

Web service discovery: Rationale, challenges, and solution directions

Ahmet Vedat Tokmak^a, Akhan Akbulut^a, Cagatay Catal^{b,*}

^a Department of Computer Engineering, Istanbul Kültür University, 34536 Istanbul, Turkey

^b Department of Computer Science and Engineering, Qatar University, 2713 Doha, Qatar

ARTICLE INFO

Keywords:

Microservices
Microservice architecture
Service discovery
SLR

ABSTRACT

Service Oriented Architecture (SOA) is a methodology that promotes cooperation between services with diverse, but connected functions. Web Service technology paved the way for microservice architecture as it is a feature of modern web applications that resulted from the rise of SOA. With the proliferation of self-contained services, the ease of finding has emerged as a critical concern. Due to the increasing number of services that perform identical tasks, it has become difficult for users to select the most feasible service. Providing the most relevant service for the customer quickly is a crucial infrastructure task, and undiscovered services increase ecosystem expenses. Syntactic, semantic-conscious, and ontology-based studies have been presented as ways to improve the effectiveness and quality of service discovery techniques. While there are many approaches that have been proposed and validated for service discovery in literature, these studies are fragmented and there is a lack of overview of the techniques of web service discovery. As such, we conduct a Systematic Literature Review (SLR) study to review the existing body of knowledge surrounding service discovery and discuss the state-of-the-art. We present an overview of the techniques and empirical evidence by identifying, analyzing, and classifying the papers. Among the 764 papers we retrieved, 54 papers were included. We provide a comprehensive analysis of methodologies and tools for discovering web services.

1. Introduction

SOA is a method for designing distributed systems that offer application functionality as language and platform-independent services [1]. Each service within SOA executes one or more functions separately. Web Services (WS) are web-based programs that may be published, distributed, and utilized independently. Because of the convenience and interoperability of data provided by Extended Markup Language (XML), Simple Object Access Protocol (SOAP), Universal Description, Discovery, and Integration (UDDI), and Web Services Description Language (WSDL) technologies, web services are widely used. While SOAP is used to communicate data, XML is utilized to define data in a systematic manner using carefully defined labels [2]. Because of the benefits of WSs, their number is expanding rapidly, making WS discovery more challenging.

SOAP, one of the fundamental communication protocols of Service-Oriented Architecture (SOA), facilitates messaging and data sharing between services. Conversely, web services often employ REST (Representational State Transfer)-based APIs for communication purposes. REST is a widely adopted architectural style for data sharing and communication in web-based applications. RESTful services utilize the HTTP protocol to provide APIs that represent resources and support CRUD

(Create, Read, Update, Delete) operations. Both SOAP and REST are protocols utilized for communication between services within a service-oriented architectural design. Microservices, on the other hand, adhere to the core principles of SOA to construct a service-based system. The distinction lies in the fact that while SOA can utilize SOAP-based communication protocols, microservices are typically implemented using REST-based APIs [3].

Typically, a web service discovery procedure consists of three steps. First, providers advertise web services in public pools by publicly disclosing programming interfaces through description files such as WSDL files. Second, a user submits a web service request, and finally, a web service discovery model is used to locate a set of web service candidates, which are typically sorted and displayed to the user in a particular order. The user, then, invokes one of the web services. The discovery phase is the most important step in the utilization of web services, as undiscovered services cannot be utilized [4]. The list of retrieved web services may include many services with equivalent functionality. Service Quality (QoS) is decisive in this instance.

QoS data is typically used to determine the quality grade of prospective web services. These QoS-related details may include performance (e.g., response time, latency), availability, accessibility, and security,

* Corresponding author.

E-mail addresses: 200005655@stu.iku.edu.tr (A.V. Tokmak), a.akbulut@iku.edu.tr (A. Akbulut), ccatal@qu.edu.qa (C. Catal).

<https://doi.org/10.1016/j.csi.2023.103794>

Received 28 March 2023; Received in revised form 25 July 2023; Accepted 17 September 2023

Available online 25 September 2023

0920-5489/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

among others [5]. This information significantly affects the use of a web service and their expectations of this service. Although it assures validity to a certain extent, QoS information published by service providers is not always accurate. Data mining and machine learning approaches are employed to enhance the WS discovery procedure. Clustering techniques are used to reduce the search space for online services based on a variety of similarity criteria and to group web services into distinct clusters. To construct a recommendation model, association rule algorithms are utilized to identify correlations between relationships between services and users [6].

In order to better understand the discovery process of web services (WS) and develop more efficient methods, it is necessary to examine previous studies conducted in this field. Previous research on the discovery process of WS can provide guidance for new methods, while the approaches used to overcome challenges in these studies can also assist in addressing the difficulties encountered by new methods. A comprehensive review of previous studies on the discovery process can provide valuable insights into various aspects, such as the methods, algorithms, technologies, and challenges encountered.

The primary goal of this study is to conduct a Systematic Literature Review (SLR) in order to facilitate a thorough understanding of the methods employed in the web service discovery process. In this study, we examined and analyzed 54 scientific publications, focusing primarily on publications that included experimental results and were conducted within the last 10 years. We focused our SLR exclusively on the microservice ecosystem's web service discovery. Over the past decade, we compiled this collection of papers from electronic databases. We mined for keywords to ensure that all applicable literature was considered. Exclusion criteria guide the study's selection procedure. Quality evaluation questions [7] were used to evaluate the quality of papers. Due to the selection of search keywords and electronic databases, it is possible that we may have missed some relevant literature on online service discovery. The authors worked together on this comprehensive literature evaluation, increasing its credibility. During the SLR's data extraction stage, the authors developed a data extraction file. The authors examined and double-checked all study questions and responses to ensure that they accurately synthesize and infer from the available data.

The following sections are organized as follows: Section 2 explains the context and related work. Section 3 shows the research methodology. Section 4 discusses the results. Section 5 concludes the paper and shows future work.

2. Background and related work

In Section 2.1, we present the steps involved in discovering and using web services. The studies that are of interest are discussed in Section 2.2.

2.1. Web service discovery

Web services are structures that provide support to other systems over the HTTP protocol. The use of web services has expanded significantly due to their simplicity and independence from certain platforms. As the quantity of accessible services increases, it becomes increasingly difficult for consumers to locate relevant services. Insignificant services that cannot be located by the user are not utilized. Therefore, the discovery process for online services is significant. This discovery is difficult because of the increasing number of web services and services with similar purposes. In general, researchers employ three fundamental methodologies during the service discovery procedure. These include the syntactic technique based on keyword matching, the semantic-conscious approach addressing the semantic description of services, and ontology-based learning.

In the web service discovery procedure, ontologies such as Owl-S (Web Ontology Language for Services), WSMO (Web Service Modeling Ontology), and SAWSDL (Semantic Annotations for WSDL) with

UDDI (Universal Identification, Discovery, and Integration) and WSDL (Web Services Description Language) protocols are used [8]. Various strategies are performed during the exploring procedure. Traditional keyword-mapping-based WS discovery, for instance, is based on Service Quality (QoS) data or context data of the WS description structure. Semantic-conscious WS discovery comprises of hidden semantic-conscious service-finding techniques based on fundamental descriptive data. Ontology-matching-based Web Services (WS) exploration is mostly based on input-output-prediction-effect-centric state mapping. Using deep learning and machine learning, researchers endeavour to improve service discovery quality.

2.2. Related work

Research on web service discovery is extensive. Taking into account the published articles, it is apparent that some challenges arise in service discovery methodologies. Using their similarities, studies combine the techniques used to determine the most accurate service in order to shorten the time required for service discovery. The grouping and categorization of services and semantic service discovery methodologies, which are prominent in the service discovery strategy, are also prominent in the review papers, based on the investigations.

Obeidallah and Raahemi (2017) present a survey of distinct web service discovery techniques, focusing on the clustering and classification techniques used to speed the discovery time of services. This survey examines context-warehousing, user-side, clustering, and recommendation-based online service discovery techniques [9]. Alam et al. (2016) provide a comprehensive assessment of the most recent technological advances in clustering and classification processes for web services. This article expands on the current state of clustering and classification approaches [10]. In the survey study of Sagayaraj and Santhoshkumar (2017), the clustering model is structured according to the clustering of web Service discovery methods. This paper examines the drawbacks of web service discovery technologies and provides a summary of the issues posed by various methodologies [11].

One of the most influential works in this area is by Elsayed and Salah (2015), in which the authors conduct a systematic survey of the literature on the topic of semantic web services discovery to address the question of how to retrieve semantic web services that are a match for a user's query. This review, in which 19 articles from the literature are analyzed, focuses on four points. The first point focuses on the techniques used to retrieve semantic web services that match a user query. The second point focuses on semantic service sections that consider matching. The third point focuses on the pre-processing stage before the pairing stage. The final point focuses on measurement to evaluate techniques [1]. Malaimalavathani and Gowri (2013) evaluate and discuss service discoverability in terms of semantic web service discovery to find the most appropriate service for a given application to use [12]. Phalnikar and Khutade (2012) present a detailed investigation of web service discovery systems utilizing service quality(QoS) factors. It also emphasizes the significance of ontology in representing non-functional requirements in web service discovery [5]. Agarwal et al. (2022) presented a Systematic Literature Review paper on web service clustering approaches aimed at enhancing service discovery, selection, and recommendation [13].

2.3. Contributions

Considering the above-mentioned background and related work, we identified three contributions in this study.

Comprehensive overview of web service discovery techniques: This paper presents a systematic review of the existing body of knowledge covering service discovery in the context of web services. It analyzes a significant number of papers to provide a comprehensive overview of the methodologies and tools employed in web service discovery. By synthesizing this diverse set of techniques, the paper presents a

Table 1
Identified research questions.

RQ	Research questions
RQ1	What method is employed for service discovery?
RQ2	What complementary functions does service discovery serve?
RQ3	What meta-heuristic techniques and algorithms are employed?
RQ4	To what extent is the implementation strategy cloud-based?
RQ5	To what extent is the implementation strategy data or service-centric?
RQ6	What are the challenges, and potential solutions in this domain?

Table 2
Exclusion criteria.

ID	Exclusion criteria
1	The paper is not fully accessible (i.e., papers that include only abstract are excluded).
2	The language of the article is not English.
3	The article does not contain experimental results.

comprehensive understanding of the current state-of-the-art in web service discovery.

Identification of gaps in the existing literature: Through the systematic literature review, this paper identifies a lack of overview in the studies related to web service discovery. By highlighting this gap, the paper brings attention to the need for a more consolidated approach to advancing web service discovery techniques. This contribution provides a foundation for future research efforts.

Empirical evidence of service discovery methodologies: The paper not only reviews the existing literature but also includes empirical evidence of the methodologies employed in web service discovery. This study discusses the effectiveness of different service discovery techniques. This contribution provides valuable insights, enabling researchers and practitioners to make informed decisions when implementing service discovery methodologies.

3. Research methodology

Procedures for gathering data, posing research questions, and deciding which articles to include are presented in this section. By narrowing the pool of potential publications down using the criteria established in this study, 54 papers were selected. Both the publications and the research questions (RQ) were carefully examined. We used [Table 1](#) to list our research questions. Resolving these uncertainties can pave the way for more primary research and can help to create new theoretical frameworks to address some of the issues raised in this research.

This SLR used an automated search in IEEE Xplore, Science Direct, Springer Link, and Scopus databases to find papers, conference proceedings, and books published on the topic of web service discovery in the last 10 years. The search was conducted with the keywords “web service discovery” and “microservice discovery” specifically to list studies that focus on the discovery stages of web services and microservices. The accessibility of the full texts of these publications and the ability to fully understand the language used in the studies have been the primary considerations in their selection. The inclusion of experimental results in the publications is necessary for a better understanding of the methods, algorithms, and technologies employed in the conducted studies. Experimental studies also provide guidance to researchers in future studies. Applying the exclusion criteria ([Table 2](#)) to the initial pool of 764 publications helped narrow down the relevant studies. Our exclusion criteria were developed based on recommendations provided in existing guidelines and SLR articles [14–18]. The inclusion criterion for SLR is selecting papers that specifically address web service discovery topics and provide experimental results.

Table 3
The distribution of papers per database.

Database	Number of papers after query search	Selected primary studies
1.Science Direct	99	5
2.IEEE Xplore	394	34
3.Scopus	219	13
4. Springer Link	52	2
Total	764	54

Table 4
Quality evaluation questions. Yes, scores 2; Partial scores 1; No scores 0.

ID	Questions
Q1	Is the purpose of the research specified?
Q2	Are the scope and context of the study clearly defined?
Q3	Are the variables used in the research valid and reliable?
Q4	Is the research process sufficiently documented?
Q5	Are the main findings clearly stated in terms of validity and reliability?

Important studies cited in the literature are also presented here. [Table 3](#) displays the outcomes of this analysis, showing the distribution of papers per database.

Following the aforementioned criteria, all papers were assessed for relevancy. In order to identify the articles that would contribute the most to the objective of the study, it is necessary to rank the publications based on their degree of relevance to the objective and prioritize the ones that provide the highest contribution. To rank the publications, a quality score was assigned to each publication. The scoring was done by answering specific questions, where a “yes” response was assigned a score of 2, a “partially” response a score of 1, and a “no” response a score of 0. The quality score of a publication was determined as the total score obtained from all the questions. The questions were selected to assess the relevance, validity, reliability, and documentation of the study. It is important for the purpose and objectives of the study to be clear in order to understand the aim of the publication. The validity and accuracy of the research methods and data analysis are related to the validity of the study. Factors such as the methodological approach, data collection and analysis processes, and measurement tools are associated with the reliability of the study. Documentation refers to factors such as correctly citing references, explaining data tables or graphs, and providing detailed descriptions of the methods used. We used the questions in [Table 4](#) to assess an article’s quality once we applied our exclusion criteria. We eliminated the low-scoring papers and narrowed it down to 54 papers. The IEEE Xplore database utilized provided most of the selected papers.

Publications are then synthesized after the application of the quality rating questions. [Fig. 1](#) shows the distribution of publication years of the papers. This table shows that the importance of the discovery process, the most crucial process for web services, is reflected in the research and publications as shown in [Fig. 1](#).

The variety of published works is depicted in [Fig. 2](#). About 50% are articles published in journals, 44% are presented at conferences, and 6% are published in books. In addition to these findings, [Fig. 3](#) illustrates the distribution of studies according to various search databases. Research on the discovery process of web services is growing its popularity as its significance becomes more apparent.

By reading each article in detail, we were able to collect the information we needed to answer our research questions. Each RQ was placed in a separate column and each publication in its own row throughout the extraction process, and we reviewed both separately and as a whole afterwards. The basic objective of data synthesis is to answer research questions by bringing together data from multiple original sources. The data analysis followed a systematic approach,

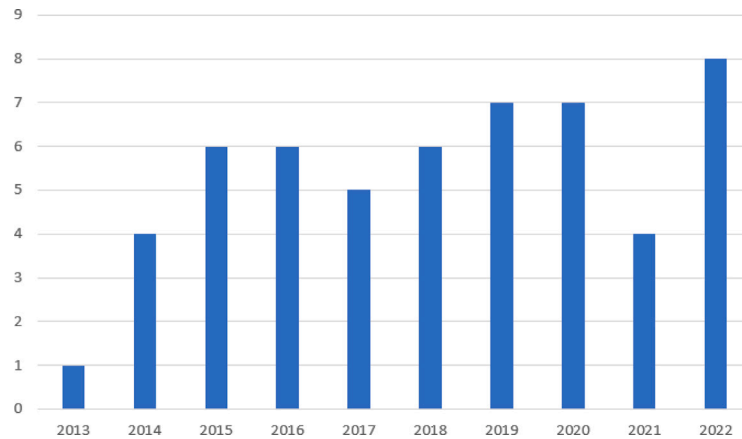


Fig. 1. Distribution of papers per year.

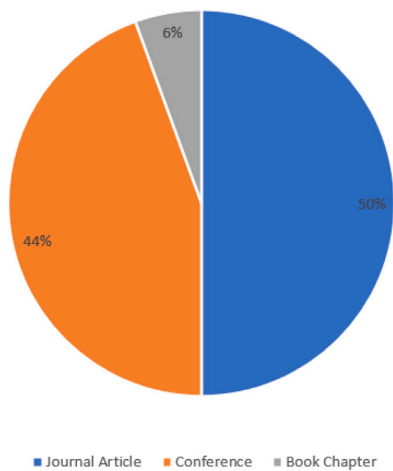


Fig. 2. Types of Publications.

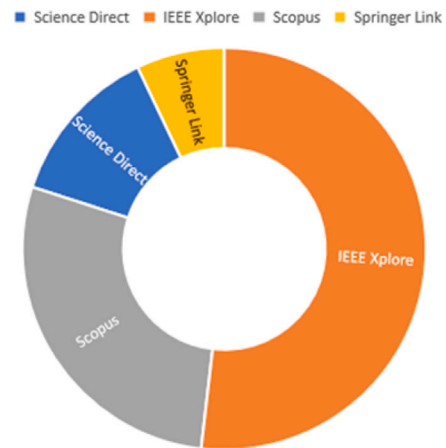


Fig. 3. The distribution of studies per search database.

which involves categorization and data synthesis. Initially, the data was organized based on key categories to observe the trends over time. Later, a qualitative synthesis approach was applied to interpret the main findings. This data analysis approach provided a solid foundation for drawing conclusions.

4. Results

After preparing a form that included the research questions, the selected publications were individually reviewed according to this form to systematically examine them. Once the review process was completed, the data were synthesized by considering similarities and differences among the publications. Finally, the data were compared, analyzed, and interpreted to answer the research questions. In this section, we report our findings per the research question.

4.1. RQ-1: What method is employed for service discovery?

This RQ describes methods for web service discovery. In the literature, syntactic [S1], semantic-conscious [S4,S5,S10,S15,S23], and ontology-based [S16,S17,S18,S19] exploration methodologies predominate. Syntactic techniques are the oldest keyword-based strategy, which is still in use. It is vital to compare keywords and web service descriptions in this strategy. Choosing the correct keywords in the syntactic approach is, therefore, crucial for the effectiveness of the discovery procedure. If the keywords employed in the discovery process are insufficient or inaccurate, it is likely that the user may encounter a large

number of unwanted services or no services at all. In this method, the use of informal language in interrogative phrases is discouraged. It will be extremely difficult for the discovery procedure to be successful given how difficult it is to connect words from ordinary English with their WS meanings. The semantic method, which arose after the syntactic approach and focuses on the semantic definition of services, enables the user to more quickly and reliably locate the desired services. The semantic aware method identifies their WS using semantic web technologies such as OWL-S, WS-Policy, and WSDLs. In the semantically conscious method, semantic inferences are employed instead of direct matching to determine whether the question and explanation are similar. In this manner, the user can also use natural language instead of attempting to employ keywords in the search question. Moreover, in this method, services with comparable characteristics are grouped according to their WS description texts. In the ontology-based approach, which is evolved from semantic-based approaches, ontology techniques are utilized to enable the user to locate the most relevant service in the context of the service and user. In this method, service profiles containing information on each service’s function, operation, and accessibility are built. After each run, not only the cookies but also the state information are defined. Consequently, each operation of a WS more closely resembles the concept of a state change. In addition to focusing on inputs and outputs, the ontology-based discovery method also emphasizes predictive and effect-oriented matching. When examining the selected articles, the semantic conscious technique was employed 40 times, the ontology-based approach 13 times, and the syntactic approach once. Fig. 4 demonstrates the distribution of approaches to

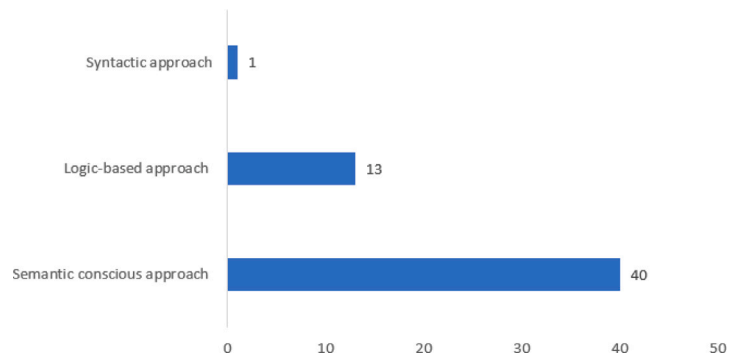


Fig. 4. Distribution of approach types of publications.

service discovery utilized in the literature. These data indicate that the most popular method for discovering online services is the semantic-conscious method. While syntactic techniques are an important aspect of web service discovery, their representation in Fig. 4 is very limited, with only one publication included. This is a result of the specific criteria we applied during the selection process. It is possible that the publications meeting our criteria for inclusion did not extensively focus on syntactic techniques, or they may have presented their results in a different manner.

4.2. RQ-2: What complementary functions does service discovery serve?

The discovery of web services relies mostly on three concepts. These are discovery [S2,S3,S4,S5,S6,S7,S8,S9,S10,S11,S12,S13,S14,S15,S16,S17,S18,S19,S20,S21,S22,S23,S24,S25,S26,S27,S28,S29,S30,S31,S32,S33,S34,S35,S36,S37,S38,S39,S40,S41,S42,S43,S44,S45,S46,S47,S48,S49,S50,S51,S52,S53,S54], health check, and load balancing [S1]. All 54 selected papers centred on the concept of discovery, as well as one additional load balancing and one health check feature, as decided by an analysis of the collected literature. It should be ensured that the data collected about the services are used to solve other ecosystem problems, despite the fact that discovery will be regarded as a key issue today. For instance, reverse proxy services and the circuit breaker mechanism employed by microservice-based software are supported.

Considering these functions, this study aims to increase its overall effectiveness and efficiency in the web services ecosystem. This approach enables the collected data to be used to solve a wider range of challenges, encouraging the development of robust and scalable systems.

4.3. RQ-3: What meta-heuristic techniques and algorithms are employed?

Since the number of available web services continues to grow exponentially, it is becoming increasingly challenging for consumers to quickly and accurately locate the best service that suits their needs. Instead of using rule-based approaches, researchers have been using a variety of strategies and algorithms to make strides in the field of service discovery. This RQ presents the most common algorithms and approaches used for the discovery of services in selected publications. Grouping services for web service discovery is critical for quickly finding the most accurate service. LDA (Latent Dirichlet Allocation) [S5,S8,S24], BTM (Bi-term Topic Model) [S6,S34], LSTM (Long Short-Term Memory) [S3], Bi-LSTM [S37], and Gauss LDA [S24] are the most common methods employed for the clustering or classification of web services. In Fig. 5, the publications indicate the methods used for clustering in the web service discovery process. As seen in Fig. 5, the most commonly preferred methods for clustering in publications are LDA and BTM.

When grouping the services, they are clustered according to the similarities between the services by considering the keywords and the words used in the service descriptions. Insufficient explanation texts with selected keywords pose a problem for the discovery process. Differences between the natural language used by the user while searching and the language used are another problem. Researchers use different algorithms to overcome these problems. As shown in Fig. 6, the most commonly used algorithms of the researchers are WordNet [S6,S20,S21], Matchmaker [S12,S21,S29] and K-Means [S2,S12] algorithms.

When examining the selected articles, it becomes evident that WordNet, matchmaking, and K-means algorithms are the most commonly used ones. These algorithms classify services based on similarities and mutual relationships. By considering the similarities between clustered web services and the user's query, the most relevant services are provided to the user. Various approaches and meta-heuristic algorithms aim to deliver the fastest and most accurate service to the user in the expanding web service environment. Due to the continuously increasing number and dynamic nature of web services, researchers need to constantly enhance these studies.

4.4. RQ-4: To what extent is the implementation strategy cloud-based?

Because of ongoing improvements to the underlying network, users now have better access to the Internet at higher speeds. Similarly, cloud computing is expanding rapidly. The transition of service providers to cloud technology will expand the number and quality of services available to users. In this subsection, the cloud computing connections of the approaches employed for the discovery of services in the selected publications are examined. Fig. 7 shows that cloud technology is becoming increasingly popular among web services, but only a small percentage of the papers we looked at dealt with cloud computing. The lack of knowledge about the advantages of cloud computing and the lack of confidence in this technology are the main reasons for its insufficient use. Only 28% of the selected papers employ cloud computing, whereas 72% do not. With cloud computing's rising flexibility and performance, cloud computing's application in the service discovery process will rise.

4.5. RQ-5: To what extent is the implementation strategy data or service-centric?

Web services can be defined as structures that provide services to other systems over the HTTP protocol. In general, web services can be grouped into three basic categories: service providers, users, and intermediaries between service providers and users. Service providers advertise their services through intermediaries, which can lead to multiple services being present on intermediaries, resulting in complexity. Upon analyzing the published papers, we observed that certain operations are performed on the intermediaries' databases before a user's query is processed. These operations include clustering services

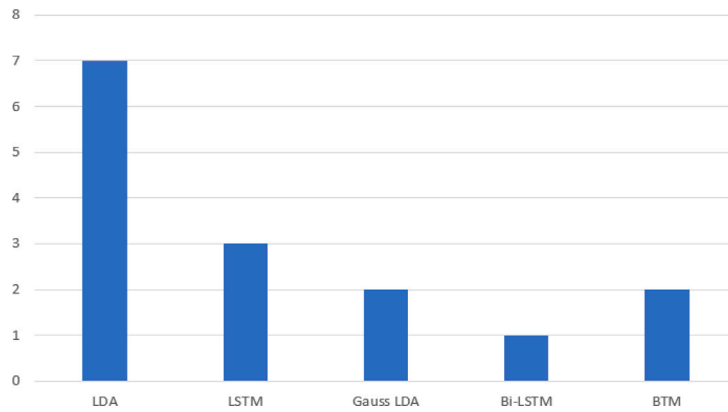


Fig. 5. Clustering and classification approaches.

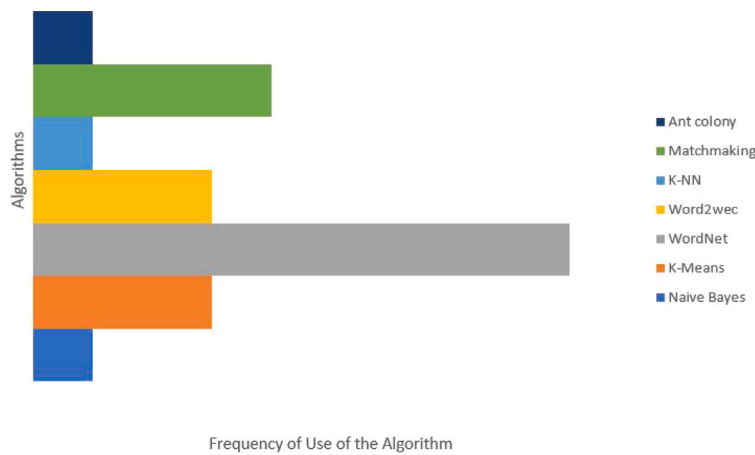


Fig. 6. Types of algorithms.

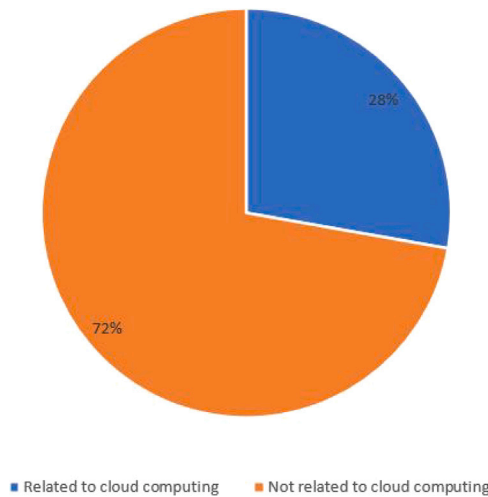


Fig. 7. Cloud computing usage chart.

based on their similarities, indexing services, employing word embedding techniques for missing description texts, and translating from WSDL (Web Services Description Language) to OWL-S (Ontology Web Language for Services). However, as depicted in the Fig. 8, many researchers tend to carry out these operations in response to the user's query and do not perform any proactive actions prior to receiving the query.

Typically, agents take no action before receiving a user request, and their search activity begins with a user query. While 67 percent of the selected publications did not take any action before the investigation, it is seen that pre-processing was performed on the database in other publications. These activities, which are carried out independently of the user's request, can sometimes be advantageous and sometimes harmful to the process due to the additional burden they bring. The fact that the majority of the researchers in the publications we reviewed did not implement the pre-examination procedures suggests that the procedures to be performed are more likely to adversely affect the discovery process or have low cost-benefit ratios. User queries are the most convenient way to start the dynamic service discovery process.

4.6. RQ-6: What are the challenges and possible solutions in this domain?

Examining the challenges reveals that they may be categorized into two distinct categories: dataset and user inquiry. In general, researchers confront five challenges that fall into two categories. The fact that certain problems have several solutions shows that researchers employ innovative approaches. The first obstacle that researchers face is a lack of datasets to analyze. Because rebuilding a dataset is a time-consuming and costly procedure, it is crucial for researchers to locate an acceptable dataset for their studies. In the analyzed papers [2,4,S22,S30,S31,S33], researchers resolved this issue by utilizing a publicly available dataset. The most often utilized datasets are those from SAWSDL-TC3, OWLS-TC4, and ProgrammableWeb.

The second issue identified with the dataset is erroneous content in the services and replica services [15,18,S20]. For the sake of the study's integrity, our researchers attempted to eliminate from the study

Table 5
Challenges and solution suggestions.

Category	Challenges (C1-C5)	Proposed solutions (S1-S7)	Reference
Dataset	C1.Obtaining a dataset	S1.Using a public dataset	2,4,22,30, 31,33
	C2.Services and replicas with incorrect content	S2.Removing services and replicas with incorrect or incomplete content	15,18,20
	C3.Insufficient description of services in the dataset	S3.Improving service definition by using words with similar meaning and syntactic properties with active word	40
Query		S4.Adding random sentences to the category to which the service belongs	13
	C4.User's natural language input (Everyday Language Usage)	S5.Removing stop words and abbreviations	3,11,12,18, 20,33,36,40
	C5.Restricted data entry in query	S6.Using synonyms/variants of keywords in the query	5,7,40
		S7.Using the Compromise Library (consisting of synonyms and abbreviations)	2,12

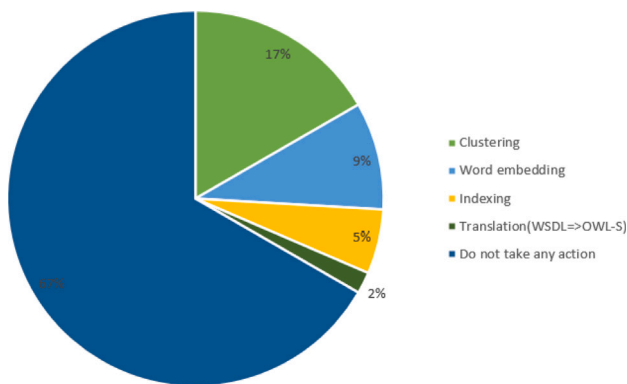


Fig. 8. Pre-inquiry activities.

the services in the dataset with incorrect content and duplicates. The final issue we find with datasets is insufficient explanations of services in the dataset [13,S40]. In the process of service discovery, it is difficult to identify the correct service due to the lack of adequate service descriptions. In order to address this issue, researchers improved the service description [S40] by adding terms with similar semantic and syntactic features to the part containing the service description. Alternatively, the definition of the service could be enhanced by using random sentences from the category to which the service belongs [13].

A further difficulty noticed during the service discovery phase is the usage of ordinary language and abbreviations within the query [3, 11,12,18,S20,S33,S36,S40]. When employing deep learning algorithms and machine learning algorithms, the presence of superfluous words and abbreviations in the query prolongs the process and reduces the precision of the output. Researchers resolved this issue by excluding superfluous words and abbreviations from the query. Another issue noticed [2,5,7,12,S40], is short (insufficient) data entry in the query. Incomplete query data prevents the user's request from being properly processed. As a result of the query, it is possible that the user will receive too many services or that no services will be found. Two solutions have been identified by researchers for this issue. The first is to use synonyms and variants of the query's keywords [5,7,S40], and the second is to leverage a consensus library comprised of synonyms and abbreviations [2,12]. Table 5 contains a listing of the issues faced in the selected articles and the offered solutions. It is of the utmost importance for researchers to be predictable and to have a general understanding of solution methodologies.

4.7. An overview of responses to research questions

In this SLR study, several research questions related to service discovery are addressed. According to the findings of the study, syntactic, semantic-conscious and ontology-based methods are widely used among service discovery methods. The semantic conscious method is the most popular among these studies. Although service discovery is discovery-oriented, it also includes complementary functions such as health checking and load balancing. Among meta-heuristic techniques and algorithms, algorithms such as WordNet, Matchmaker, K-Means are used for clustering (LDA, BTM) and Web service discovery. Important metrics such as response time, success rate, and scalability are used to evaluate service discovery performance. The challenges with service discovery include scalability, heterogeneity, and information overload. There are also limitations such as dependency on static information and insufficient support for dynamic environments. Table 6 provides an overview of the research questions in this study.

5. Conclusion

The difficulty in locating useful web services is growing as the number of available services grows. However, web services might be closed or their functionalities might alter during the process of introducing new services. A load balancer may be required if several services are responsible for the same function. Users disregard a service that cannot be located. Researchers who attempted to find answers to the fundamental issues plaguing the discovery of services aided the discovery process and acted as a guide for subsequent investigators. We observed that the syntactic approach, which is based on the classic keyword-matching premise, has been largely replaced by the semantically-aware approach. Examining the literature on such discovery process cornerstones as "Health check", "load balance", and "service discovery" leads one to the conclusion that "service discovery" receives a lot of attention from researchers. The majority of studies using methodologies and algorithms for clustering services in the online service discovery process favor LDA and WordNet. The majority of researchers do not study cloud computing, despite its rapid development, according to an analysis of published works on the topic of cloud computing's application in the discovery of new services. It is expected, however, that as cloud computing continues to advance, service providers will increasingly favor it, and the cloud's role in locating relevant services will grow. Last but not least, while looking at research on online services that are stored without user querying, it can be noticed that clustering and indexing are usually conducted on the services, but the rate of no activity is considerably greater. This suggests that the researchers do not prefer procedures that are executed regardless of the user's request. For researchers who will work in the field of web service discovery in the future, there are other approaches not covered in the

Table 6
A summary of main findings.

Research Question	Findings
RQ-1: Service Discovery Method	Syntactic, semantic-conscious, and ontology-based methods are commonly used for service discovery. The semantic-conscious method is the most popular.
RQ-2: Complementary Functions	Service discovery is primarily focused on discovery itself, but also includes health check and load balancing features.
RQ-3: Meta-heuristic Techniques and Algorithms	Commonly used techniques include clustering (LDA, BTM) and algorithms (WordNet, Matchmaker, K-Means) for web service discovery.
RQ-4: Performance Evaluation Metrics	Key metrics for evaluating service discovery performance include response time, success rate, and scalability.
RQ-5: Challenges and Limitations	Challenges in service discovery include scalability, heterogeneity, and information overload. Limitations include the reliance on static information and lack of support for dynamic environments.
RQ-6: Future Research Directions	Potential future research directions include improving efficiency, addressing security concerns, and integrating AI techniques for intelligent service discovery.

papers discussed that can be used to improve web service discovery. One such approach is the Semantic-Aware Web Service Discovery Approach, which utilizes semantic inferences to enhance the accuracy and effectiveness of service discovery. This approach has been employed in numerous studies and has yielded successful results. Researchers can consider advanced classification algorithms such as Boosting and Convolutional Neural Networks (CNN) that have shown promising results in various domains to improve the accuracy of service classification. Additionally, alternative methods such as Latent Semantic Analysis (LSA) and ConceptNet can be explored to capture word relationships and semantic similarities. LSA, in particular, has proven successful in uncovering hidden semantic relationships in textual data. By adopting these recommended approaches, researchers can contribute to the advancement of web service discovery.

Funding

No funding received.

CRediT authorship contribution statement

Ahmet Vedat Tokmak: Research, Writing the manuscript. **Akhan Akbulut:** Research, Writing the manuscript. **Cagatay Catal:** Research, Writing the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

- [1] D.H. Elsayed, A. Salah, Semantic web service discovery: a systematic survey, in: 2015 11th International Computer Engineering Conference, ICENCO, IEEE, 2015, pp. 131–136.
- [2] R. Phalnikar, P.A. Khutade, Survey of QoS based web service discovery, in: 2012 World Congress on Information and Communication Technologies, IEEE, 2012, pp. 657–661.
- [3] C. Pautasso, E. Wilde, RESTful web services: principles, patterns, emerging technologies, in: Proceedings of the 19th International Conference on World Wide Web, 2010, pp. 1359–1360.

- [4] W. Rong, K. Liu, A survey of context aware web service discovery: from user's perspective, in: 2010 Fifth IEEE International Symposium on Service Oriented System Engineering, IEEE, 2010, pp. 15–22.
- [5] V.X. Tran, H. Tsuji, A survey and analysis on semantics in QoS for web services, in: 2009 International Conference on Advanced Information Networking and Applications, IEEE, 2009, pp. 379–385.
- [6] W.J. Obidallah, B. Raahemi, U. Ruhi, Clustering and association rules for web service discovery and recommendation: a systematic literature review, SN Comput. Sci. 1 (2020) 1–33.
- [7] B. Kitchenham, P. Brereton, M. Turner, M. Niazi, S. Linkman, R. Pretorius, D. Budgen, The impact of limited search procedures for systematic literature reviews—A participant-observer case study, in: 2009 3rd International Symposium on Empirical Software Engineering and Measurement, IEEE, 2009, pp. 336–345.
- [8] J. Kopecký, T. Vitvar, C. Bournez, J. Farrell, Sawsdl: Semantic annotations for wsd and xml schema, IEEE Internet Comput. 11 (6) (2007) 60–67.
- [9] W.J. Obidallah, B. Raahemi, A survey on web service discovery approaches, in: Proceedings of the Second International Conference on Internet of Things, Data and Cloud Computing, 2017, pp. 1–8.
- [10] K.A. Alam, V. Chang, R. Ahmad, M. Tahir, A. Akhuzada, A.V. Vasilakos, Clustering and classification techniques for web service discovery: A systematic review, Int. J. Inf. Manage. (2016) 1–29.
- [11] S. Sagayaraj, M. Santhoshkumar, A survey on clustering methods in web service discovery, in: 2017 4th International Conference on Electronics and Communication Systems, ICECS, IEEE, 2017, pp. 189–194.
- [12] M. Malaimalavathani, R. Gowri, A survey on semantic web service discovery, in: 2013 International Conference on Information Communication and Embedded Systems, ICICES, IEEE, 2013, pp. 222–225.
- [13] N. Agarwal, G. Sikka, L.K. Awasthi, A systematic literature review on web service clustering approaches to enhance service discovery, selection and recommendation, Comp. Sci. Rev. 45 (2022) 100498.
- [14] A. Alkhateeb, C. Catal, G. Kar, A. Mishra, Hybrid blockchain platforms for the Internet of Things (IoT): A systematic literature review, Sensors 22 (4) (2022) 1304.
- [15] C. Catal, On the application of genetic algorithms for test case prioritization: a systematic literature review, in: Proceedings of the 2nd International Workshop on Evidential Assessment of Software Technologies, 2012, pp. 9–14.
- [16] R. van Dinter, B. Tekinerdogan, C. Catal, Predictive maintenance using digital twins: A systematic literature review, Inf. Softw. Technol. (2022) 107008.
- [17] N. Slob, C. Catal, A. Kassahun, Application of machine learning to improve dairy farm management: A systematic literature review, Prevent. Vet. Med. 187 (2021) 105237.
- [18] R. Dohmen, C. Catal, Q. Liu, Computer vision-based weight estimation of livestock: a systematic literature review, N. Z. J. Agric. Res. 65 (2–3) (2022) 227–247.

Primary studies (Sources reviewed in the SLR)

- [S1] Z. Houmani, D. Balouek-Thomert, E. Caron, M. Parashar, Enhancing microservices architectures using data-driven service discovery and QoS guarantees, in: 2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing, CCGRID, IEEE, 2020, pp. 290–299.
- [S2] S. Garba, R. Mohamad, N. Saadon, Self-adaptive mobile web service discovery framework for dynamic mobile environment, J. Syst. Softw. 184 (2022) 111120.

- [S3] L. Guodong, Q. Zhang, Y. Ding, W. Zhe, Research on service discovery methods based on knowledge graph, *IEEE Access* 8 (2020) 138934–138943.
- [S4] M. Fang, D. Wang, Z. Mi, M. Obaidat, Web service discovery utilizing logical reasoning and semantic similarity, *Int. J. Commun. Syst.* 31 (10) (2018) e3561.
- [S5] C. Sun, L. Lv, G. Tian, Q. Wang, X. Zhang, L. Guo, Leverage label and word embedding for semantic sparse web service discovery, *Math. Probl. Eng.* (2020) (2020).
- [S6] J. Wang, P. Gao, Y. Ma, K. He, P. Hung, A web service discovery approach based on common topic groups extraction, *IEEE Access* 5 (2017) 10193–10208.
- [S7] Z. Huang, W. Zhao, Combination of ELMo representation and CNN approaches to enhance service discovery, *IEEE Access* 8 (2020) 130782–130796.
- [S8] C. Li, R. Zhang, J. Huai, X. Guo, H. Sun, A probabilistic approach for web service discovery, in: 2013 IEEE International Conference on Services Computing, *IEEE*, 2013, pp. 49–56.
- [S9] M. Adarme, M. Jimeno, QoS-based pattern recognition approach for web service discovery: *Ar_WSDS*, *Appl. Sci.* 11 (17) (2021) 8092.
- [S10] J. Dantas, H. Lira, B. Muniz, T. Nunes, P. Farias, Semantic web services discovery adopting serin, in: 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing, *IEEE*, 2015, pp. 387–394.
- [S11] C. Pushpa, G. Deepak, A. Kumar, J. Thriveni, K. Venugopal, OntoDisco: improving web service discovery by hybridization of ontology focused concept clustering and interface semantics, in: 2020 IEEE International Conference on Electronics, Computing and Communication Technologies, *CONECCT*, *IEEE*, 2020, pp. 1–5.
- [S12] B. Madhu, J. Sinchana, Mining underlying interface semantics based web services discovery approach, in: 2019 1st International Conference on Advances in Information Technology, *ICAIT*, *IEEE*, 2019, pp. 453–461.
- [S13] S. Li, H. Luo, G. Zhao, Bi-hptm: An effective semantic matchmaking model for web service discovery, in: 2020 IEEE International Conference on Web Services, *ICWS*, *IEEE*, 2020, pp. 433–440.
- [S14] H. Fethallah, S. Ismail, M. Mohamed, T. Zeyneb, An outranking model for web service discovery, in: 2017 International Conference on Mathematics and Information Technology, *ICMIT*, *IEEE*, 2017, pp. 162–167.
- [S15] B. Cheng, C. Li, J. Chen, Semantics mining&indexing-based rapid web services discovery framework, *IEEE Trans. Serv. Comput.* 14 (3) (2018) 864–875.
- [S16] S. Samir, A. Sarhan, A. Aalgaway, Context-based web service discovery framework with qos considerations, in: 2017 11th International Conference on Research Challenges in Information Science, *rcis*, *IEEE*, 2017, pp. 146–155.
- [S17] A. Eddine, B. M'hamed, T. Sifaqes, Y. Hamza, Evaluation tool for contextual similarity measures in web services discovery approaches, in: 2017 International Conference on Mathematics and Information Technology, *ICMIT*, *IEEE*, 2017, pp. 168–176.
- [S18] B. Cheng, S. Zhao, C. Li, J. Chen, MISDA: web services discovery approach based on mining interface semantics, in: 2016 IEEE International Conference on Web Services, *ICWS*, *IEEE*, 2016, pp. 332–339.
- [S19] H. Naim, M. Aznag, M. Quafafou, N. Durand, Probabilistic approach for diversifying web services discovery and composition, in: 2016 IEEE International Conference on Web Services, *ICWS*, *IEEE*, 2016, pp. 73–80.
- [S20] B. Cheng, S. Zhao, C. Li, J. Chen, A web services discovery approach based on mining underlying interface semantics, *IEEE Trans. Knowl. Data Eng.* 29 (5) (2016) 950–962.
- [S21] R. Hammami, H. Bellaaj, A. Kacem, A novel approach for semantic web service discovery, in: 2016 IEEE 25th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises, *WETICE*, *IEEE*, 2016, pp. 250–252.
- [S22] Y. Wu, C. Yan, Z. Ding, G. Liu, P. Wang, C. Jiang, M. Zhou, A multilevel index model to expedite web service discovery and composition in large-scale service repositories, *IEEE Trans. Serv. Comput.* 9 (3) (2015) 330–342.
- [S23] N. Seghir, O. Kazar, K. Rezeg, S. Bouekkache, A semantic web services discovery approach based on a mobile agent using metadata, *Int. J. Intell. Comput. Cybern.* (2017).
- [S24] G. Tian, P. Liu, Y. Peng, C. Sun, Tagging augmented neural topic model for semantic sparse web service discovery, *Concurr. Comput.: Pract. Exper.* 30 (16) (2018) e4448.
- [S25] L. Duan, H. Tian, Collaborative web service discovery and recommendation based on social link, *Future Internet* 9 (4) (2017) 63.
- [S26] P. Rodriguez-Mier, C. Pedrinaci, M. Lama, M. Mucientes, An integrated semantic web service discovery and composition framework, *IEEE Trans. Serv. Comput.* 9 (4) (2015) 537–550.
- [S27] G. Khan, S. Sengupta, A. Sarkar, N. Debnath, Web service discovery in enterprise cloud bus framework: T vector based model, in: 2015 IEEE 13th International Conference on Industrial Informatics, *INDIN*, *IEEE*, 2015, pp. 1672–1677.
- [S28] K. Moradyan, O. Bushehrian, R. Akbari, A query ontology to facilitate web service discovery, in: 2015 2nd International Conference on Knowledge-Based Engineering and Innovation, *KBEI*, *IEEE*, 2015, pp. 202–206.
- [S29] S. Kamath, V. Ananthanarayana, Change propagation based incremental data handling in a web service discovery framework, in: 2014 IEEE International Symposium on Signal Processing and Information Technology, *ISSPIT*, *IEEE*, 2014, pp. 000474–000479.
- [S30] W. Lo, J. Yin, Z. Wu, Accelerated sparse learning on tag annotation for web service discovery, in: 2015 IEEE International Conference on Web Services, *IEEE*, 2015, pp. 265–272.
- [S31] A. Alnahdi, S. Liu, A. Melton, Enhanced web service matchmaking: A quality of service approach, in: 2015 IEEE World Congress on Services, *IEEE*, 2015, pp. 341–348.
- [S32] M. Alshafaey, A. Saleh, M. Alrahamawy, A new cloud-based classification methodology (CBCM) for efficient semantic web service discovery, *Cluster Comput.* 24 (3) (2021) 2269–2292.
- [S33] A. Baskara, R. Sarno, Web service discovery using combined bi-term topic model and W DAG similarity, in: 2017 11th International Conference on Information & Communication Technology and System, *ICTS*, *IEEE*, 2017, pp. 235–240.
- [S34] M. Omidipoor, A. Toomanian, N.N. Samany, A. Mansourian, Knowledge discovery web service for spatial data infrastructures, *ISPRS Int. J. Geo-Inf.* 10 (1) (2020) 12.
- [S35] A. Abdelli, W. Serrai, L. Mokdad, A novel and efficient index based web service discovery approach, *Comput. Stand. Interfaces* 80 (2022) 103586.
- [S36] H. Ye, B. Cao, J. Chen, J. Liu, Y. Wen, J. Chen, A web services classification method based on GCN, in: 2019 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking, *ISPA/BDCLOUD/SocialCom/SustainCom*, *IEEE*, 2019, pp. 1107–1114.
- [S37] H. Ye, B. Cao, Z. Peng, T. Chen, Y. Wen, J. Liu, Web services classification based on wide & Bi-LSTM model, *IEEE Access* 7 (2019) 43697–43706.
- [S38] J. Sha, Y. Du, L. Qi, A user requirement oriented Web service discovery approach based on logic and threshold Petri net, *IEEE/CAA J. Autom. Sin.* 6 (6) (2019) 1528–1542.
- [S39] A. Sullivan, M. Candra, Web service recommendation system using history and quality of service, in: 2019 International Conference on Data and Software Engineering, *ICoDSE*, *IEEE*, 2019, pp. 1–6.
- [S40] G. Tian, S. Zhao, J. Wang, Z. Zhao, J. Liu, L. Guo, Semantic sparse service discovery using word embedding and Gaussian LDA, *IEEE Access* 7 (2019) 88231–88242.
- [S41] F. Zeshan, R. Mohamad, M. Ahmad, M. Othman, A. Elhag, S. Hussain, I. Babar, Context-aware ontology and web services discovery for distributed embedded real-time systems, *Malaysian J. Comput. Sci.* 32 (3) (2019) 186–208.
- [S42] L. Zhang, OWL-s based web service discovery in distributed system, in: 2014 IEEE Workshop on Electronics, Computer and Applications, *IEEE*, 2014, pp. 882–885.
- [S43] A. Benghida, M. Boufaida, Web services discovery in P2P networks based on clustering, in: 2014 1st International Conference on Information and Communication Technologies for Disaster Management, *ICT-DM*, *IEEE*, 2014, pp. 1–7.
- [S44] Y. Du, X. Qin, Multi-strategy web service discovery for smart government, in: *Applied Mechanics and Materials*, Vol. 536, *Trans Tech Publications Ltd*, 2014, pp. 625–631.
- [S45] X. Han, X. Li, A service discovery algorithm based on network topology management, in: *Applied Mechanics and Materials*, Vol. 513, *Trans Tech Publications Ltd*, 2014, pp. 1182–1186.
- [S46] X. Tang, D. Luo, J. Huan, An optimized method for web service discovery based on semantics, in: *ITITS*, 2017, pp. 1–7.
- [S47] A. Nikolov, D. Petrova-Antonova, Repository platform for RESTful web services, in: 2022 International Conference Automatics and Informatics, *ICAI*, *IEEE*, 2022, pp. 292–297.
- [S48] S. Garba, R. Mohamad, N. Saadon, Self-adaptive mobile web service discovery approach based on modified negative selection algorithm, *Neural Comput. Appl.* (2022) 1–23.
- [S49] H. Bu, J. Xia, Q. Wu, L. Chen, Relationship discovery and hierarchical embedding for web service quality prediction, *Comput. Intell. Neurosci.* 2022 (2022).
- [S50] M. Halilali, E. Gouarderes, M. Gaio, F. Devin, Geospatial web services discovery through semantic annotation of WPS, *ISPRS Int. J. Geo-Inf.* 11 (4) (2022) 254.
- [S51] G. Senthil Kumar, K. Ramana, R. Mohana, R. Aluvalu, V. Gunjan, N. Singh, An effective bootstrapping framework for web services discovery using trigram approach, *Mob. Inf. Syst.* 2022 (2022) 1–12.
- [S52] W. Obidallah, B. Raahemi, W. Rashideh, Multi-layer web services discovery using word embedding and clustering techniques, *Data* 7 (5) (2022) 57.
- [S53] Z. Huang, W. Zhao, A semantic matching approach addressing multidimensional representations for web service discovery, *Expert Syst. Appl.* 210 (2022) 118468.
- [S54] N. Swetha, G. Karpagam, Lexicon ontology driven concept lattice framework for semantic web service discovery, in: 2022 6th International Conference on Computing Methodologies and Communication, *ICCMC*, *IEEE*, 2022, pp. 1428–1435.