology interaction using large language models and domain-specific standards. In *2024 IEEE 29th international conference on emerging technologies and factory automation (ETFA)* (pp. 1–4). IEEE.

Slimani, T. (2015a). A study investigating typical concepts and guidelines for ontology building. *Journal of Emerging Trends in Computing and Information Sciences*, *5*, 886–893. [CrossRef]

Slimani, T. (2015b). Ontology development: A comparing study on tools, languages and formalisms. *Indian Journal of Science and Technology*, *8*, 1–12. [CrossRef]

Sugumaran, V. (2016). Semantic technologies for enhancing knowledge management systems. In *Successes and failures of knowledge management* (pp. 203–213). Morgan Kaufmann. [CrossRef]

Uschold, M., & Gruninger, M. (1996). Ontologies: Principles, methods and applications. *The Knowledge Engineering Review*, *11*, 93–136. [CrossRef]

Val-Calvo, M., Aranguren, M. E., Mulero-Hernández, J., Almagro-Hernández, G., Deshmukh, P., Bernabé-Díaz, J. A., Espinoza-Arias, P., Sánchez-Fernández, J. L., Mueller, J., & Fernández-Breis, J. T. (2025). OntoGenix: Leveraging large language models for enhanced ontology engineering from datasets. *Information Processing & Management*, *62*(3), 104042. [CrossRef]

Vinkler, P. (2009). The π-index. A new indicator for assessing scientific impact. *Journal of Information Science*, *35*, 602–612. [CrossRef]

Vinkler, P. (2010). *The evaluation of research by scientometric indicators* (pp. 5–25). Elsevier.

Vîiu, G. A. (2016). A theoretical evaluation of Hirsch-type bibliometric indicators confronted with extreme self-citation. *Journal of Informetrics*, *10*, 552–566. [CrossRef]

Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, *10*, 365–391. [CrossRef]

Xiao, M., Shi, Z., & Wang, S. (2018). The impact on citation analysis based on ontology and linked data. In *Scientometrics* (pp. 196–208). IntechOpen. [CrossRef]

Xing, X., Zhong, B., Luo, H., Li, H., & Wu, H. (2019). Ontology for safety risk identification in metro construction. *Computers in Industry*, *109*, 14–30. [CrossRef]

Zakka, W. P., Lim, N. H. A. S., & Khun, M. C. (2021). A scientometric review of geopolymer concrete. *Journal of Cleaner Production*, *280*, 124353. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.