

Review

Cloud Computing Services: Taxonomy of Discovery Approaches and Extraction Solutions

Fathey Mohammed ^{1,*}, Abdullah Marish Ali ², Abdullah Saad Al-Malaise Al-Ghamdi ²,
Fawaz Alsolami ², Siti Mariyam Shamsuddin ³ and Fathy E. Eassa ²

¹ School of Computing, Universiti Utara Malaysia (UUM), Sintok 06010, Kedah Darul Aman, Malaysia

² Faculty of Computing and Information Technology, King Abdulaziz University (KAU),
Jeddah 21589, Saudi Arabia; ammali@kau.edu.sa (A.M.A.); aalmalaise@kau.edu.sa (A.S.A.-M.A.-G.);
falsolami1@kau.edu.sa (F.A.); feassa@kau.edu.sa (F.E.E.)

³ School of Computing, Universiti Teknologi Malaysia (UTM), Skudai 81310, Malaysia; mariyam@utm.my

* Correspondence: fathey.mohammed@uum.edu.my or fathey.m.ye@gmail.com

Received: 19 July 2020; Accepted: 11 August 2020; Published: 13 August 2020



Abstract: Cloud computing offers new features of sharing resources and applications to meet users' computing requirements. It is a model by which the users can access computing resources as services offered on the Internet (cloud services). Cloud service providers offer a highly diverse range of asymmetric cloud services with heterogeneous features, which makes it difficult for the users to find the best service that fits his needs. Many research studies have been done on cloud service discovery, and several models and solutions that applied different techniques have been proposed. This paper aims at presenting the state of the art in the area of cloud services discovery by exploring the current approaches, techniques, and models. Furthermore, it proposes a taxonomy of cloud service discovery approaches. An integrative review approach was used to explore the related literature. Then, by analyzing the existing cloud service discovery solutions, a taxonomy of discovery approaches was suggested based on several perspectives including the discovery environment and the discovery process methods. The proposed taxonomy allows easily classifying and comparing cloud services discovery solutions. Moreover, it may reveal issues and gaps for further research and expose new insights for more innovative and effective cloud services discovery solutions.

Keywords: discovery approaches; cloud services; classification; extraction

1. Introduction

In recent years, the essential communication media that humans use for fulfilling their needs is the Internet. People can perform tasks related to education, business, social networking, and many others using Internet. Advances in information and communication-related technologies have encouraged organizations and enterprises to re-engineer their process to efficiently use these technologies for lowering operational cost, increasing the scalability, improving the performance, and utilizing their resources in efficient ways [1]. Cloud computing emerged to meet users' computing requirements by offering new features designed to enhance the sharing of resources and applications [2].

Cloud computing is a model in which the users are allowed to access computing resources (hardware and software) via the Internet as hosted services that can be scaled dynamically according to users' needs. These services are called cloud services. The cloud service provider and cloud customer are the two main players in the cloud computing environment. The cloud service providers hosts cloud services in their data centers and allow users (cloud customers) to access and use these resources on a rent basis (pay-per-use) [3]. Cloud services have become very important for users, more specifically organizations and enterprises; thus, there is competition between providers to offer services with

various features and performance attributes [4]. Accordingly, asymmetric services with heterogeneous types and features have been developed [5].

Due to the growth in the number of cloud providers and the offered cloud services, discovering the best cloud service that satisfies the user requirements with specific Quality of Service (QoS) attributes has become a challenge for the cloud customers [5,6]. Cloud service discovery is a process of addressing the problem of identifying the most suitable service or the set of services that best match with specific requirements. In the literature, cloud services discovery has been studied, and several models have been proposed using different approaches. Ghazouani and Slimani [7] conducted a review study on the approaches and solutions of cloud services discovery. The focus was on comparing the cloud services discovery approaches suggested in the literature based on the delivery model, the applied technique, and the service representation method. Alkalbani and Hussain [8] also explored cloud service discovery trends and issues. The existing discovery approaches were reviewed and classified based on approach model, type of service, representation, model of evaluation, user's preferences and repositories. In addition, Sun, et al. [9] presented a review study on cloud service discovery solutions and analyzed the discovery approaches based on the techniques of decision making, the model of data representation, the characteristics of cloud services, the context, and the purposes. Furthermore, Ali, et al. [10] reviewed the existing cloud service discovery approaches and identified the limitations and weaknesses of existing solutions. However, more exploration and investigation on the related literature may promote research on cloud service discovery approaches and expose new insights for more innovative and effective solutions. Accordingly, the objectives of this study are:

1. Exploring the exciting cloud service discovery solutions, and
2. Proposing a taxonomy of the current approaches based on different discovery characteristics.

The next section introduces cloud computing and its characteristics and models. This is followed by cloud service discovery concepts. Then, in the following section, the existing solutions of cloud service discovery are presented. Next, a classification of the cloud services discovery approaches is suggested. After that, the proposed taxonomy of discovery approaches is discussed. Finally, the paper ends with the conclusion.

2. Cloud Computing

Cloud computing is an innovative model of sharing computing resources. It provides users with virtualized hardware and software resources as a service via the Internet using the pay-as-you-go method. The pay-as-you-go method allows users to pay for the resources they use rather than pay for the cost of all resources they need [11]. As reported by the National Institute of Standards and Technology (NIST), cloud computing is a model of accessing shared and configurable computing resources including servers, storage, platforms, and applications via the Internet. These resources can be accessed on-demand and rapidly provisioned with minimal interaction with the service provider [12]. The cloud computing providers provide end users with IT resources (physical or virtual) and applications as services over the Internet [13,14]. The cloud represents a network of data centers, where each data center contains several thousand computers that are connected together to accomplish the required task. These data centers allow users to access the applications, platforms, and services available on the Internet.

Cloud computing model mainly comprises three types of services: Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). First, in SaaS, consumers are given the capability of using the provider's applications through the web (such as Salesforce.com, Google Docs, Gmail, and Spreadsheets) operating on the cloud infrastructure [15]. Therefore, it is not required for end users to download, install, configure, and run the software applications on their own computing terminals. SaaS is a real software service that is accessed and used by end users through a user interface such as an Internet browser. The user may be given permission to control the application using some limited configuration settings. The underlying resources including operating

systems and servers are managed and manipulated by the providers. Second, PaaS is built on the cloud infrastructure and involves providing capabilities to the consumer to help developing and deploying applications using platforms and programming languages supported by the provider [16]. PaaS provides a higher-level platform (such as Google App Engine, Microsoft Azure, and Salesforce platform) where developers can develop customized applications without involving in managing or configuring the cloud infrastructure. Finally, in IaaS, computing resources are provided virtually (in the form of virtual machines) to the consumer such as processing (such as Amazon EC2 service), storage (such as Amazon S3 service), servers, and other essential computing resources [17]. So, consumers can deploy and run arbitrary software including applications and/or operating systems on these resources.

In addition, a cloud computing delivery model can be classified according to the cloud data centers ownership or type of cloud combined for single or multiple cloud environments into private clouds, public clouds, hybrid clouds, and community clouds. A company or an organization that fully owns a cloud data center is referred to as a private cloud [18]. The infrastructure and the place of the running applications and data of people or organizations using these applications are fully controlled by the cloud owner [19]. When the data centers of hardware and software are run by a third party, the cloud is referred to as a public cloud. Public clouds offer their services to public customers and companies on the basis of pay-per-use [18]. Combination of private and public clouds constitutes a hybrid cloud, which allows an organization or a company to run certain applications on public clouds and others on an internal infrastructure. The benefit of this type of cloud is that the important applications and data will be kept inside the firewall, whereas the other applications run on the public clouds [20]. Many organizations from a specific community with common concerns collaborate their efforts in a shared cloud infrastructure called community clouds [21].

3. Cloud Service Discovery

Cloud service provisioning is the process of enabling the users to access and use the computing resources as services [22–24]. The service provider offers computing services based on the Service Level Agreement (SLA), which includes the QoS attributes. However, different service providers may offer different levels of QoS attributes including availability, scalability, elasticity, security, and exact billing. This creates a challenge for the users to find the service that fits his needs [6,25,26]. Accordingly, one of the main objectives of service discovery is to offer a reasonable comparison between the provided services, so that the users can compare and decide to select the service that satisfies their needs based on several predefined specifications [27]. The user should identify the functional, technical, and security specifications that the required service should meet [28]. Functional specifications may include tasks to be performed, cost policy, and service domain. Technical specifications include software service, software compatibility constraints, hardware requirements (storage capacity, server, multicore processing, and others), cloud deployment (private, public, hybrid, and community), operating system (e.g., single and multiple OS support), and cloud service model (IaaS, PaaS and SaaS). Security specification include rules and permissions, data encryption and deletion constraints, cloud/service provider location constraints, virtualization (virtual machine (VM) separation), and multitenancy policies.

As cloud providers offer a diversity of cloud services, it becomes a challenge for customers to select the service provider that can offer the most suitable services [29]. A lot of effort has been put into web services description and standardization. WSDL (Web Service Definition Language) was developed to describe web services. However, WSDL provides a technical description that includes service features and interface operations [30]. This description may be not sufficient to be used for describing cloud services, which are business services that need to be described from a business point of view. Business services are concerned with time, cost, and the SLA of the delivery process [31]. Therefore, to discover the most suitable cloud services that match the predefined requirements, a discovery system is necessary. A discovery system should be able to identify a similarity between the user service specifications and service provider offerings.

4. Methodology

Since the cloud services discovery is a mature research topic and the aim of this study is to propose a taxonomy of discovery approaches, the integrative review approach is followed in this study. The integrative review approach does not follow strict standards, and the general aim of the data analysis in this approach is to critically analyze and examine the literature [32]. Based on the objectives of this study, the search keywords (including cloud services discovery/description/extraction/selection) were selected to access appropriate studies. Initial literature searches yield many studies. To identify which studies are actually relevant, the authors scanned the titles and abstracts of retrieved articles. Only the studies that investigated cloud services discovery were included. Then, by reading full-text articles to ensure they meet the inclusion criteria, the relevant literature has been collected. Additionally, references in the selected studies were scanned to identify any article that may potentially be relevant. By having the final sample, appropriate information from each article was abstracted and reported. To propose a taxonomy of cloud services discovery approaches, the proposed solutions were analyzed against the main discovery characteristics including the environment and service selection or extraction process. Figure 1 represents how the study was conducted.

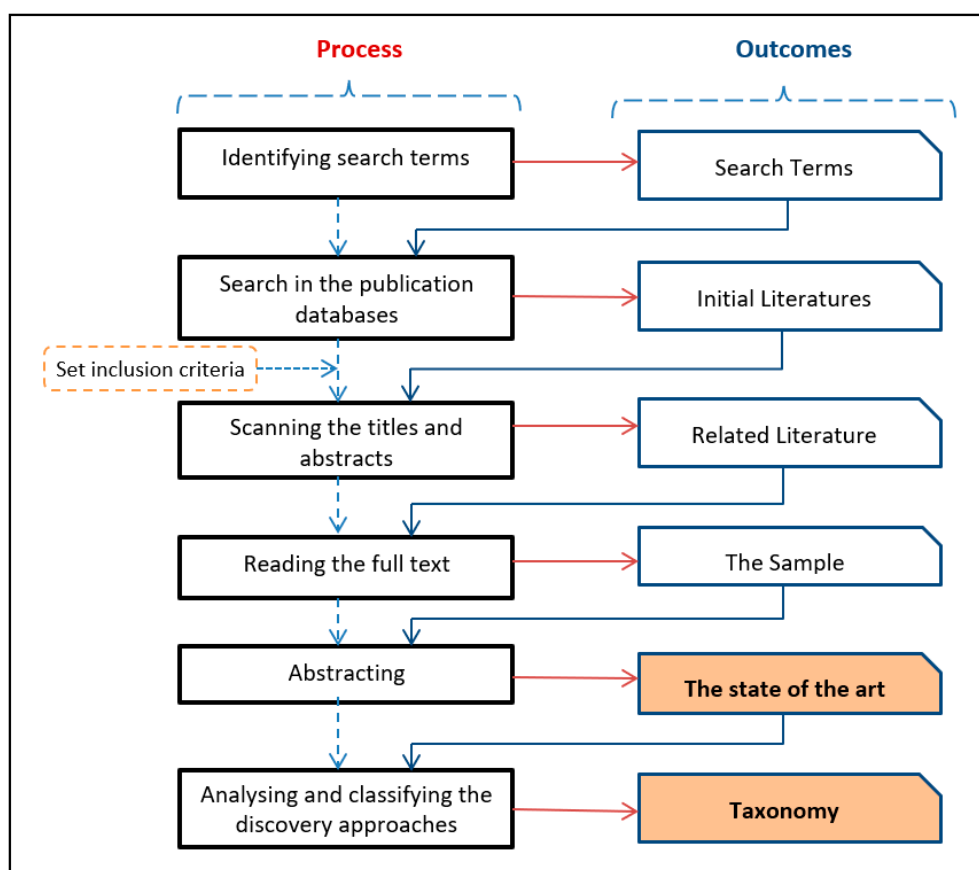


Figure 1. Review process.

5. Cloud Service Discovery Solutions: State of the Art

Many works have been done in the literature of cloud services discovery. This section explores these works and analyzes the proposed solutions to provide the state of the art in the area of cloud services discovery approaches. Although some researchers adapted the techniques used in SOA (Service-Oriented Architecture) and web services description [33,34], others applied more specific techniques that fit with cloud computing characteristics.

Kanagasabai [34] adopted a semantic-based approach using Web Ontology Language for Service (OWL-S) and proposed a novel cloud broker. An OWL-based matchmaking system was developed that can dynamically discover complex and constraint-based cloud services. However, the offered and requested services should be translated using a shared ontology among cloud providers. Tahamtan, et al. [35] developed cloud business ontology to help organizations with searching and selecting suitable cloud services. A model was proposed to integrate a unified business service and cloud ontology, which enables organizations to find the suitable cloud service using querying capabilities. A unified ontology helps record the inquired cloud services and build a link between the business functions and the available cloud services. This framework could also function as a service repository.

Moreover, a SaaS discovery system was developed by Afify, et al. [36] to establish SaaS service attributes using ontology. The SaaS discovery system comprised three key components: service registration, filtering and ranking, and discovery. Further, a model comprising ontology and semantic-based feature similarity matching was proposed to implement the matching process. However, the automatic discovering of SaaS services was not supported in this system and the cloud provider needs to subscribe and register its services in the system. Vasudevan, et al. [37] suggested integrating the semantic representation with the service profile to improve the automatic services discovery. The service profile includes the name of the service, price, and other features, and it is represented as ontology. The ontologies of all services are merged together to construct a general ontology. Further, Parhi, et al. [38] proposed a model for cloud service discovery using ontology and multi agents. The ontology, which was used to describe cloud services, contributed to increasing the retrieved cloud services that match the user requirements.

Kang and Sim [39] suggested a discovery approach by publishing the addresses of cloud services and making them available through a search engine. For extracting and retrieving the required services, ontology-based semantic filtering is applied. Semantic filtering identifies the relations between cloud services and advises the matching service. In addition, Kang and Sim [40] developed an ontology-based search engine that includes different reasoning techniques to identify the similarity among cloud services. The user should insert the information such as service type, functions, and price through an interface; then, the search engine retrieves the matching services based on the user enquiry using ontology. The retrieved services are presented according to a service feature such as the price. This search engine was improved by Sim [41] through involving a crawler that searches cloud services on the web and stored the retrieved services in a database instead of registering the services by cloud services providers.

Han and Sim [42] suggested a cloud service discovery approach that applies agents and ontology. Users search the services on the web using Google. However, an online search makes the discovery a time-consuming process, as it is done before matching the queried with retrieved services. Then, the system was improved by Han and Sim [43] concentrating on measuring the similarity among services. The improved system uses three reasoning methods: similarity, equivalent, and numerical reasoning. However, the system is still time-consuming, as it does online searching before beginning to check the similarity.

Chang, et al. [44] presented a prototype for an intelligent cloud service discovery system that integrates ontology with the mobile agent. It was assumed that the crawlers have direct access to cloud data centers by knowing the structure of the cloud data centers, the location, the schema, and other information. However, there is no naming convention or description standard for cloud services [5]. The evaluation showed a high degree of precision and low degree of recall.

Noor, et al. [45] developed a cloud crawler engine that uses cloud ontology to crawl various cloud data centers. The cloud services attributes are described, and a dataset is created as a repository to store cloud service descriptions. However, the evaluation revealed that some cloud service information in the dataset, including name and URL, were not linked with semantic meaning.

Nabeeh, et al. [46] suggested a discovery framework that is comprised of a set of software agents accomplishing tasks such as crawling, analyzing, and reasoning cloud services ontology, and suggesting recommendations. Users' profile and evaluators ranking reports were used to recommend cloud services. Nevertheless, the crawling process needs to be enhanced by using an extended ontology.

Hamza, et al. [47] suggested a cloud services discovery approach based on a mobile agent. An algorithm was proposed and implemented to compare the user query (based on search keywords) with the service description and perform filtering on the retrieved services to return the requested services. However, as cloud services provided by different cloud services providers may have different names, this algorithm may not be able to return the best matching services. Furthermore, it does not consider all attributes.

In addition, a centroid-based search engine (CB-Cloudle) was developed by [4,48]. CB-Cloudle includes crawlers to discover services offered by different cloud providers. It applies a k-means clustering algorithm to congregate cloud services into clusters based on the similarities among services from different providers. The crawling process in CB-Cloudle was done on the most popular cloud providers using their own crawlers.

Wheal and Yang [49] developed a cloud service discovery based on a recommender system that includes a search engine and a recommender. It applies collaborative and content-based methods to recommend the suitable cloud services for the user. The user can use the system through a website to search for cloud services. The high ranked cloud services, based on other users' ratings, would be returned. However, the reported results showed that a few types of cloud services were covered, and the accuracy was relatively low.

Alfazi, et al. [50] developed a search engine by which users can search a cloud service based on their own intention of use, cost, and other features. The proposed search engine can extract and identify cloud services automatically on the Web. The search engine includes a service identifier and a service feature extractor. The identifier uses a classification method to automatically identify the cloud service. The main function of the feature extractor is to extract the cloud service features. It uses the cluster method, as well as a service detection and tracking model to effectively identify cloud services with relevant service features and cost.

Karthikeyan and RS [51] proposed a cloud service recommendation system using semantic technologies. A fuzzy service ontology structure was suggested for the standard cloud services of PaaS, IaaS, and SaaS. Users can enquire of a cloud service in natural language; then, a query refiner refines the requirements using a fuzzy connection (AND, OR), and Resource Description Framework (RDF) queries are built to retrieve cloud services from the ontology repository. Information about the cloud services is generated by syntax analysis of the cloud service description documents using a natural language processing method. Then, this information is represented in the structure fuzzy service ontology.

Boukadi, et al. [52] presented a service-oriented architecture-based crawler for cloud service discovery. The crawler can parse the Internet and discover cloud providers and services by considering functional and non-functional properties. It provides easy modification capability for better performance. Moreover, the crawler includes a dedicated cloud service description ontology to save the search time and provide a better exploitation of the provider features.

Nabli, et al. [53] proposed a self-adaptive semantic based crawler that applies Latent Dirichlet Allocation (LDA) to effectively discover cloud services. In addition, an ontology is included to categorize cloud services. This ontology includes a set of concepts to enable the crawler to collect and categorize cloud services automatically. In addition, the crawler applies ranking techniques to list the URLs that will be parsed, so the relevant cloud services are retrieved in an efficient way. Then, a self-adaptive semantic-based crawler was created. The crawler also includes a learning function to update the cloud service ontology and enhance the system performance.

Parhi, et al. [54] proposed a multi-agent framework for the effective discovering and selecting of cloud services based on a standardized service registry that employs a semantic searching process.

Cloud ontology was used to represent a functional and non-functional QoS description for cloud services. In addition, an efficient semantic service matchmaking technique was developed to identify the degree of similarity between two cloud service descriptions by considering the customer preferences. This helps discover the suitable cloud service successfully with low response time.

Modi and Garg [55] proposed an integrated cloud service discovery based on QoS and semantic web. Cloud ontologies were developed to provide semantic descriptions of cloud services. QoS parameters including availability, throughput, response time, and cost were considered for quality assurance and user satisfaction.

Jiang, et al. [56] proposed a two-stage model to discover cloud services by integrating the service description and service tags. A Hierarchical Dirichlet Processes (HDP) model was proposed to assort cloud services into an optimal number of groups based on service description. Then, cloud services are ranked and recommended in each group using a Personalized PageRank algorithm based on the service tags. Cloud services with high ranking in the same group are identified as relevant services and recommended to the user.

Md, et al. [57] proposed an algorithm to discover cloud services based on the QoS using a decision tree classification algorithm. A cloud service resource registry algorithm was developed to enable service providers to register their services with the QoS attributes. Then, a search engine was implemented using the Split and Cache (SAC) algorithm to search the suitable cloud services with QoS attributes that meet the user application requirements. The proposed approach enables users to identify the cloud service possible via web GUI (Graphical User Interface) based on dynamic QoS attributes.

Alkalbani, et al. [58] proposed a cloud service repository model that includes two modules: a service repository and Harvesting as a Service (HaaS). The HaaS module is a user-friendly harvester that enables crawling cloud services information from various structured web portals and extracting real-time data from available different file formats. The service repository module that was implemented using web ontology language provides a centralized repository to ensure an efficient and effective cloud service discovery.

Table 1 summarizes the current cloud services discovery approaches. It presents the applied techniques, the cloud service model, and whether SLA or QoS attributes were considered.

In summary, works in the literature suggested various cloud service discovery solutions based on different techniques. Table 1 shows that three techniques were applied in the literature; ontology, agents, and machine learning. Ontology was widely used to develop cloud service discovery systems. It was used to represent and describe the retrieved services [34,35,37,38,52–55] and also was applied as a knowledge base for the matching purposes in the discovery process [35,36,39,40]. Table 1 also shows that some cloud services discovery solutions were implemented using agents [41,43,46,47,59]. In addition, other researchers applied agents and ontology [39,59–61]. They examined the cloud service discovery by developing cloud service brokers. Agents were used to search and check the similarities between the inquired service by the user and the available cloud services based on the ontology. Moreover, other researchers applied intelligent and machine learning techniques to identify the appropriate cloud service that matches the user's needs [34,50,51,53,56–58]. Mainly, solutions were applied to extract services of all cloud models (SaaS, PaaS, and IaaS), but some were specific for one model [5,36]. Other solutions did not classify the services based on the cloud models [34,35,38,46,47,49,54]. Furthermore, while most of studies considered SLA or QoS attributes, Noor, Sheng, Alfazi, Ngu and Law [45], Parhi, Pattanayak and Patra [38], Karthikeyan and RS [51], Boukadi, Rekik, Rekik and Ben-Abdallah [52], and Nabli, Djemaa and Amor [53] did not. In the next section, the cloud service discovery solutions are analyzed and classified constructing a taxonomy of different discovery approaches.

Table 1. Summary of Cloud Services Discovery Solutions in the Literature. IaaS: Infrastructure as a Service, PaaS: Platform as a Service, QoS: Quality of Service, SaaS: Software as a Service, SLA: Service Level Agreement.

Study	Techniques		Cloud Service Models			SLA/QoS
	Ontology Agents	Machine Learning	SaaS	PaaS	IaaS	
Kang and Sim [59]		√	√	√	√	√
Han and Sim [43]		√	√	√	√	√
Chang, Juang, Chang and Yen [44]	√	√	√	√	√	√
Kanagasabai [34]	√		-	-	-	√
Hamza, Aicha-Nabila, Okba and Youssef [47]		√	-	-	-	√
Sim [41]		√	√	√	√	√
Tahamtan, Beheshti, Anjomshoaa and Tjoa [35]	√		-	-	-	√
Noor, Sheng, Alfazi, Ngu and Law [45]			√	√	√	
Afify, Moawad, Badr and Tolba [36]			√			√
Vasudevan, Haleema and Iyengar [37]	√		√	√	√	√
Gong and Sim [4] and Gong and Sim [48]			√	√	√	√
Wheal and Yang [49]			-	-	-	√
Nabeeh, El-Ghareeb and Riad [46]		√	-	-	-	√
Parhi, Pattanayak and Patra [38]	√	√	-	-	-	-
Alfazi, Sheng, Babar, Ruan and Qin [50]			√	√	√	√
Karthikeyan and RS [51]	√	√	√	√	√	
Boukadi, Rekik, Rekik and Ben-Abdallah [52]	√		√	√	√	
Nabli, Djemaa and Amor [53]	√	√	√	√	√	
Parhi, Pattanayak and Patra [54]	√	√	-	-	-	√
Modi and Garg [55]	√		√	√	√	√
Jiang, Tao, Liu, Sun and Ling [56]		√	√	√	√	√
Md, Varadarajan and Mandal [57]		√	√	√	√	√
Alkalbani, Hussain and Kim [58]	√	√	√	√	√	

6. Classification of Cloud Service Discovery Approaches: Taxonomy

Many approaches have been proposed in the literature to construct cloud services discovery models/systems. These approaches are different in terms of the environment within which the discovery process is conducted. Based on some approaches, providers register their cloud services attributes in the repository of the cloud services discovery system (centralized). In other approaches, cloud services discovery is performed on cloud providers' websites (decentralized). Moreover, to model the collected information and represent the result, ontology or other formats are used. The ontology may be used as a knowledge base to perform semantic matching or to provide a semantic description of cloud services. Further, in terms of the methods/techniques used to implement the discovery process, researchers applied different techniques e.g., agents and machine learning. Table 2 represents the classification of the proposed solutions in the literature based on the discovery environment and the discovery technique.

Table 2. Classification of the Existing Cloud Service Discovery Approaches.

Discovery Approaches' Types		Centralized	Decentralized
Semantic	Agent based	Kang and Sim [60], Boukadi, Rekik, Rekik and Ben-Abdallah [52], Parhi, Pattanayak and Patra [54], and Parhi, Pattanayak and Patra [38]	Han and Sim [43], Chang, Juang, Chang and Yen [44]
	Intelligent	Kanagasabai [34], Nabli, Djemaa and Amor [53]	Alfazi, Sheng, Babar, Ruan and Qin [50], Karthikeyan and RS [51]
Non-Semantic	Agent based	-	Hamza, Aicha-Nabila, Okba and Youssef [47], Sim [41], Nabeeh, El-Ghareeb and Riad [46]
	Intelligent	Md, Varadarajan and Mandal [57], Alkalbani, Hussain and Kim [58]	Jiang, Tao, Liu, Sun and Ling [56]

6.1. Discovery Environment

Based on the discovery environment, the cloud services discovery solutions can be classified into four types: centralized semantic-based approaches, centralized non-semantic-based approaches, decentralized semantic-based approaches, and decentralized non-semantic-based approaches.

6.1.1. Centralized Semantic-Based Approaches

This type of discovery solution offers detailed information for all cloud services already existing in the market via a central node. A repository is an essential element in this approach to maintain the cloud services attributes. This repository typically allows cloud providers to add or update the descriptions of a list of services they provided. In addition, a repository provides a centralized place for consumers so they may select the preferred services. The repository works as a mediator between cloud providers and customers, and it helps to set up the association between them. The cloud providers publish their services by registering them in a repository and the user can choose based on their requirements. In this kind of discovery approach, the cloud services semantic description is represented by an ontology. A cloud service requester ontology is used to request executing the application demanding resources, and a service provider ontology is used to allocate resources in the form of services. Examples of this type of approach include the models proposed by Kang and Sim [60], Boukadi, Rekik, Rekik and Ben-Abdallah [52], Parhi, Pattanayak and Patra [54], Parhi, Pattanayak and Patra [38], Kanagasabai [34], Nabli, Djemaa and Amor [53], Md, Varadarajan and Mandal [57] and Alkalbani, Hussain and Kim [58].

6.1.2. Centralized Non-Semantic-Based Approaches

In this type of approach, a registry acts as a central node to keep all information for existing cloud services, and consumers can select the services based on their requirements. However, ontology is neither used in the process of discovery nor in representing the services. Instead, machine learning techniques are applied to implement the discovery process [57,58], and the services are represented using text, or description languages such as WSDL or RDF [33,62].

6.1.3. Decentralized Semantic-Based Approaches

In this type of approach, a direct connection between a cloud user and a provider is established. In this approach, the service description is published by the providers on their official website, which includes price and other information. Users can select the suitable service based on their requirements. In this type of approach, either a general search engine such as Google is used, or a private crawler is launched to find and search the offered cloud services through many providers' websites. Moreover, ontology is utilized as a useful knowledge representation tool of the underlying search concepts to enhance the crawlers' performance. Once the different cloud services are retrieved, it will be ranked based on the user criteria. This approach allows the cloud providers to update service attributes at any

time. Han and Sim [43], Chang, Juang, Chang and Yen [44], Alfazi, Sheng, Babar, Ruan and Qin [50], and Karthikeyan and RS [51] proposed models of this approach (see Table 2).

6.1.4. Decentralized Non-Semantic-Based Approaches

In this approach, a direct connection between a cloud user and a provider is established, so the user can search the services based on his requirements through a search engine that provides a description of services by the providers. In addition, this approach does not use ontology in the process of cloud services discovery. [4], Hamza, Aicha-Nabila, Okba and Youssef [47,48], Wheal and Yang [49], and Nabeeh, El-Ghareeb and Riad [46] are examples of this type.

6.2. Discovery Process Methods

In the literature, researchers applied different methods to implement cloud services discovery systems. These methods can be classified into two main approaches: agent-based and intelligent.

6.2.1. Agent-Based Approaches

To process the discovery and selection of cloud services, agents can be used to get input from the user. Then, the service is searched according to the requirements defined by the user. The agent gets the request and performs the matchmaking and reasoning the relations among the identified cloud services. Then, the agent may select the most relevant cloud service from a list of services that satisfies the user functional requirements. In the literature, many researchers applied this method to help in the discovery of cloud services [38,41,43,44,46,47,52,54,60].

6.2.2. Intelligent Approaches

In this approach of implementing a cloud service discovery, data mining and machine learning techniques are used. Classification methods are used to identify the appropriate cloud service toward users' needs. Then, cluster methods are applied to extract the features of the identified cloud service. This approach was applied in the literature by Kanagasabai [34], Nabli, Djemaa and Amor [53], Alfazi, Sheng, Babar, Ruan and Qin [50], Karthikeyan and RS [51], Md, Varadarajan and Mandal [57], Alkalbani, Hussain and Kim [58], and Jiang, Tao, Liu, Sun and Ling [56].

7. Discussion

Although there are three cloud service models (IaaS, PaaS, and SaaS), some studies only discover and identify whether the web page is related to cloud services or not without classifying it to the corresponding service model [34,35,38,46,47,49,54] (see Table 1). The classification of retrieved cloud services based on their service models (IaaS, PaaS, SaaS) can increase the efficiency and effectiveness of cloud services discovery. The consumer can identify the model of the cloud service, and the system will search only the services of that model. In addition, some studies proposed cloud service discovery systems to extract services of specific cloud models such as IaaS [5] or SaaS [36]. However, users may search for the integrated services of more than one model, which makes these systems insufficient. Furthermore, as users need to evaluate the services based on their requirements, SLA or QoS attributes should be considered. However, some studies provided a solution to discover cloud services without extracting their attributes [51–53,58], while others considered the extraction of only few attributes [4,44,48]. Three types of formats to represent cloud service attributes are used: tables, JSON (JavaScript Object Notation) and text. Representing the attributes of an extracted cloud service in a unified way may increase the effectiveness of cloud services discovery.

The complexity of user requirements (functional and non-functional) and the diversity of provided services' SLAs and QoS attributes as well as cloud computing characteristics motivated researchers to propose and develop different cloud service discovery models and solutions. Different methods and techniques were applied for service discovery such as ontology, which was used for different purposes

such as similarity checking and reasoning [43], and cloud service representation [5]. In addition, agents were applied to search the service enquired by the user. The agent gets the request, performs the matchmaking, reasoning the relations among the identified cloud services, and returns the results [41,46,47]. To improve the results, other researchers applied agents with ontology [39,59–61]. A search engine with a database of cloud services descriptions was developed. The search engine uses agents to search and check the similarities between the inquired service and the available cloud services. The similarity is checked using ontology. Moreover, machine learning techniques have emerged as a new trend to implement the cloud service discovery systems in recent years. Classification and clustering methods are used to identify the appropriate cloud services and extract their features [34,50,51,53,56–58]. These intelligent methods can be applied with ontology to enhance the clustering of the cloud services based on different features [34,53].

Additionally, reviewing the literature showed that cloud service discovery approaches may differ in the environment within which the services are searched. A central repository can be used to store all the services information provided by the cloud providers [45,59–61], or the services information can be searched for on the Web [44,46]. In addition, as these approaches can be implemented in ontology or non-ontology-based environments, the existing proposed models, systems, and frameworks are categorized based on discovery environment into four classes: centralized semantic based, centralized non-semantic based, decentralized semantic-based, and decentralized non-semantic-based.

Table 2 shows that the most common models were centralized semantic approaches that use ontology and agents during the cloud service discovery phases (searching, comparing, selecting and representing) [38,52,54,60]. Agents also were applied in the semantic decentralized environment [43,44]. Additionally, agent-based cloud discovery approaches were applied in non-semantic decentralized environments [41,46,47], but not in non-semantic centralized environments. On the other hand, methods based on machine learning were applied to implement the discovery process in centralized semantic [34, 53] and non-semantic [57,58] as well as decentralized semantic [50,51] and non-semantic [56].

8. Conclusions

In this study, the existing cloud services discovery solutions have been explored. By critically reviewing the related literature, the state of the art in the area of cloud services discovery was presented. Three different techniques were identified including ontology, agents, and machine learning. Different cloud models and QoS attributes of cloud services as well as SLAs were considered. Additionally, a taxonomy of cloud services discovery approaches was proposed. Four main classes of cloud services discovery approaches were identified based on the discovery environment, which are centralized semantic-based, centralized non-semantic-based, decentralized semantic-based, and decentralized non-semantic-based approaches. Moreover, cloud services discovery was classified based on the implementation method into two groups: agent based and intelligent solutions. Accordingly, this study presents a clear picture of the current cloud service discovery solutions and provides a conceptual structure to classify and compare cloud service discovery solutions.

Although there are many existing proposed approaches, there is still a need for more efficient systems that can consider all cloud models as well as QoS attributes or SLAs. The effectiveness of cloud service discovery can be improved by extracting all cloud service attributes that are described in different formats and represent them in a standardized way using a comprehensive ontology. Ontology can be used in the crawling procedure or in the matching process by identifying the similarity between retrieved and queried cloud services. Furthermore, cloud service discovery can use intelligent methods to extract cloud services represented in different formats. Furthermore, the cloud service discovery process can be carried out on all cloud services and classify them into their corresponding model (SaaS, PaaS, IaaS) using the classification method. Accordingly, for the future work, research applying machine learning techniques to classify retrieved cloud services into their models is suggested. Further, studies that propose ways to extract different types of formats representing cloud service attributes for

all cloud services are recommended. Finally, a way to standardize the format of representing cloud service attributes is suggested.

Author Contributions: Conceptualization, F.M. and A.M.A.; Data curation, F.M., A.M.A., F.A. and A.S.A.-M.A.-G.; Formal analysis, F.M.; Methodology, F.M., S.M.S. and F.E.E.; Project administration, S.M.S. and F.E.E.; Resources, F.A.; Supervision, S.M.S. and F.E.E.; Validation, A.S.A.-M.A.-G., S.M.S. and F.E.E.; Writing—original draft, A.M.A.; Writing—review & editing, F.M. Funding acquisition, A.S.A.-M.A.-G. All authors have read and agreed to the published version of the manuscript.

Funding: This project was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under grant No. (RG-18-611-38). The authors, therefore, gratefully acknowledge the DSR technical and financial support.

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

References

1. Soundararajan, V.; Govil, K. Challenges in building scalable virtualized datacenter management. *ACM SIGOPS Oper. Syst. Rev.* **2010**, *44*, 95–102. [[CrossRef](#)]
2. Emeakaroha, V.C. *Managing Cloud Service Provisioning and SLA Enforcement via Holistic Monitoring Techniques*; Technische Universität Wien: Karlsplatz, Wien, Austria, 2012.
3. Lee, G. Resource Allocation and Scheduling in Heterogeneous Cloud Environments. In *Computer Science*; University of California: Berkeley, CA, USA, 2012.
4. Gong, S.; Sim, K.M. CB-Cloudle and cloud crawlers. In Proceedings of the 2014 IEEE 5th International Conference on Software Engineering and Service Science, Beijing, China, 27–29 June 2014.
5. Nepal, S.; Zhang, M.; Ranjan, R.; Haller, A.; Georgakopoulos, D. An Ontology-based System for Cloud Infrastructure Services' Discovery. In Proceedings of the IEEE 8th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), Pittsburgh, PA, USA, 14–17 October 2012.
6. Garg, S.; Versteeg, S.; Buyya, R. A framework for ranking of cloud computing services. *Future Gener. Comput. Syst.* **2013**, *29*, 1012–1023. [[CrossRef](#)]
7. Ghazouani, S.; Slimani, Y. A survey on cloud service description. *J. Netw. Comput. Appl.* **2017**, *91*, 61–74. [[CrossRef](#)]
8. Alkalbani, A.M.; Hussain, F.K. A comparative study and future research directions in cloud service discovery. In Proceedings of the 2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA), Hefei, China, 5–7 June 2016; pp. 1049–1056.
9. Sun, L.; Dong, H.; Hussain, F.K.; Hussain, O.K.; Chang, E. Cloud service selection: State-of-the-art and future research directions. *J. Netw. Comput. Appl.* **2014**, *45*, 134–150. [[CrossRef](#)]
10. Ali, A.; Shamsuddin, S.M.; Eassa, F.E.; Mohammed, F. Cloud Service Discovery and Extraction: A Critical Review and Direction for Future Research. In *Proceedings of the Advances in Intelligent Systems and Computing*; Springer: Cham, Switzerland, 2018; pp. 291–301.
11. Bhardwaj, S.; Jain, L.; Jain, S. Cloud computing: A study of infrastructure as a service (IAAS). *Int. J. Eng. Inf. Technol.* **2010**, *2*, 60–63.
12. Mell, P.M.; Grance, T. *The NIST Definition of Cloud Computing*; NIST Special Publication: Gaithersburg, MD, USA, 2011.
13. Aoun, R.; Abosi, C.E.; Doumith, E.A.; Nejabati, R.; Gagnaire, M.; Simeonidou, D. Towards an optimized abstracted topology design in cloud environment. *Future Gener. Comput. Syst.* **2013**, *29*, 46–60. [[CrossRef](#)]
14. Calheiros, R.N.; Ranjan, R.; Beloglazov, A.; De Rose, C.A.F.; Buyya, R. CloudSim: A toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. *Softw. Pract. Exp.* **2010**, *41*, 23–50. [[CrossRef](#)]
15. Shimba, F. *Cloud Computing: Strategies for Cloud Computing Adoption*. In *Information Technology*; Dublin Institute of Technology: Dublin, Ireland, 2011.
16. Wei, Y.; Junpeng, C. Ontology management in cloud computing. In Proceedings of the 2010 Second International Conference on Computational Intelligence and Natural Computing, Wuhan, China, 13–14 September 2010.
17. Sakr, S.; Liu, A.; Batista, D.M.; AlOmari, M. A Survey of Large Scale Data Management Approaches in Cloud Environments. *IEEE Commun. Surv. Tutor.* **2011**, *13*, 311–336. [[CrossRef](#)]

18. Fox, A.; Griffith, R.; Katz, R.H.; Joseph, A.D.; Konwinski, A.; Lee, G.; Patterson, D.A.; Rabkin, A.; Stoica, I.; Zaharia, M.; et al. *Above the Clouds: A Berkeley View of Cloud Computing*; Department of Electrical Engineering and Computer Sciences, University of California: Berkeley, CA, USA, 2009; p. 28.
19. Reese, G. *Cloud Application Architectures: Building Applications and Infrastructure in the Cloud*; O'Reilly Media: Sebastopol, CA, USA, 2009.
20. Patidar, S.; Rane, D.; Jain, P. A Survey Paper on Cloud Computing. In Proceedings of the 2012 Second International Conference on Advanced Computing & Communication Technologies, Los Angeles, CA, USA, 27–29 June 2012.
21. Sun, L.; Dong, H.; Ashraf, J. Survey of Service Description Languages and Their Issues in Cloud Computing. In Proceedings of the 2012 Eighth International Conference on Semantics, Knowledge and Grids, Beijing, China, 22–24 October 2012.
22. Ahmed, E.; Khan, S.; Yaqoob, I.; Gani, A.; Salim, F. Multi-objective optimization model for seamless application execution in mobile cloud computing. In Proceedings of the 2013 5th International Conference on Information and Communication Technologies, Ohrid, Macedonia, 12–15 September 2013.
23. Ahmed, E.; Shiraz, M.; Gani, A. Spectrum-aware distributed channel assignment for cognitive radio wireless mesh networks. *Malays. J. Comput. Sci.* **2013**, *26*, 232–250.
24. Peiris, C.; Sharma, D.; Balachandran, B. C2TP: A service model for cloud. *Int. J. Cloud Comput.* **2011**, *1*, 3. [[CrossRef](#)]
25. Shamshirband, S.; Anuar, N.B.; Kiah, M.L.M.; Rohani, V.A.; Petković, D.; Misra, S.; Khan, A.N. Co-FAIS: Cooperative fuzzy artificial immune system for detecting intrusion in wireless sensor networks. *J. Netw. Comput. Appl.* **2014**, *42*, 102–117. [[CrossRef](#)]
26. Yassa, S.; Chelouah, R.; Kadima, H.; Granado, B. Multi-Objective Approach for Energy-Aware Workflow Scheduling in Cloud Computing Environments. *Sci. World J.* **2013**, *2013*, 1–13. [[CrossRef](#)] [[PubMed](#)]
27. Rahimi, M.R.; Venkatasubramanian, N.; Mehrotra, S.; Vasilakos, A.V.; Mehrotra, S. MAPCloud: Mobile Applications on an Elastic and Scalable 2-Tier Cloud Architecture. In Proceedings of the 2012 IEEE Fifth International Conference on Utility and Cloud Computing, Chicago, IL, USA, 5–8 November 2012.
28. Joshi, K.; Finin, T.; Yesha, Y. Automating Cloud Services Lifecycle Through Semantic Technologies. U.S. Patent US20160149769A1, 26 May 2016.
29. Zhao, L.; Ren, Y.; Li, M.; Sakurai, K. Flexible service selection with user-specific QoS support in service-oriented architecture. *J. Netw. Comput. Appl.* **2012**, *35*, 962–973. [[CrossRef](#)]
30. Chinnici, R.; Moreau, J.J.; Ryman, A.; Weerawarana, S. *Web Services Description Language (wsdl) Version 2.0 Part 1: Core Language*; W3C Recommendation: Cambridge, MA, USA, 2007; p. 26.
31. Cardoso, J.; Barros, A.; May, N.; Kylau, U. Towards a Unified Service Description Language for the Internet of Services: Requirements and First Developments. In Proceedings of the 2010 IEEE International Conference on Services Computing, Miami, FL, USA, 5–10 July 2010.
32. Snyder, H. Literature review as a research methodology: An overview and guidelines. *J. Bus. Res.* **2019**, *104*, 333–339. [[CrossRef](#)]
33. Chen, F.; Bai, X.; Liu, B. Efficient Service Discovery for Cloud Computing Environments. *Databases Inf. Syst.* **2011**, *153*, 443–448. [[CrossRef](#)]
34. Ngan, L.D.; Kanagasabai, R. OWL-S Based Semantic Cloud Service Broker. In Proceedings of the 2012 IEEE 19th International Conference on Web Services, Honolulu, HI, USA, 24–29 June 2012.
35. Tahamtan, A.; Beheshti, S.A.; Anjomshoaa, A.; Tjoa, A.M. A Cloud Repository and Discovery Framework Based on a Unified Business and Cloud Service Ontology. In Proceedings of the 2012 IEEE Eighth World Congress on Services, Honolulu, HI, USA, 24–28 June 2012.
36. Afify, Y.M.; Moawad, I.F.; Badr, N.L.; Tolba, M.F. A semantic-based Software-as-a-Service (SaaS) discovery and selection system. In Proceedings of the 2013 8th International Conference on Computer Engineering & Systems (ICCES), Cairo, Egypt, 26–28 November 2013.
37. Vasudevan, M.; Haleema, P.K.; Iyengar, N.C.S.N. Semantic Discovery of Cloud Service Catalog Published Over Resource Description Framework. *Int. J. Grid Distrib. Comput.* **2014**, *7*, 211–220. [[CrossRef](#)]
38. Parhi, M.; Pattanayak, B.K.; Patra, M.R. A Multi-agent-Based Framework for Cloud Service Description and Discovery Using Ontology. In *Proceedings of the Advances in Intelligent Systems and Computing*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2014; Volume 1, pp. 337–348.

39. Kang, J.; Sim, K.M. Cloudle: A Multi-criteria Cloud Service Search Engine. In Proceedings of the 2010 IEEE Asia-Pacific Services Computing Conference, Hangzhou, China, 6–10 December 2010.
40. Kang, J.; Sim, K.M. Cloudle: An Agent-based Cloud Search Engine that Consults a Cloud Ontology. In *Proceedings of the International Conference on Cloud Computing & Virtualization 2010 CCV 2010*; Global Science and Technology Forum: Singapore, 2010.
41. Sim, K.M. Agent-Based Cloud Computing. *IEEE Trans. Serv. Comput.* **2011**, *5*, 564–577. [[CrossRef](#)]
42. Han, T.; Sim, K.M. An ontology-enhanced cloud service discovery system. In Proceedings of the International MultiConference of Engineers and Computer Scientists, Hong Kong, China, 17–19 March 2010.
43. Han, T.; Sim, K.M. *An Agent-Based Cloud Service Discovery System that Consults a Cloud Ontology*; Springer: Cham, Switzerland, 2011; Volume 70, pp. 203–216.
44. Chang, Y.-S.; Juang, T.-Y.; Chang, C.-H.; Yen, J.-S. Integrating intelligent agent and ontology for services discovery on cloud environment. In Proceedings of the 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Seoul, Korea, 14–17 October 2012.
45. Noor, T.H.; Sheng, Q.Z.; Alfazi, A.; Ngu, A.H.; Law, J. CSCE: A Crawler Engine for Cloud Services Discovery on the World Wide Web. In Proceedings of the 2013 IEEE 20th International Conference on Web Services, Santa Clara, CA, USA, 27 June–2 July 2013.
46. Nabeeh, A.N.; El-Ghareeb, H.A.; Riad, A. Integrating software agents and web services in service oriented architecture based cloud services discovery framework. *J. Conver. Inf. Technol* **2015**, *10*, 67.
47. Hamza, S.; Aicha-Nabila, B.; Okba, K.; Youssef, A.; Hamza, S.; Youssef, A. A Cloud computing approach based on mobile agents for Web services discovery. In Proceedings of the Second International Conference on the Innovative Computing Technology (INTECH 2012), Rabat, Morocco, 11–13 September 2012.
48. Gong, S.; Sim, K.M. CB-Cloudle: A Centroid-based Cloud Service Search Engine. In Proceedings of the International MultiConference of Engineers and Computer Scientists, Hong Kong, China, 12–14 March 2014; Newswood Limited: Hong Kong, China.
49. Wheal, J.; Yang, Y. CSRecommender: A Cloud Service Searching and Recommendation System. *J. Comput. Commun.* **2015**, *3*, 65–73. [[CrossRef](#)]
50. Alfazi, A.; Sheng, Q.Z.; Babar, A.; Ruan, W.; Qin, Y. Toward Unified Cloud Service Discovery for Enhanced Service Identification. In *Service Research and Innovation*; Springer: Cham, Switzerland, 2015; pp. 149–163.
51. Balaji, S.; Kumar, R.S.R.; Karthikeyan, N.K. Fuzzy service conceptual ontology system for cloud service recommendation. *Comput. Electr. Eng.* **2018**, *69*, 435–446. [[CrossRef](#)]
52. Boukadi, K.; Rekik, M.; Rekik, M.; Ben Abdallah, H. FC4CD: A new SOA-based Focused Crawler for Cloud service Discovery. *Computing* **2018**, *100*, 1081–1107. [[CrossRef](#)]
53. Nabli, H.; Ben Djemaa, R.; Ben Amor, I.A. Efficient cloud service discovery approach based on LDA topic modeling. *J. Syst. Softw.* **2018**, *146*, 233–248. [[CrossRef](#)]
54. Parhi, M.; Pattanayak, B.K.; Patra, M.R. A multi-agent-based framework for cloud service discovery and selection using ontology. *Serv. Oriented Comput. Appl.* **2017**, *12*, 137–154. [[CrossRef](#)]
55. Modi, K.; Garg, S. A QoS-based approach for cloud-service matchmaking, selection and composition using the Semantic Web. *J. Syst. Inf. Technol.* **2019**, *21*, 63–89. [[CrossRef](#)]
56. Jiang, Y.; Tao, D.; Liu, Y.; Sun, J.; Ling, H. Cloud service recommendation based on unstructured textual information. *Future Gener. Comput. Syst.* **2019**, *97*, 387–396. [[CrossRef](#)]
57. Quadir, A.; Varadarajan, V.; Mandal, K. Efficient Algorithm for Identification and Cache Based Discovery of Cloud Services. *Mob. Netw. Appl.* **2019**, *24*, 1181–1197. [[CrossRef](#)]
58. Alkalbani, A.M.; Hussain, W.; Kim, J.Y. A Centralised Cloud Services Repository (CCSR) Framework for Optimal Cloud Service Advertisement Discovery From Heterogenous Web Portals. *IEEE Access* **2019**, *7*, 128213–128223. [[CrossRef](#)]
59. Kang, J.; Sim, K.M. Towards Agents and Ontology for Cloud Service Discovery. In Proceedings of the 2011 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, Beijing, China, 10–12 October 2011.
60. Kang, J.; Sim, K.M. A Cloud Portal with a Cloud Service Search Engine. In *International Conference on Information and Intelligent Computing IPCSIT*; IACSIT Press: Singapore, 2011.

61. Kang, J.; Sim, K.M. Ontology and search engine for cloud computing system. In Proceedings of the 2011 International Conference on System Science and Engineering, Dalian, China, 6 August 2011.
62. Mittal, S.; Joshi, K.P.; Pearce, C.; Joshi, A. Automatic Extraction of Metrics from SLAs for Cloud Service Management. In Proceedings of the 2016 IEEE International Conference on Cloud Engineering (IC2E), Luxembourg, 12–15 December 2016.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).