

A Semantic Collaboration Method Based on Uniform Knowledge Graph

Qi Li, Zehong Cao*, *Member, IEEE*, M. Tanveer, *Senior Member, IEEE*, H.M. Pandey, *Senior Member, IEEE*, and Chen Wang

Abstract—The Semantic Internet of Things is the extension of the Internet of Things and the Semantic Web, which aims to build an interoperable collaborative system to solve the heterogeneous problems in the Internet of Things. However, the Semantic Internet of Things has the characteristics of both the Internet of Things and the Semantic Web environment, and the semantic data presents many new data features. We analyze the characteristics of semantic data and propose the concept of uniform knowledge graph, which can be smoothly applied in the environment of the Semantic Internet of Things. In the meantime, a semantic collaboration method based on uniform knowledge graph is designed to take the uniform knowledge graph as the form of knowledge organization and representation, which provides a good data basis for semantic collaboration by constructing semantic links to complete semantic relation between different data sets. Then, the semantic collaboration in the Semantic Internet of Things can be achieved in our study. The experiments show that our method has the potential to analyze and understand the semantics of user requirements and provide more satisfactory results.

Index Terms—Semantic Internet of Things, semantic collaboration, knowledge graph, Internet of Things.

I. INTRODUCTION

With the rapid development of Internet of Things (IoT) technology [1], a variety of information sensing technology [2] has been widely used. Communication between devices has become the key to the study of Internet of Things technology. How to translate real-world data and objects into machine-readable information has become one of the key contents [3]. The semantic technologies [4] are of great significance to the extension of human perception and the establishment of a global collaborative ecosystem. The semantic technologies like Ontologies, Resource Description Framework (RDF), Resource Description Framework Schema (RDFS), Web Ontology Language (OWL), and Simple Protocol And RDF Query Language (SPARQL), semantic annotations, and semantic reasoning to interoperable IoT smart data. It also supports integrating, acquiring, unifying the sensors generated smart data. Data is represented by symbols, and symbols themselves

have no meaning. Data can only be used if these symbols are given special meanings. The purpose of semantic technology is to acquire the semantics of data, through natural language processing [5], data mining [6] and other technologies [7], to obtain the true meaning of data for machine understanding.

Tim Berners-Lee [8] first introduced semantic technology into the Internet in 1998. He proposed the concept of semantic web for intelligent information processing and extraction [9]. The semantic web is a network that adds semantic support on the basis of internet, it describes things in a machine-understandable way and enables machines to automatically understand the content of the World Wide Web.

The IOT is developing towards intellectualization, and people's expectations and demands for IOT are constantly improving. At present, the application of various terminals in the IOT is developing towards the demand of multi-equipment linkage and multi-data fusion [10]. These changes need to be reflected in the software management platform of the IOT. Through the intelligent analysis of large data transmitted by various IOT terminal devices [11], the statistical analysis results are presented to users or interact well with the terminal devices. It has become the focus of the development of the IOT in the direction of intellectualization. However, with the continuous development of the IOT, various problems are also highlighted [12]. In order to keep the research of the IOT in line with people's real-life needs. The IOT needs to develop in the direction of intellectualization and form accurate expression in information transmission and display. Semantic technology is introduced into the European FP7/IERC research project [13] to solve the operability of communication and information exchange between heterogeneous devices in the IOT [14]. The introduction of semantic technology will greatly improve the problem of insufficient machine understanding ability. Using the core of semantic technology (ontology) [15] to label the information of things semantically can provide users with more understandable information, so as to solve the problem of isomerization and enhance the ability to acquire related information, thus forming the Semantic Internet of Things (SIOT).

SIOT aims to build a global collaborative ecosystem [16]. Which integrates the objective entities of the physical world, terminal detection equipment, information transmission media and application analysis system to form a collaborative system of interconnection. Only by solving the heterogeneous problems in the development of the IOT, and developing towards the direction of intellectualization, can we truly realize the extension of human perception. In order to achieve the goal

Qi Li is with the Department of Computer Science and Engineering, Shaoxing University, Shaoxing, Zhejiang, 312000 China.

Zehong Cao is with the Department of Discipline of ICT, School of Technology, Environments and Design, University of Tasmania, Hobart, Australia. * Corresponding author (e-mail: Zehong.Cao@utas.edu.au or zhcaonctu@gmail.com)

M. Tanveer is with Discipline of Mathematics Indian Institute of Technology Indore, India.

H.M. Pandey is with Computing, Department of Computer Science, Edge Hill University, United Kingdom.

Chen Wang is with College of Computer Science, Chongqing University, Chongqing, China.

of the SIOT, we must realize the interoperability between these things and achieve the semantic collaboration in the SIOT.

The essence of semantic collaboration is to reduce data friction in the IOT [17], so that data is no longer limited to one website or application, and speed up data flow. Reducing data friction is mainly aimed at users. Users themselves provide data, but also need to consume data from the IOT. If the machine can not understand the meaning of the data, the data returned to the user must be a mechanical combination rather than an organic combination. After the user acquires the data, he needs to process and understand the data again through the human brain, which results in data friction between the human and the machine. At the same time, Because of the schema heterogeneity, data friction will also occur between machines. To speed up data flow is to let machines participate in the decision-making process, instead of part of the work of the human brain, to speed up decision-making.

In order to reduce data friction and accelerate data flow, knowledge graph technology [18] emerged as the times require. Firstly, knowledge graph is essentially a semantic network. It is constructed on the basis of semantic technology [19]. Nodes represent entities or concepts, while edges represent relationships between entities or concepts. In this way, the machine can understand the knowledge well. The graphical representation of data is conducive to the user's understanding. This can reduce data friction between human and machine. Meanwhile, By means of semantic mapping, knowledge graph can connect heterogeneous data sources and reduce data friction between machines. It links data from different data sources and formats to make data flow throughout the network, effectively facilitating and accelerating the data flow. Our main contributions can be summarized as follows:

- We analyzed the data characteristics of semantic data in the SIOT. The characteristics of semantic data represent the characteristics of the SIOT, which require to be considered in theoretical research and application. According to these characteristics, appropriate data organization and representation are adopted, for better use of the advantages of semantic data in representation and reasoning.
- According to the particularity of semantic data, we proposed a unified knowledge graph to organize and express the knowledge in the Semantic Internet of Things. Compared with the traditional ontology and semantic network, the unified knowledge map has higher entity coverage and more complex and comprehensive semantic relations. It achieves the transformation of the objective world from string description to structured semantic description.
- We design a semantic collaboration method based on the unified knowledge graph in the SIOT. This method achieves the corresponding data sources according to the user's needs, extracts the candidate entities (concepts) and their attribute associations needed to construct the knowledge graph, and forms isolated extraction graphs. Then, through semantic mapping, the semantic links between data sources are constructed, and a unified knowledge graph is finally formed. Supporting with the unified knowledge graph, we can share the semantic information and access interactions by semantic mapping.

- We choose a user query as a specific application environment to verify if the semantic collaboration method is effective. In the experimental verification stage, the standard ontology data set CoBra ontology is used as the experimental test data. Furthermore, the method is validated by the specific application scenario of the query system. The experimental results show that the semantic collaboration method based on the uniform knowledge graph can effectively achieve the sharing and interoperability of semantic data, and better mine the deep meaning of user needs, reduce redundant information and achieve high accuracy.

The rest of this paper is organized as follows; Section II, briefly outlined the details of related work and existing problems and challenges. Section III, the overall semantic collaboration framework based on the uniform knowledge graph is presented. Particularly, we mainly introduced the framework from four modules, including semantic Internet of things ontology parsing module, consistency maintenance module based on minimal adjustment set, construct module of unified knowledge graph based on semantic links and user requirement mapping module based on the semantic mapping. the experimental results and compared performance with other state-of-art methods is presented in Section IV. Section V, presents an example to illustrate our proposed model. Finally, this paper is concluded and addressed future work in Section VI.

II. RELATED WORK

The concept of semantic Internet of things [20]–[22] will play a great role in promoting the development of the Internet of things. Its main goal is to achieve resource integration, information sharing and interoperability in the Internet of things through semantic technology. In the traditional Internet of things, due to the openness, inclusiveness and flexibility of the Internet of things, the information obtained by the terminal equipment of the Internet of things has obvious personalized characteristics [23]. Different institutions, organizations and devices have different access to and expression of information. When this information is uploaded to the Internet, it has a negative impact on the analysis of data. Information providers are not consistent in their understanding of information, resulting in the diversification of information representations. The semantic Internet of Things adds semantic labeling to the information [24] in the Internet of Things, which makes the information on the Internet of Things clear lysine and facilitates the understanding of people and machines. That is to say, eventually information users can transform the information of objects into personal understandable knowledge according to their own knowledge background, and information users also have certain control ability over the information.

In recent years, many scholars have explored how to add semantics to the Internet of Things [21], and the Semantic Internet of Things is its representative system framework. In the research of semantic Internet of Things, many scholars have made some achievements. In 2003, Vagan Terziyan [25] believed that the current users of the Internet of Things were

not limited to humans, but that smart devices had emerged as new "users" in the Internet of Things. To achieve interoperability between these devices, a "global understanding environment" [26] needs to be achieved. Therefore, it is necessary to add semantic information to the data in the Internet of Things so that intelligent devices can process and utilize this information more intelligently.

Konstantinos Kotis [27], [28] believes that introducing semantic technologies into the environment of the Internet of things can effectively realize the integration of data in the semantic Internet of things. They propose to represent the information in the Internet of Things by ontology, so as to realize the interoperability of heterogeneous information in the Internet of Things by virtue of the advantages of ontology in knowledge fusion. Payam Barnaghi [29] believes that the heterogeneity of the Internet of Things makes interoperability between heterogeneous resources a huge challenge. Describing objects, sharing and integrating information through machine-understandable semantic technology, and reasoning out new knowledge together with other intelligent technologies will be an effective way to solve the heterogeneous problems in the Internet of Things. Michele Rute et al. [30] defined the semantic Internet of things as an organic combination of the Internet of things and the semantic web. Perera et al. [31] analyzed the context-aware applications are needed to propagate the sensor semantic data for a machine-to-machine interactions. Semantic web technologies can effectively integrate and reuse existing resources. However, the Internet of things needs to realize information sharing and interaction of things through pervasive computing. By embedding rich semantics into the Internet of things, the machine can effectively understand heterogeneous information and derive new information through logical reasoning.

Semantic collaboration is also called semantic interoperability. The purpose of semantic collaboration is to acquire semantic interoperability, that is, the ability of application entities to exchange information and understand information semantics correctly so as to use this information correctly. Paul Barom Jeon et al. [32] solved the semantic collaboration between heterogeneous machines by introducing semantic technology, and proposed a service semantic negotiation environment based on intelligent machines without user intervention. SOiA, as a service-oriented architecture of grammar (semantics), allows information integration of any type of object or device. The framework introduces the concept of context and the method of semantic tuple space to realize the interoperability between heterogeneous devices. Through the in-depth study of existing agent cluster methods, Salvatore Garruzzo [33] proposed a semantic collaboration method based on agent cluster in multi-agent environment. Through the concept and structure similarity matching of Agent ontology, the problem that communication between agents of the same group is difficult to understand each other is solved, and the semantic collaboration between agents is realized.

A. The extension of knowledge graph in SIOT

For the traditional knowledge graph, knowledge graph is the whole set of knowledge base, which not only needs to acquire

and store all the data in the possible data sources [34], but also needs to store the semantic association between these data. Because the data in the SIOT is dynamic and massive, it becomes a challenge to acquire and store semantic data in all data sources. Therefore, in order to organize and express knowledge in the form of knowledge graph in the SIOT, it is necessary to expand and transform the knowledge graph.

At the same time, the purpose of SIOT is to achieve semantic collaboration between heterogeneous systems, that is, to achieve semantic interoperability [35]. Semantic interoperability is to realize the sharing and interaction of semantic information. In order to achieve this goal, entities in the SIOT need to have the same knowledge background. If the various candidate entities (concepts) and their attribute associations required to build a knowledge graph are extracted from one or more data sources alone, only isolated extraction graphs can be formed. To form a true knowledge graph, we need to integrate these isolated extraction graphs together, which creates a unified knowledge graph. A unified knowledge graph is a representation of cross-domain common knowledge established across different fields.

Compared with traditional ontology and semantic network, unified knowledge graph has higher entity coverage and more complex and comprehensive semantic relationship. It realizes the knowledge mapping of the objective world from string description to structured semantic description.

III. SEMANTIC COLLABORATION FRAMEWORK BASED ON UNIFORM KNOWLEDGE GRAPH

SIOT is the extension and optimization of IOT. Its main purpose is to integrate resources, share information and interoperate with each other through semantic technology. SIOT uses the advantages of semantic technology in knowledge representation and knowledge sharing, and introduces semantic technology into the research field of IOT to solve the problem of resource interoperability. Its ultimate goal is to achieve semantic collaboration.

The use of uniform knowledge graph to organize and express knowledge in the SIOT is the basis for achieving semantic interpretability. Therefore, we use the unified knowledge graph to implement the semantic collaboration method in the SIOT. Under the condition of satisfying the characteristics of semantic data in the SIOT and giving full play to the advantages of these data characteristics, semantic data sharing and interoperability are realized. Fig.1 gives a framework of semantic collaboration based on unified knowledge graph. Fig. 2 represents data layer extraction and model layer extraction.

As shown in Fig.1, The framework takes knowledge graph as the organization and description means of information, and provides the bottom support and logic control for the upper application. In this framework, semantic collaboration is treated as a special upper application, so the framework still adopts the general bottom control logic based on the application of knowledge graph, following the general application rules, and has the relevant characteristics of the application of knowledge graph. However, knowledge graph is a special way of knowledge organization. Therefore, the data acquisition,

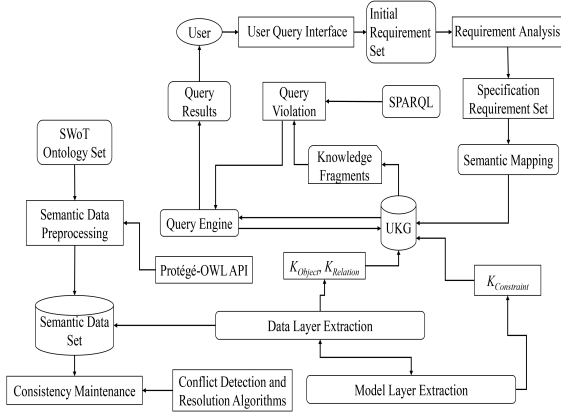


Fig. 1. Uniform knowledge graph based semantic collaboration framework in SIOT

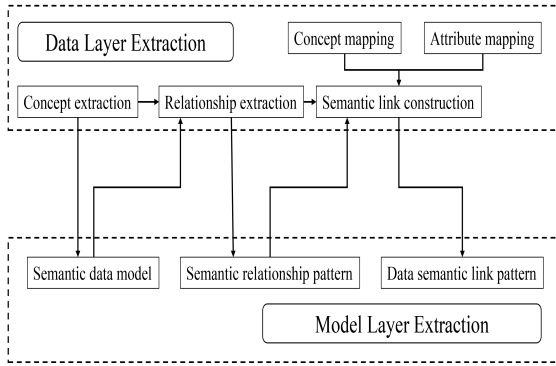


Fig. 2. Data layer extraction and model layer extraction

preprocessing and implementation of the framework still need to be redesigned to meet the special needs of the knowledge graph. The framework mainly includes the following four modules.

- 1) **SIoT ontology parsing module.** This module is the basis of semantic collaboration. Through loading and analyzing the ontology of SIOT, the concepts, attributes, instances and other related information in the ontology are obtained, which provides the data basis for semantic data interpretability.
- 2) **Consistency maintenance module based on minimal adjustment set.** This module is a necessary condition to achieve semantic collaboration. In order to get the results that meet the needs of users, it is necessary to maintain the consistency of the knowledge base to avoid errors caused by inconsistent information.
- 3) **The construct module of unified knowledge graph based on semantic links.** This module is the bottom support to realize semantic collaboration. It obtains the relationship between the concepts and attributes of SIoT ontology provided by ontology parsing module. The concepts, attributes and instances in the knowledge graph are related through relationships to form extraction graphs. By constructing semantic links, these extraction graphs are linked to each other to form a final uniform knowledge graph, which provides support for semantic

collaboration.

- 4) **User requirement mapping module based on semantic mapping.** This module is an application practice of semantic collaboration. It obtains the user's requirement set and analyses it to form a standard requirement description. By means of semantic mapping, the user requirements are mapped to the unified knowledge graph, and the user requirements are expanded, so as to better understand the user requirements and provide users with more satisfactory results.

In the framework, the various levels and functional modules are connected mainly through four functional operations, which enable the interface, interaction and control between the modules. It can also be performed in such as relationship discovery and semantic link construction. From the practical programming point of view, each functional module provides the corresponding interface to the upper module. Through these interfaces, resources in the framework can be called, organized and generated, which is related to whether semantic collaboration can be realized. It is the core of the framework.

The framework is the basic method of constructing knowledge graph. With semantic data in the SIOT as the object, the process of program execution from the semantic data from the bottom-up, and gradually realize the sharing and exchange of semantic data, conform to the general processing of data sharing and exchange, and finally realize the semantic collaboration goal of SIOT. Nature combines the construction of knowledge graph and the goal of semantic collaboration, and better plays the advantage of knowledge graph to solve the problem of semantic collaboration.

The key steps and specific algorithms involved in the semantic collaborative approach based on uniform knowledge graph are described in detail below.

A. Preprocessing of Semantic Data

In view of the semantic data characteristics of the Semantic Internet of Things, firstly, the ontology data of the Semantic Internet of Things need to be preprocessed. The main purpose of semantic data preprocessing is to eliminate the heterogeneity of semantic data. Semantic collaboration is essentially a knowledge-driven process. Existing knowledge needs to be constantly updated and organized with the participation of users. Therefore, this paper adopts OWL as the standard format of source data. OWL format is a declarative way of semantic description, which can better adapt to the uncertainty in the process of semantic collaboration. At the same time, OWL is based on RDF and RDFS proposed by W3C. It can be expressed by XML schema, which also provides syntactic consistency for collaborative process.

Semantic data preprocessing is mainly divided into two steps: SIoT ontology parsing and semantic consistency maintenance. Ontology parsing provides data basis for subsequent processing, while semantic consistency maintenance is the necessary condition for the final return of results that meet user needs.

Ontology parsing module is the basis of semantic collaboration model. In this paper, Jena API is used to obtain

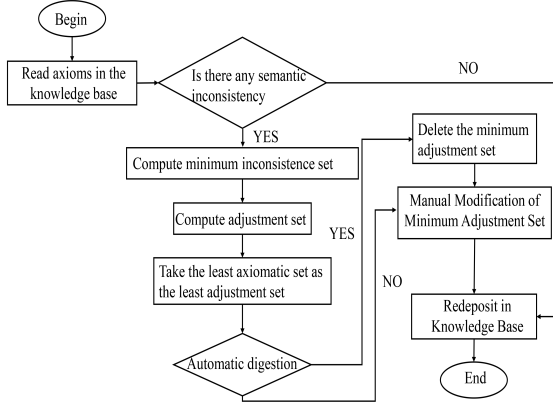


Fig. 4. Flow chart of semantic consistency maintenance

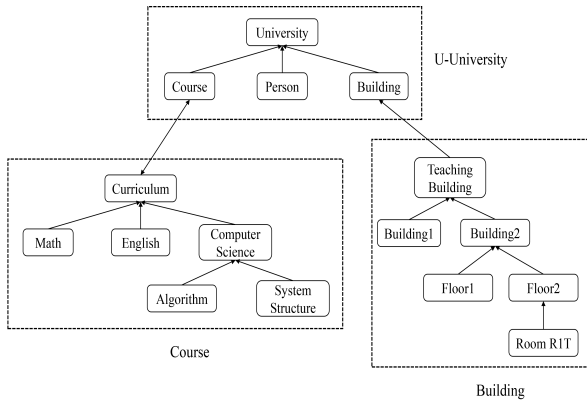


Fig. 5. Instance of semantic linked data

B. Construction process of uniform knowledge graph

In the ontology parsing module in the previous section, various candidate entities (concepts) and their attribute associations needed to construct the knowledge graph are extracted, thus the extracted graphs are formed. However, these extracted graphs are isolated from each other and have no connection with each other. In order to achieve the consistency of knowledge understanding among collaborative ontologies, it is necessary to establish links between extracted graphs, and to share and reuse knowledge among collaborative entities by sharing concepts. The knowledge in the collaborative entity is organized by semantic link, the mapping relationship between the data sets is established, and the separated data set links are formed into a uniform knowledge graph, which ensures the semantic consistency of the collaborative information.

This section establishes semantic association between extracted graphs by means of semantic links. Through these semantic links, different extracted graphs can be linked together to form a uniform concept network. As shown in Fig.5, the University dataset, Course dataset, and Building dataset are three OWL ontology fragments constructed through protege. Through semantic links, they are linked to form a uniform knowledge graph.

The basic idea of constructing a uniform knowledge graph is shown in Fig.6, which can be divided into four steps.

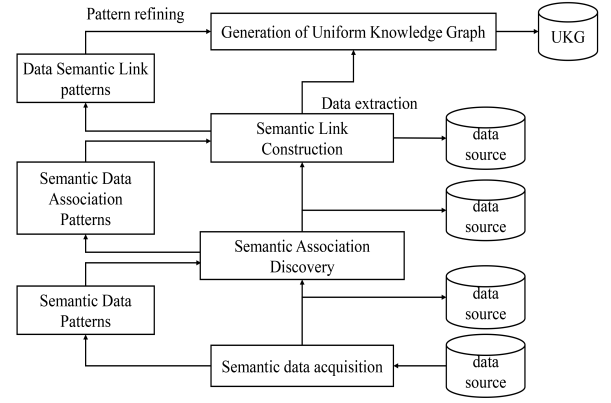


Fig. 6. Construction of uniform knowledge graph

- 1) **Semantic Data Acquisition:** Obtain semantic data from semantic internet of things;
- 2) **Semantic Data Association:** Extract the correlations between semantic data, forming many isolated extracted graphs;
- 3) **Construct Semantic Links:** Obtain the association relationship of data between extracted graphs and establish semantic links;
- 4) **Construct Uniform Knowledge Graph:** In the previous step, the schema level was not involved. In fact, the schema is the extraction of knowledge. Following the pre-given schema helps to standardize knowledge, and is more conducive to subsequent processing such as knowledge query. Constructing schemas for knowledge graphs is equivalent to construct ontologies for them. The most basic ontology includes concept, concept hierarchy, attribute, attribute value type, relation, relation domain concept set and relation domain concept set. On this basis, we can add additional rules or axioms to represent more complex constraints at the schema level.

In order to construct semantic links, we need to use ontology mapping to construct semantic links between data sets.

Definition 4. (Ontology Mapping, OM) For given two ontologies O_1 and O_2 , the mapping of the ontology O_2 to O_1 is expressed as $M_{O_1, O_2} = (MON, MOU, MOS, M_{O_1, O_2}^C, M_{O_1, O_2}^A, M_{O_1, O_2}^{RT})$. If there is a mapping relationship between ontology O_1 is $\{O_2, O_3, \dots, O_n\}$ and ontology O_1 , the mapping set of ontology O_1 is represented by M_{O_1} . $M_{O_1} = (ON, OU, OS, M_{O_1, O_2}, M_{O_1, O_3}, \dots, M_{O_1, O_n})$. Where

- 1) MON represents the name of ontology O_2 .
- 2) MOU represents the storage address of ontology O_2 .
- 3) MOS represents the storage form of ontology O_2 .
- 4) M_{O_1, O_2}^C represents the set of conceptual mappings between ontologies O_1 and O_2 . $M_{O_1, O_2}^C = \{M_{C_x, C_y}^C | c_x \in C_x, c_y \in C_y\}$.
- 5) M_{O_1, O_2}^A represents the set of attribute mappings between ontologies O_1 and O_2 . $M_{O_1, O_2}^A = \{M_{C_x, C_y}^A | c_x \in C_{O_1}, c_y \in C_{O_2}\}$.

- 6) M_{O_1, O_2}^{RT} denotes the set of relational mappings. $M_{O_1, O_2}^{RT} = \{m_1^{RT}, m_2^{RT}, \dots, m_m^{RT}\}$, and $n = |M_{O_1, O_2}^{RT}|$ denotes the number of relationship type mappings. The relationship type mapping is expressed as $m_i^{RT} = (rt_{i_1}, rt_{i_2})$. Generally, the number of relationship types in ontology is small, therefore, the credibility is not set.
- 7) ON represents the name of ontology O_1 .
- 8) OU represents the storage address of ontology O_1 .
- 9) OS represents the storage form of ontology O_1 .

Ontology semantic mapping includes attribute set mapping and concept mapping. Given two or more ontologies. For any one concept in one ontology, concept mapping refers to the attempt to find one or more corresponding concepts in another ontology, so that they have the same or similar semantics. The attribute set mapping of concepts refers to the attribute set mapping of two concepts between ontologies. Next, we introduce the attribute set mapping algorithm, the concept mapping algorithm and the ontology semantic mapping algorithm respectively.

1) *Attribute set mapping algorithm for concepts*: The input is attribute set of two concepts ($Attr_x, Attr_y$) and the output is attribute mapping set $AM(x, y)$. The algorithm uses an approximation mechanism. when the similarity of attributes is greater than the limit of the similarity of attributes lim_L , it is added to the set of attribute mapping. As shown in algorithm 1, where $Attr_c$ denotes the set of attributes of concept C . $AM(x, y)$ represents the attribute mapping between concepts x and y . a represents attribute in attribute set A .

Algorithm 1 Attribute Set Mapping Algorithm for Concepts

Input: Attribute set of concepts $Attr_x, Attr_y$

Output: Attribute mapping set $AM(x, y)$

```

1:  $AM(x, y) = \text{NULL}$ 
2: for  $ax$  in  $Attr\_x$  do
3:    $lim = 0$ 
4:    $k = 0$ 
5:   for  $ay$  in  $Attr\_y$  do
6:      $\{v_j \text{ is the adjacent vertex of } v_i\}$ 
7:      $lim(ax, ay) = Sim(ax, ay)$ 
8:     if  $lim(ax, ay) > lim$  then
9:        $lim = lim(ax, ay)$ 
10:    end if
11:  end for
12: end for
13: if  $(Attr\_y \neq \text{NULL}) \&\& (lim \geq lim\_L)$  then
14:    $AM(x, y) \leftarrow (ax, ay, lim)$ 
15: end if
16: return  $AM(x, y)$ 
```

2) *Concept mapping algorithm*: Concept is the core element of ontology, therefore, the concept mapping is the basis of ontology mapping. The concept mapping algorithm determines the semantic relationship between concepts according to the set of attribute mappings. For relationship " \cap ", the concept similarity is further determined by the combination of basic similarity, attribute set similarity and related concept similarity. As shown in algorithm 2, the set of concept mappings is represented as $CM(x, y)$. $believe(x, y)$ denotes the mapping

credibility of concepts x to y . $relation(x, y)$ denotes the semantic relationship between concepts x and y .

Algorithm 2 Concept Mapping Algorithm

Input: The attribute sets of concepts C_x and C_y are $Attr_x$ and $Attr_y$, respectively. And attribute mapping set $AM(x, y)$

Output: Concept mapping $CM(x, y)$

```

1:  $Attr\_x\_temp = Attr\_y\_temp = \text{NULL}$ 
2:  $Sum = 0$ 
3:  $Count = 0$ 
4: if  $AM(x, y) == \text{NULL}$  then
5:   Return  $\text{NULL}$ 
6: end if
7: for  $ax, ay, lim(ax, ay)$  in  $AM(x, y)$  do
8:    $Attr\_x\_temp = Attr\_x\_temp \cup ax$ 
9:    $Attr\_y\_temp = Attr\_y\_temp \cup ay$ 
10:   $Sum = Sum + lim(ax, ay)$ 
11:   $Count = Count + 1$ 
12: end for
13:  $believe(x, y) = Sum / Count$ 
14: if  $(Attr\_x\_temp == \text{NULL}) \&\& (Attr\_y\_temp == \text{NULL})$  then
15:   Return  $\text{NULL}$ 
16: end if
17: if  $(Attr\_x\_temp == Attr\_x) \&\& (Attr\_y\_temp == Attr\_y)$  then
18:    $C_x$  is equivalent to  $C_y$ 
19: end if
20: if  $(Attr\_x\_temp == Attr\_x) \&\& (Attr\_y\_temp \text{ in } Attr\_y)$  then
21:    $C_x$  is a subordinate concept of  $C_y$ .
22: end if
23: if  $(Attr\_x\_temp \text{ in } Attr\_x) \&\& (Attr\_y\_temp == Attr\_y)$  then
24:    $C_x$  is the superordinate concept of  $C_y$ 
25: else
26:    $C_x$  and  $C_y$  have an intersection
27:    $believe(x, y) = Sim(C\_x, C\_y)$ 
28: end if
29:  $CM(x, y) \leftarrow (C\_x, C\_y, relation(x, y), believe(x, y))$ 
30: return  $CM(x, y)$ 
```

3) *Semantic link construction algorithm*: First, the basic similarity of concepts is calculated. When the basic similarity of concepts is greater than the basic similarity limit a_B of concepts, the attribute mapping set of concepts is obtained by using the attribute set mapping algorithm of concepts. Then the concept mapping algorithm is used to get the concept mapping. As shown in algorithm 3.

C. Semantic Cooperative Process Based on uniform knowledge graph

For the problem of unclear information for users' needs, in order to be able to return the results that meet the user' needs. We propose a semantic collaboration method based on uniform knowledge graph. In this method, the query accuracy is improved by using the advantage of uniform knowledge graph in knowledge representation and semantic reasoning. The main framework of user requirements mapping is shown in Fig.7.

Algorithm 3 Semantic Link Construction Algorithm**Input:** Ontology O_1, O_2 **Output:** $M(O_1, O_2)$

```

1:  $M(O_1, O_2) = \text{NULL}$ 
2:  $CM(O_1, O_2) = \text{NULL}$ 
3:  $AM(O_1, O_2) = \text{NULL}$ 
4: for  $c_x$  in  $C_{O1}$  do
5:   for  $c_y$  in  $C_{O2}$  do
6:      $m(x, y)$ 
7:     if  $m(x, y) \geq a$  then
8:        $AM(x, y) = \text{Attr\_Mapping}(\text{Attr}_x, \text{Attr}_y)$ 
9:        $CM(x, y) = \text{Attr\_Mapping}(\text{Attr}_x, \text{Attr}_y, AM(x, y))$ 
10:       $AM(O_1, O_2) = \text{emph}AM(O_1, O_2) \cup AM(x, y)$ 
11:       $CM(O_1, O_2) = \text{emph}CM(O_1, O_2) \cup CM(x, y)$ 
12:    end if
13:  end for
14: end for
15: if  $CM(O_1, O_2) \neq \text{NULL} \ \&\& \ AM(O_1, O_2) \neq \text{NULL}$  then
16:    $M(O_1, O_2) \leftarrow MON + MOU + MOS + CM(O_1, O_2) + AM(O_1, O_2)$ 
17: end if
18: return  $M(O_1, O_2)$ 

```

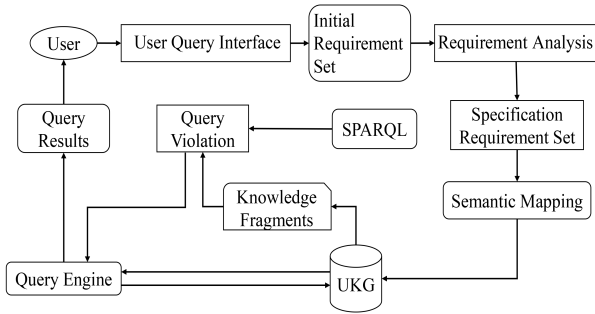


Fig. 7. User requirement mapping framework

In order to find as many results as possible to meet the user's needs. We use a maximum combination theory to initialize the demand information. The idea of this theory is as follows.

- 1) Use a keyword as a retrieval condition and query it in the knowledge base. If a valid result is returned, the keyword will be saved in the valid requirement set, otherwise it will be deleted. The effective requirement set stores the data in the form of stack.
- 2) Two keywords are randomly selected in the keyword set for validity checking, and gradually added to the selection of all keywords for validity checking.
- 3) Check the set of valid requirements until all combinations in the keyword set have been validated, and the combinations in the requirement set are checked one by one out of the stack. Check the combinations in the requirement set one by one until the number of keywords in combination A is greater than that in combination B, then combination A and the combination found before are the maximum combination.

The effectiveness of requirement combination needs to be judged by knowledge graph, which is mainly to judge whether the requirement keywords in the combination can find the

corresponding concepts, attributes or examples in knowledge graph. If it can be found, it is valid, otherwise, the combination is considered invalid.

After the user's requirement information is analyzed, it can form a standardized requirement information, which is represented by a set of keywords, such as "John's Weblog", after word division and normalization to form a collection {John, weblog}. The key words are regarded as part of the knowledge fragments in the knowledge graph of concepts, attributes or instances. Based on these fragments, the knowledge fragments are mapped to the uniform knowledge graph by means of semantic mapping, and the relevant knowledge in the knowledge graph is found. In this process, the maximum combination theory is used to take the keywords in the user's requirement set as the main concept of the query, and the other keywords as supplementary attributes, the semantic mapping, to obtain the maximum combination to meet the requirement. Then, these combinations are used as semantic query conditions, and the keyword query is transformed into a semantic query by SPARQL query statement.

IV. EXPERIMENT

Semantic reasoning plug-in Jena¹ and ontology editing software Protégé² are used in the experiment. Jena is mainly used to support the application of Semantic Web. It provides the operation of ontology and the reasoning engine based on semantics. Protégé is a Java-based ontology building software developed by Stanford University Biology Research Center. It is an open source tool software and the core tool of ontology building.

The model takes the semantic data in the SIOT as the source data. The source data and configuration file (Config. xml) are used as input of the system. Source data is first pre-processed and parsed into RDF triples by Jena. And then stored in the initial ontology knowledge base.

The initial ontology knowledge base is verified by the consistency detection algorithm, and the inconsistency concepts and rules are eliminated by the inconsistency resolution algorithm, then a consistent ontology knowledge base is formed. Consistency ontology knowledge base is used as data source, and the constructed association file is as the input of the unified knowledge graph generation module. According to the concept and attribute association of constrained extraction ontology, the module constructs a small extraction graph one by one. Then, according to the ontology mapping relationship in the association relationship, these extraction graphs are integrated to form a unified knowledge graph. With a unified knowledge graph as an intermediary, semantic data can be shared and interoperable, and eventually the results of the collaborative work back to the users.

A. dataset

As the current research on the SIOT is still not deep enough, there is no recognized standard data set. We choose some data

¹<http://jena.apache.org/>

²<https://mailman.stanford.edu/pipermail/protege-owl/2010-November/015605.html>

of ServiceXchange website, which provides a search engine for external Web services. At present, there are more than 16,000 Web services on the website. We select some Web services in the field of intelligent conference as instance-level data of SWoT ontology. At the same time, the CoBra ontology shared in the website³ is used as the model layer data of SWoT. The ontology library is constructed by Protg and saved as CoBra ontology. Finally, the source data suitable for the environment of the Semantic Internet of Things is formed.

We simulated ten user query statements, as shown in TABLE II. Among them, There is a syntax error in the fourth set of query statements. Invalid properties in the fifth and ninth queries. The rest of the query statements are sorted by complexity. The CoBra ontologies constructed are selected as the experimental data, which mainly contains the related concepts and attributes in the field of intelligent conference. It is more in line with the relevant characteristics in the field of semantic Internet of Things. The main attributes and concepts are shown in TABLE III.

B. Evaluation Index

The performance evaluation of semantic collaboration method based on unified knowledge graph proposed in this paper is mainly considered from two aspects. On the one hand, it needs to evaluate the accuracy of collaborative results [36], called query accuracy rate; on the other hand, it also needs to evaluate the integrity of collaborative results [37], called query integrity rate. query accuracy rate and query integrity rate are commonly used indicators to evaluate users' satisfaction with query results. They directly reflect the function and efficiency of the system.

Query Accuracy Rate (QAR). If the results set returned for a query is R . S_q is the complete set of services required by users, the QAR value of the system is shown in the following formula.

$$QAR = \frac{S_q}{R} \times 100\% \quad (2)$$

Query Integrity Rate (QIR). For a query, $S=\{S_1, S_2, \dots, S_m\}$ is the set of results, and S_q is the complete set of results required by users. The QCR value of the system is shown in the following formula.

$$QCR = \frac{S_1 \cup S_2 \cup \dots \cup S_m}{S_q} \times 100\% \quad (3)$$

C. experimental results

The trimmed CoBra ontology [38] was used as the data source, and the average value was obtained by 10 experiments. The following experimental results are obtained. The bar chart shown in Fig. 8 represents the number of results returned by three different query methods.

From Fig. 8, we can see that the number of query results of the proposed method is significantly less than that of ServiceXchange, which reduces the redundancy of the results. At the same time, it is more than the traditional ontology matching

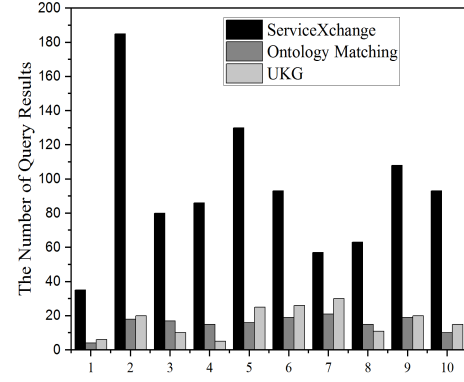


Fig. 8. The number of query results comparison

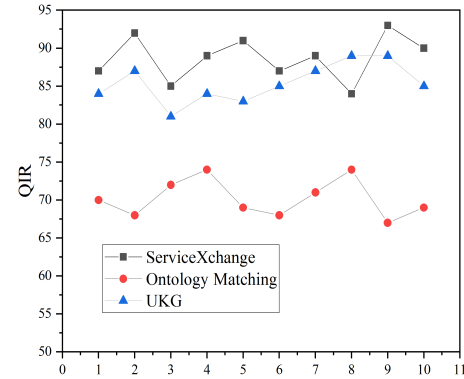


Fig. 9. Query integrity rate comparison

method. Our method (UKG) ensures the completeness of the query results and avoids missing the results that meet the requirements. The values QAR and QIR can better illustrate the advantages of the proposed method.

Fig. 9 shows the query integrity rate comparison of different methods, we can see that the query result integrity rate by the proposed semantic collaboration method is slightly lower than that of keyword-based queries, but the difference is not significant. However, the query integrity rate of UKG is significantly higher than that of traditional ontology matching methods. As can be seen from the QIR value, the method proposed in this paper reduces the redundancy of query results and takes into account the satisfaction of users needs.

Fig. 10 shows a broken-line graph of query accuracy for three query modes. When there are more data sources to query, the query accuracy rate of keyword-based query decreases more obviously. Relatively speaking, the QAR value of ontology matching-based queries decreases slowly. However, when the data complexity of queries increases gradually, the QAR of UKG does not change much and the decline is not obvious. It shows that in the complex environment of multi-data sources, the method proposed in this paper is obviously better than the other two methods. Which can better mine the deep meaning of user needs, reduce the redundant information and have higher accuracy. SIOT is a multi-data source environment with strong dynamic and obvious changes. Therefore, the proposed method is more suitable for the SIOT environment.

In summary, the semantic collaboration method based on

³<http://daml.umbc.edu/ontologies/cobra/0.2/cobra-ont>

Order Number	Query Statement
1	NAME like '%translation%' property like '%englishchinese%'
2	NAME like '%pay%' property like '%online%'
3	NAME like '%weather%' property like '%forecast%' property like '%america%'
4	NAME like '%forecast%' property like '%weather%' property like '%america%'
5	NAME like '%pay%' property like '%online%' property like '%RMB%'
6	NAME like '%person%' property like '%speakerrole%' property like '%speakerofpersentation%'
7	NAME like '%notes%' property like '%journalentry%' property like '%date%'
8	NAME like '%notes%' property like '%journalentry%' property like '%name%'
9	NAME like '%pay%' property like '%online%' property like '%RMB%' property like '%secure%'
10	NAME like '%person%' property like '%speakerrole%' property like '%RMB%' property like '%speakerofpersentation%' property like '%personinroom%'

TABLE II
QUERY STATEMENTS

The Concept of CoBra Ontology			
PlaceAtomicPlace	Agent	ThingInBuilding	PresentationSchedule
CompoundPlace	Person	SoftwareAgentInBuilding	EventHappeningNow
CampusBuilding	SoftwareAgent	PersonInBuilding	PresentationHappeningNow
AtomicPlace	Role	ThingNotInBuilding	SpeakerofPresentation
InBuilding	SpeakerRole	SoftwareAgent	HappeningNow
RoomParkingLot	AudienceRole	NotInBuilding	PersonFilesSpeakerRole
StairwayRestroom	IntentionalActive	PersonNotInBuilding	PersonFiles AudienceRole
OtherPlace-InBuilding	ActionFound	PersonInBuilding	InPresentation
Attributes of CoBra Ontology			
latitude	hasContactInformation	locatedIn	participatesIn
longitude	hasFullName	locatedIn AtomicPlace	startTime
hasPrettyName	hasEmail	locatedInRoom	endTime
isSpatiallySubsumedBy	hasHomePage	locatedInRestroom	location
SpatiallySubsumes	hasAgentAddress	locatedInParkingLot	hasEventHappeningNow
accessRestricted	fillRole	locatedInBuilding	invitedSpeaker
ToGender	intendsToPerform	locatedInCampus	expectedAudience

TABLE III
ONTOLOGY FOR EXPERIMENT

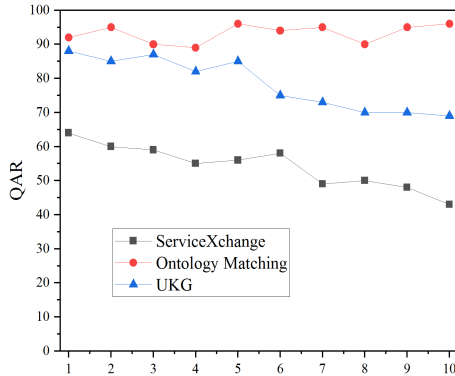


Fig. 10. Query accuracy rate comparison

the uniform knowledge graph proposed in this paper takes into account the completeness and conciseness of the query results, while query accuracy rate and query integrity rate are high, which is superior to the traditional query method. And it is more suitable for the SIOT environment. Therefore, the semantic collaboration method proposed in this paper can better satisfy users query requests.

V. APPLICATION EXAMPLE

We select a specific scenario of collaborative manufacturing as an application example to illustrate the semantic collaboration process described in this paper.

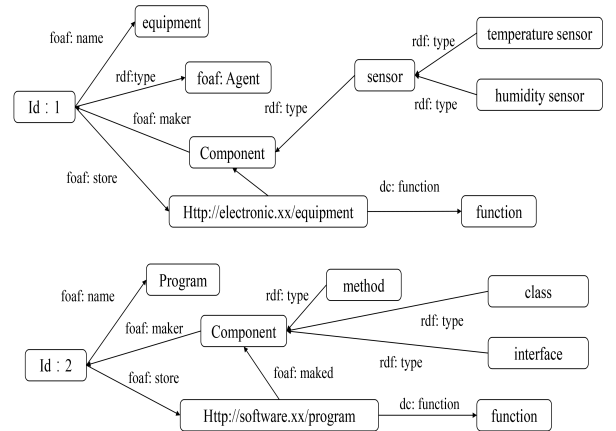


Fig. 11. Semantic Data Express

The development of an embedded device requires mechanical engineers, electronic engineers and software engineers to work together. Software engineers need to find functional modules to modify, and then input component keyword "Component". At the same time, the design of the device also needs to be modified. In order to ensure the correctness of the search results, we also input "Program" and "Equipment". The user's actual needs are functional modules and related electronic components.

The part of the semantic data set involved in this query is shown in Fig. 11.

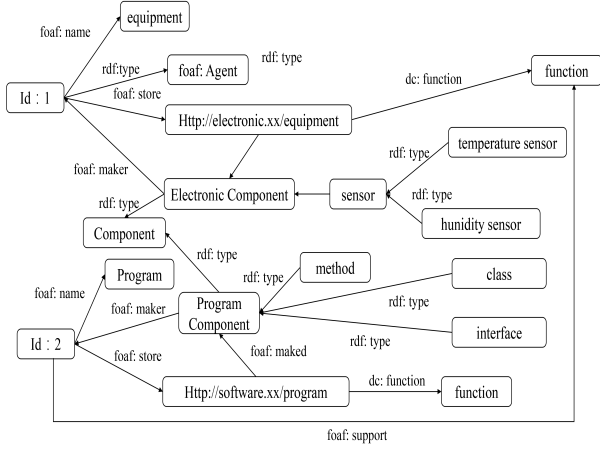


Fig. 12. Construction of a Uniform Knowledge Graph

However, in the field of electronic design, $Component \subseteq Equipment$, $Component$ is a sub-concept of $Equipment$. In the field of software development, $Component \subseteq Program$, $Component$ is a sub-concept of $Program$. " $Equipment$ " and " $Program$ " are two unrelated concepts, $Equipment \subseteq \neg Program$. Because the user's needs are not clear, the return results are often all the information related to these three keywords, which cannot meet the user's needs very well.

In the design method of this paper, the two ontologies are firstly loaded for consistency detection. The following axioms exist.

$$\begin{aligned} ax_1 : Component &\subseteq Equipment \\ ax_2 : Component &\subseteq Program \\ ax_3 : Equipment &\subseteq \neg Program \end{aligned} \quad (4)$$

These axioms are semantic inconsistency and consistency maintenance is required. By using the method described in Section III, the minimum adjustment set is $\{ax_1, ax_2\}$. We choose manual processing to resolve inconsistencies. Modify $Component$ in electronic domain ontology to $ElectronicComponent$ and add relational type ($ofComponent$). $Component$ in the software domain is modified to $ProgramComponent$, and relational type ($ofComponent$) is added as well. After that, we use the algorithm in Section III to establish semantic links, and finally form a uniform knowledge graph (UKG) as shown in Fig. 12.

After standardization of user requirements, a set of keywords $\{Component, Program, Equipment\}$ is formed, which is maximized and combined to form the final set of query keywords $\{Component, Program, Equipment\}$. Through Semantic Mapping, a knowledge fragment is returned. Its graphical representation is shown in Fig. 13. The original keyword set has been transformed into semantic query, and its corresponding SPARQL query statement is:

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?url
FROM <Embedded device>
WHERE {
  ?contributor foaf:name "Program Component"
```

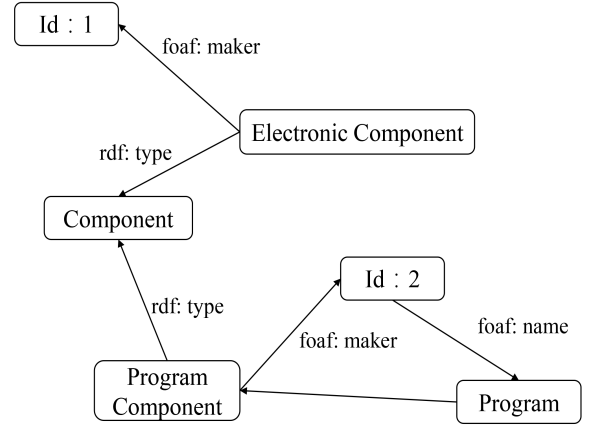


Fig. 13. Knowledge Fragment Returned by Semantic Mapping

```
?contributor foaf:make "Program"
?contributor foaf:nsupport "Equipment"
?contributor foaf:store ?url
}
```

Finally, through SPARQL query, the result is returned to the user as the URL of $ProgramComponent$ and $Equipment$.

In summary, the semantic collaboration method based on the uniform knowledge graph proposed in this paper has been completed.

VI. CONCLUSION

Semantic Internet of Things is the expansion and optimization of the Internet of Things. Its main purpose is to integrate resources, share information and interoperate in the Internet of Things through semantic technology. It takes advantage of semantic technology in knowledge representation and knowledge sharing. Semantic technology is introduced into the field of Internet of Things to solve the problem of resource interoperability in the Internet of Things. Its ultimate goal is to achieve semantic collaboration.

In this paper, after analyzing the data characteristics of the semantic Internet of Things, the unification of knowledge graph as the organization form of semantic data, combined with the advantage of knowledge graph on knowledge representation and processing, designed the semantic collaboration method based on the uniform knowledge graph, finally realizes the semantic unity based knowledge graph in the Internet of things the basic function of the semantics of the synergy.

REFERENCES

- [1] H. Yang, S. Kumara, S. T. S. Bukkapatnam, and F. Tsung, "The internet of things for smart manufacturing: A review," *IJSE Transactions*, pp. 1–35, 2019.
- [2] X. Zhao, D. E. Lucani, X. Shen, and H. Wang, "Reliable iot storage: Minimizing bandwidth use in storage without newcomer nodes," *IEEE Communications Letters*, vol. PP, no. 99, pp. 1–1, 2018.
- [3] L. Buhmann, R. Usbeck, and A. N. Ngomo, "Assess — automatic self-assessment using linked data," pp. 76–89, 2015.
- [4] "Semantic technology," *Design & Test IEEE*, vol. 31, no. 5, pp. 66–66.
- [5] V. Raskin and S. Nirenburg, "An applied ontological semantic microtheory of adjective meaning for natural language processing," *Machine Translation*, vol. 13, no. 2/3, pp. 135–227, 1998.

- [6] B. Chen, X. Dong, D. Jiao, H. Wang, Q. Zhu, Y. Ding, and D. J. Wild, "Chem2bio2rdf: a semantic framework for linking and data mining chemogenomic and systems chemical biology data," *Bmc Bioinformatics*, vol. 11, no. 1, p. 255, 2010.
- [7] B. Zhou, Z. Hang, X. Puig, T. Xiao, S. Fidler, A. Barriuso, and A. Torralba, "Semantic understanding of scenes through the ade20k dataset," vol. 127, no. 3, pp. 302–321, 2016.
- [8] N. Shadbolt, T. Berners-Lee, and W. Hall, "The semantic web revisited," *IEEE Intelligent Systems*, vol. 21, no. 3, pp. 96–101, 2006.
- [9] T. Bernerslee and J. A. Hendler, "Publishing on the semantic web," *Nature*, vol. 410, no. 6832, pp. 1023–1024, 2001.
- [10] Y. Liu, "Multi-source heterogeneous data fusion based on perceptual semantics in narrow-band internet of things," *Personal and Ubiquitous Computing*, no. 7, pp. 1–8, 2019.
- [11] Z. Yorio, R. Oram, S. El-Tawab, A. Salman, M. H. Heydari, and B. B. Park, "Data analysis and information security of an internet of things (iot) intelligent transit system," pp. 24–29, 2018.
- [12] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," *Computer Networks*, vol. 54, no. 15, pp. 2787–2805, 2010.
- [13] I. P. Zarko, "Interoperability and open-source solutions for the internet of things: International workshop, fp7 openiot project, held in conjunction with softcom 2014, split, croatia, september 18, 2014, invited p," *Lecture Notes in Computer Science*, vol. 9001, 2015.
- [14] J. R. D. M. Nogueira, N. C. L. de Oliveira Ahiadzro, L. T. Cavalini, and T. W. Cook, "Knowledge management of controlled vocabularies for semantic interoperability of healthcare applications," in *International Conference on Healthcare Informatics*, 2015.
- [15] F. Loukil, C. Ghediraguean, K. Boukadi, and A. N. Benharkat, "Liopy: A legal compliant ontology to preserve privacy for the internet of things," in *IEEE Computer Software & Applications Conference*, 2018.
- [16] S. S. Reka and T. Dragicevic, "Future effectual role of energy delivery: A comprehensive review of internet of things and smart grid," *Renewable & Sustainable Energy Reviews*, vol. 91, pp. 90–108, 2018.
- [17] E. Arias-De-Reyna, P. Closas, D. Dardari, and P. M. Djuric, "Crowd-based learning of spatial fields for the internet of things: From harvesting of data to inference," *IEEE Signal Processing Magazine*, vol. 35, no. 5, pp. 130–139, 2018.
- [18] A. Pomp, A. Paulus, S. Jeschke, and T. Meisen, "Eskape: Platform for enabling semantics in the continuously evolving internet of things," in *IEEE International Conference on Semantic Computing*, 2017.
- [19] G. Bakal, P. Talari, E. V. Kakani, and R. Kavuluru, "Exploiting semantic patterns over biomedical knowledge graphs for predicting treatment and causative relations," *Journal of Biomedical Informatics*, 2018.
- [20] A. Palavalli, D. Karri, and S. Pasupuleti, "Semantic internet of things," in *IEEE Tenth International Conference on Semantic Computing*, 2016.
- [21] A. Sheth, "Internet of things to smart iot through semantic, cognitive, and perceptual computing," *IEEE Intelligent Systems*, vol. 31, no. 2, pp. 108–112, 2016.
- [22] S. Mayer, J. Hodges, Y. Dan, M. Kritzler, and F. Michahelles, "An open semantic framework for the industrial internet of things," *IEEE Intelligent Systems*, vol. 32, no. 1, pp. 96–101, 2017.
- [23] C. W. Tsai, C. F. Lai, and A. V. Vasilakos, "Future internet of things: open issues and challenges," *Wireless Networks*, vol. 20, no. 8, pp. 2201–2217, 2014.
- [24] W. Wang, S. De, G. Cassar, and P. K. Moessner, "Knowledge representation in the internet of things: Semantic modelling and its applications," *Automatika Journal for Control Measurement Electronics Compu*, vol. 54, no. 4, pp. 388–400, 2013.
- [25] V. Y. Terziyan, "Semantic web services for smart devices in a "global understanding environment"," *Lecture Notes in Computer Science*, pp. 279–291, 2003.
- [26] V. Terziyan, "Global understanding environment: towards self-managed web of everything," in *International Conference on Grid & Pervasive Computing*, 2011.
- [27] K. Kotis, A. Katasonov, and J. Leino, *Aligning Smart and Control Entities in the IoT*, 2012.
- [28] K. Kotis and A. Katasonov, "Semantic interoperability on the web of things: The smart gateway framework," in *Sixth International Conference on Complex*, 2012.
- [29] P. Barnaghi, W. Wang, C. Henson, and K. Taylor, "Semantics for the internet of things: early progress and back to the future," *International Journal on Semantic Web & Information Systems*, vol. 8, no. 1, pp. 1–21, 2012.
- [30] T. S. Dillon, A. Talevski, V. Potdar, and E. Chang, "Web of things as a framework for ubiquitous intelligence and computing," in *International Conference on Ubiquitous Intelligence & Computing*, 2009.
- [31] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Context aware computing for the internet of things: A survey," *IEEE Communications Surveys and Tutorials*, vol. 16, no. 1, pp. 414–454, 2014.
- [32] P. B. Jeon, J. W. Kim, S. Lee, C. Lee, and D. K. Baik, "Semantic negotiation-based service framework in an m2m environment," in *IEEE/WIC/ACM International Conference on Web Intelligence & Intelligent Agent Technology*, 2011.
- [33] M. Vega-Barbas, D. Casado-Mansilla, M. A. Valero, D. López-De-Ipina, J. Bravo, and F. Flórez, "Smart spaces and smart objects interoperability architecture (s3oia)," in *Sixth International Conference on Innovative Mobile & Internet Services in Ubiquitous Computing*, 2012.
- [34] H. Lin, Y. Liu, W. Wang, Y. Yue, and Z. Lin, "Learning entity and relation embeddings for knowledge resolution," *international conference on conceptual structures*, vol. 108, pp. 345–354, 2017.
- [35] M. Doerr, "The cidoc conceptual reference module: an ontological approach to semantic interoperability of metadata," *Ai Magazine*, vol. 24, no. 3, pp. 75–92, 2003.
- [36] N. Alexander, "Precision of rate estimation under uniform interval censoring," *Statistics in Medicine*, vol. 27, no. 17, pp. 3442–3445, 2010.
- [37] A. Ohlsson, C. Guthenberg, and U. V. Döbeln, "Galactosemia screening with low false-positive recall rate: The swedish experience," *Jimd Rep*, vol. 2, pp. 113–117, 2012.
- [38] S. Aitken and Y. Chen, *COBRA and COBRA-CT: Ontology Engineering Tools*, 2008.