Bulletin of the *Transilvania* University of Brakov ÅVol. 9 (58) No. 2 - Special Issue ï 2016
Series I: Engineering Sciences

APPLICATIONS OF THE SEARCH AS A SERVICE (SaaS)

P. DAŠIĆ[1] J. DAŠIĆ[2] B. CRVENKOVIĆ[3]

Abstract: *Search as a Service (SaaS) is a cloud service model the main focus of which is on enterprise search or site-specific web search. Modern companies require fast and accurate information from their internal databases, internal document stores, or through the content of a website. Having a reliable searching mechanism is essential for both internal company staff and for external customers. In this paper the overview of current state and technological advances of Search as a Service (SaaS) cloud service is given, as well as its security issues on current internet service platforms.*

Key words: *Cloud computing, Search as a Service (SaaS), internet technologies.*

1. Introduction

The vast amount of data available in electronic form wouldnôt be of much use if they couldnôt be searched for specific information. It is an essential function for any business database function, either through internal databases, internal document stores, or through the content of a Website. Search as a Service (SaaS) is a cloud based service which is a branch of Software as a Service (SaaS), with a purpose of performing enterprise search or site-specific web search. Search as a Service (SaaS) is a sophisticated method of retrieving specific information with great complexity behind it. Existing literature offers very little insight on Search as a Service (SaaS) functioning; however, there are numerous publications on cloud related technologies development. Some of scientific contributions include: multi-domain search [1], semantic search [2,3], digital ecosystems search [4], Everything as a Service (XaaS) [5], Digital forensics as a service (DFaaS) [6], personalized search [7], similarity-based search [8], heuristic-based search algorithm [9], many-objective visual analytics (MOVA) [10].

2. Search as a Service (SaaS)

Search as a Service (SaaS) hosts search engine capable of full-text, numerical, and faceted searching that seamlessly delivers results in real-time even from the first entered character. Full-text search examines all words within full-text fields in order to find the most relevant records. This capability allows quickly searching and returning records from large volumes

[1] High Technical Mechanical School of Professional Studies Trstenik, Serbia. E-mail: dasicp58@gmail.com
[2] SaTCIP Publisher Ltd., Vrnja¸ka Banja, Serbia. E-mail: dasic_jovan@yahoo.com
[3] SaTCIP Publisher Ltd., Vrnja¸ka Banja, Serbia. E-mail: crvenkovicbojan@gmail.com

of data. Security is ensured through access control by utilizing encryption algorithms on communication channels [11-13], permission-based index clustering [14], multi-user indexing [15], privacy-aware searching [16] and secure data sharing over untrusted cloud providerôs methods [17]. Numerical search is used for retrieving results from large datasets which consist mainly from numbers. Faceted search, also called faceted navigation or faceted browsing, is a technique for accessing information organized according to a faceted classification system, allowing users to explore a collection of information by applying multiple filters [18].

In Table 1 are shown features of Search as a Service (SaaS).

Table 1.

Relevance and ranking	SaaS offers customizable relevance and ranking of search results enabling the client to fine tune their product or service visibility.
Typo tolerance	In todayôs mobile environment it often happens that the user misspells a word or term which without an algorithm that automatically detects such errors will not yield a relevant result.
Smart highlighting	Search results have a highlighted search term or word which facilitates the users decision on selecting the right response.
Facets	Even with the first character the service offers search based facets which improves navigation, drill down, and refinements based on the userôs query.
Geo awareness	Service reads the location, and/or language used, of the user and ranks the results which are closer to him. For example, it can be highly useful for tourists seeking a coffee shop in Pairs.
Language support	Service has in-build support for all languages without any intervention from the client side.
Security	Prevention of crawling client data, account hacking, human mistakes, confidential data breach etc. Two-factor Authentication, Secure your Admin API Key, HTTPS, Unretrievable Attributes and API Key security.
Analytics	This feature gives more insights into how the search engine is used and factors like: most popular searches, average hits without typos and count, queries that returned no results, activity by countries, most popular filters.

Life cycle begins with importing data from the clientôs database into the service providerôs database. There are usually two methods of importing data, either by dashboard or by using API. Whatever the case is it is recommended to upload the data in batches of 1000 to 10000 records at a time. Usual file formats are JSON, CSV, or a TSV. The data are then sent to the cloud (all servers of the service provider). The data that is stored inside the cloud can be updated and reindexed as necessary. From that point on the databases are always in sync through a serviceôs API for any type of operation. Second step is the configuration of the index for precise ranking and relevance of the search results. The configuration is executed by the client in the Search as a Service (SaaS) control center. Most Search as a Service (SaaS) providers handle semi-structured data well; however, formatting the data properly can have significant improvements towards speed, reliability and accuracy [20]. Last step of the process is the implementation of search interface to any application that is required to have search capabilities. The described steps are graphically shown in Fig. 1.

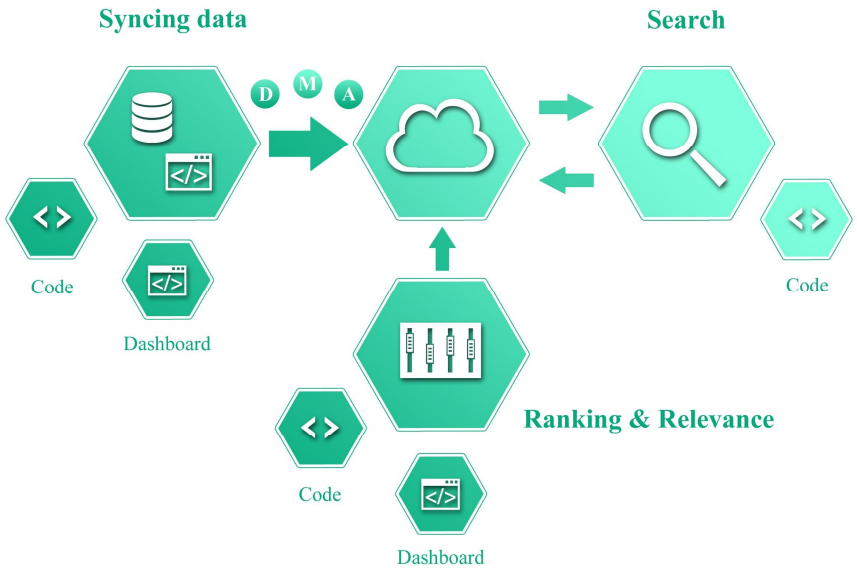

Fig. 1. *Life cycle of SaaS [18-19]*

3. Application examples of Search as a Service (SaaS)

Search as a Service (SaaS) performs enterprise search and site-specific web search [21] as its main purpose. Enterprise search is a set of operations for consolidating and processing data from multiple enterprise-type sources into formatted indices. Sources for data collections include databases, content management systems, intranets, e-mails and file systems. Operations include content awareness, content processing and analysis, indexing, query processing and matching. Site-specific web search is a process of obtaining search results from an individual web siteôs database and textual content utilizing either a local search engine or an external service through API as a product (APIaaP) or a connector.

Some of the derived search services are: Enterprise Search as a Service (ESaaS), Elastic Search as a Service (ESaaS) concept and Grid Search as a Service (GSaaS) [18].

Applications of Search as a Service (SaaS) cover a wide area of industries. Here we will shortly describe only the most used ones. However its worth mentioning that the application also covers areas of manufacturing, business & professional services, telecommunications, non-profit organizations, energy & utilities, professional associations, etc. On Fig. 2 is given a graphical representation of Search as a Service (SaaS) applications. Following is the brief description of applications:

- **High Tech**

IT companies are the largest users of Search as a Service (SaaS) especially large software developers, like Microsoft, Adobe, AutoDesk etc., which produce enormous amounts of data like manuals, articles, client lists, usage statistics etc. Without advanced data searching capabilities their existence would be hardly possible.

- **Financial Services**

Banking systems maintain a large database of private and enterprise clients, each with its own set of account attributes like account number, balance, debt, credit,

savings, investments, customer support etc. Security advantages offered by cloud hosted indices offer unprecedented capabilities compared to any locally hosted servers.

- **Mobile markets**

For over 15 years now smart phones have introduced a new way of having a mobile personal computer at hand. Today smartphones and tablets even match the computer power of a low cost PC. Application of Search as a Service (SaaS) is widely used in mobile devices especially for software markets (GooglePlay, AppStore), networked business apps, geo-services [22] etc. The main advantage is utilizing the cloud processing power saving the device resources.

Fig. 2. Search as a service (SaaS) applications

- **Public Sector**

Government bodies such as ministries, military, municipalities, legal branches, various public services etc., handle vast amounts of data on local and country level. For example, imagine the size of a consolidated database of Chinaôs consensus. Computer indexing alone would take months of work with standard techniques, while querying and results would be ambiguous. Advanced features of SaaS are a clear winner in that case.

- **Language support**

Social networks have grown to over half of the worldôs population with conversations reaching billions of words

every day. A great feature of cloud supported searching is the ability to search and receive results in any language with the ability to instantly translate the search result. This is a clear advantage of connecting even more people and sharing knowledge.

- **Geographical systems**

One of the largest and most important applications of Search as a Service (SaaS) is for geo-services like navigation, mapping applications (Google Maps and Bing Maps), manipulating, analyzing, and managing, of all types of spatial or geographical data. This set of procedures and actions is commonly called the Geographical Information Systems (GIS).

- **Digital media**

Internet has brought a new dimension to the news creation and publishing where not only professionals but also common people contribute to the worldôs media scene. Millions of articles, videos and tweets are published every day. Finding relevant information has never been as complicated but at the same time made easy with usage of cloud hosted searchable indexes.

- **Healthcare**

Healthcare system needs to manage a large number of patient records, medicine, equipment, employees, ground and online support etc. Speed and precision of search results is of crucial importance which can be successfully supported by cloud search resources.

- **Retail**

E-commerce has experienced rapid growth since it inception. Large internet stores (Amazon, E-bay, etc.) offer thousands of products and users demand fast, accurate and detailed descriptions. Some of the e-stores have even developed their own Search as a Service (SaaS), for example Amazon CloudSearch, while other extensively use this service from other providers.

- **Various analytics**

We live in the age of Big Data which is generated by billions of computer devices. In order to perform any kind of analytical and logical reasoning process it is necessary to obtain relevant data from large data sets. This is where advanced cloud search algorithms can deliver both speed and precision of results.

In Table 2 an alphabetic list of Search as a Service (SaaS) providers is given. It is a relatively new cloud service and currently there are only a dozen of providers. Service process range from 50 to 4000 USD per month depending on number of indices, number of operations and as well hardware setup a client requires. Search as a Service (SaaS) can be used by small to large clients as nature of cloud environment allows scalability at any time.

Table 2.

SaaS provider	Service description and web site
Algolia	Algolia is a fully hosted search service, available as a REST API (Representational state transfer API). API clients are also available for all major frameworks, platforms and languages. Data transmission between clients and the API is in JSON format. (URL: https://www.algolia.com/)
Amazon CloudSearch	Amazon CloudSearch is a managed service in the AWS Cloud that makes it simple and cost-effective to set up, manage, and scale a search solution for your website or application. Amazon CloudSearch supports 34 languages and popular search features such as highlighting, autocomplete, and geospatial search. (URL: https://aws.amazon.com/cloudsearch/)

Coveo Search	Coveo's advanced enterprise search technology adds the value of rich content and insights to CRM, customer service applications, intranets and websites. Coveo securely connects with legacy and cloud systems, and provides unified search, dynamic 360-degree views of information, and contextual, proactive recommendations of relevant content and experts using powerful analytics. (URL: http://www.coveo.com/)
Hubstor	Fully-managed search solution that makes storage Relevant Products/Services, indexing, and discovery low cost and completely turn-key for enterprises. Delivers searchable cloud archiving at petabyte scale. (URL: http://www.hubstor.net/search-as-a-service/)
Bonsai	Bonsai is a highly scalable search engine, fully managed by the experts in hosted search and offers advanced search services. (URL: https://bonsai.io/)
Elasticsearch	Elasticsearch is a distributed RESTful search engine built for the cloud. It offers distributed, scalable, and highly available, real-time search and analytics capabilities, sophisticated RESTful API. (URL: https://www.elastic.co/products/elasticsearch)
IndexDen	IndexDen hosted, super-fast, full-text search engine tuned specifically for searching and storing textual data. Powerful full-text search provides excellent API service. It is scalable, schema free, and easy to setup. (URL: http://indexden.com/)
Microsoft Azure	Cloud search service for web and mobile app development. Reliable throughput and storage provide fast search indexing and querying to support time-sensitive search scenarios. (URL: https://azure.microsoft.com/en-us/services/search/)
SearchBlox	SearchBlox is a leading provider of enterprise search, sentiment analysis and text analytics solutions. SearchBlox was created to provide customers with simple, affordable solutions for their data management needs including web-based administration and integrated data connectors to index enterprise and web content. (URL: http://www.searchblox.com/)
Searchify	Searchify is a Full-text Search-as-a-Service. It allows easy adding of custom full-text search, without the cost or complexity of managing search servers. It offers location-aware geo search, faceted search, custom scoring & sorting functions, snippets & highlighting for professional-looking results. (URL: http://www.searchify.com/)
Swifttype	Swiftype is built on an advanced search algorithm to deliver superior and relevant results. With language modeling intelligence such as bigram matching, spellcheck and stemming, you can expect that users will be able to find the content theyôre looking for. (URL: https://swiftype.com/)

4. Conclusion

Search as a Service (SaaS) is a newly established cloud service which due to new technologies has had positive spiraling results. Although it offers great benefits in the search area, it is a surprising fact that there is only less than a dozen companies that provides this service. Search as a Service (SaaS) is rapidly being implemented in major Web site databases such as Vevo, IBM, Amazon, Netflix etc. This paper provides the contribution of better understanding the underlying principles and methods of service functioning with clear advantages that it has to offer to modern data searching and retrieval. We also listed the applications of Search as a Service (SaaS) in the modern internet environment which showed that it influences almost every aspect of human interaction.

References

1. Bozzon, A., Brambilla, M., Corcoglioniti, F., Vadacca, S.: *A service-based architecture for multi-domain search on the Web.* In: Lect. Notes Comput. Sci. (LNCS), **6470**, (2010), p. 663ï 669.
2. Çelik, D. Elçi, A.: *A semantic search agent approach: Finding appropriate semantic web services based on user request term(S).* In: Proceedings of the 3rd International Conference on Information and Communications Technology (ICICT-2005) - Enabling Technologies for the New Knowledge Society, 2005, **2005**, p. 675ï 688.
3. Shu, G., Xu, X., Hua, H.: *A Web Service search approach based on semantic and search engine.* In: Proceedings of the 6th International Conference on Pervasive Computing and Applications (ICPCA-2011), p. 484ï 489.
4. Dong, H., Hussain, F.K., Chang, E.: *A service search engine for the industrial digital ecosystems.* In: IEEE Trans. Ind. Electron., **58** (2011) No.6, p. 2183ï 2196.
5. Duan, Y., Fu, G., Zhou, N., Sun, X., Narendra, N.C., Hu, B.: *Everything as a Service (XaaS) on the Cloud: Origins, Current and Future Trends.* In Proceedings of the 8th IEEE International Conference on Cloud Computing (CLOUD-2015), p. 621ï 628.
6. Lee, J. Un, S.: *Digital forensics as a service: A case study of forensic indexed search.* In: Proceedings of the International Conference on ICT Convergence, 2012, p. 499ï 503.
7. Hu, R., Dou, W., Liu, J.: *A personalized search approach for web service recommendation.* In: Int. J. Ad Hoc Ubiquitous Comput., **13** (2013) No. 2, pp. 83ï 95.
8. Ngu, A.H.H., Ma, J., Sheng, Q.Z., Yao, L., Julian, S.: *ServiceXplorer: A similarity-based Web service search engine.* In: Proceedings of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2014, p. 1251ï 1252.
9. Rodriguez-Mier, P., Mucientes, M., Lama, M.: *Automatic web service composition with a heuristic-based search algorithm.* In: Proceedings of the 9th IEEE International Conference on Web Services (ICWS-2011), p. 81ï 88.
10. Woodruff, M., Simpson, T., Hadka, D., Reed, P.: *Many-objective visual analytics: In search of search-as-a-service.* In: Proceedings of the AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, 2014.
11. Bouabana-Tebibel, T., Kaci, A., *Parallel search over encrypted data under attribute based encryption on*

the Cloud Computing. In: Comput. Secur., **54** (2015), p. 77ï 91.
12. Chen, F., Hong, J., Zheng, R., Jiang, M.: *Research on access control during search encrypted data.* In: Lect. Notes Electr. Eng., **155** (2012), pp. 677ï 682.
13. Kaci, A., Bouabana-Tebibel, T., Challal, Z.: *Access Control Aware Search on the Cloud Computing.* In: Proceedings of the 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI-2014), 2014, p. 1258ï 1264.
14. Micheli, E.C., Margaritis, G., Anastasiadis, S.V.: *Permission-based index clustering for secure multi-user search.* In: Trans. Data Priv., 8 (2015), No. 1, p. 29ï 53.
15. Micheli, E.C., Margaritis, G., Anastasiadis, S.V.: *Efficient multi-user indexing for secure keyword search.* In: Proceedings of the CEUR Workshop, 2014, **1133**, p. 390ï 395.
16. Pervez, Z., Awan, A.A., Khattak, A.M., Lee, S., Huh, E.-N.: *Privacy-aware searching with oblivious term matching for cloud storage.* In: J. Supercomput., **63** (2013) No. 2, p. 538ï 560.
17. Zhao, G., Rong, C., Li, J., Zhang, F., Tang, Y.: *Trusted Data Sharing over Untrusted Cloud Storage Providers*. In: 2010 IEEE Second International Conference on Cloud Computing Technology and Science, 2010, p. 97ï 103.
18. Daġİ, P., Daġİ, J., Crvenkoviĺ, B., *Service Models for Cloud Computing: Search as a Service (SaaS).* In: International Journal of Engineering and Technology, **8** (2016).
19. Algolia technical documentation. 2016. Available on [https://www.algolia.com/doc].
20. Hadjilambrou, Z., Kleanthous, M., Sazeides, Y.: *Characterization and analysis of a web search benchmark.* In: Proceedings of the IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS), 2015, p. 328ï 337.
21. Atsaros, G., Spinellis, D., Louridas, P.: *Site-Specific versus General Purpose Web Search Engines: A Comparative Evaluation.* In: Proceedings of the Panhellenic Conference on Informatics (PCIô08), 2008, p. 44-48.
22. Both, A., Ngomo, A.-C.N., Usbeck, R., Lukovnikov, D., Lemke, C., Speicher, M..: *A Service-oriented Search framework for full text, geospatial and semantic search.* In: Proceedings of the ACM International Conference Proceeding Series, 2014, p. 65ï 72.