International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 07 | July 2021 www.irjet.net p-ISSN: 2395-0072

A Study on State-of-the-art Cloud-based Enterprise Search Solutions

Karthik S Rao[1], Anitha Sandeep[1]

[1]*Department of Computer Science and Engineering, R.V College of Engineering, Karnataka, India*

Abstract - *Search has evolved to be an essential feature of any enterprise application. Having an effective enterprise search solution can fundamentally improve customer satisfaction in the case of external applications like e-commerce sites while also enhancing employee productivity in the case of internal company portals. Thus, it becomes critical that enterprises build state-of-the-art search solutions. This paper details the fundamentals of enterprise search and highlights the challenges of implementing them from scratch in an on-premises environment. The benefits of using a managed search service (or a search-as-a-service offering) for developing sophisticated enterprise search solutions are enumerated. The fact that search-as-a-service is primarily cloud-powered is underscored and the pros and cons that stem from this fact are studied. Finally, a few industry-leading cloud-based search offerings that facilitate the process of building modern cloud-first enterprise search solutions are listed.*

Key Words: **Enterprise search, Managed search service, Search-as-a-service, Cloud-search provider, Security in search, ML in search.**

1.INTRODUCTION

Search has overwhelmingly redefined and restructured how we find and use information. In a time where nearly 43% of visitors to an organization's website instantly look for and interact with a search box [1], a meaningful and accessible search experience directly results in better customer engagement and fulfilment. Searching is an imperative part of any business information retrieval system. The search can be performed on internal document stores, internal databases, or through the content of a website. This is a critical feature for both internal employees and for external partners and customers.

Even though a simple DBMS query such as "List existing employees whose salaries are greater than XYZ" is a trivial search scenario which can be easily developed in-house (probably through SQL), it happens to be a naïve example. More complex searches such as "Find all the company blog articles that contain a reference to a product XYZ" or "Search the customer-provided reviews for any synonyms of 'manufacturing defect' or 'waste of money' are more problematic to implement. The complexity is intensified with users increasing their demand for cognitive-like AI-enabled search experiences that deeply understand the intent behind their queries. It would be a herculean task for organizations to build a sophisticated search engine for each of their disparate business scenarios from scratch. Fortunately, managed search services come to the rescue.

Managed search services are also commonly termed as search-as-a-service. In this set-up, a search provider company offers a search service, and a contract is struck with the search tenant to support the latter's search scenarios. Then the tenant uses the provider's API for instance to up-load content data or indexing metadata for the content to be searched. It is the provider's responsibility to construct a search index for this content. If the content happens to be free text or similar unstructured data, it can be tokenized by Lucene or a similar process.

This concept is hardly new since web-mediated search has existed for a long time. Web-mediated search is limited to a company's publicly visible websites. As such, an existing search engine (Google, Bing, etc.) is customized to web crawl only the company's site(s). When a search query is submitted, the customized search engine is triggered. This works like any other web search but returns results only from the company's particular web site(s). The web search engine is customized easily by listing the company's web site(s) using the provider's search console. Despite the simplicity of web-mediated search, it could not be used as a comprehensive enterprise search solution since it could not incorporate the private resources that are not visible to the public web. This problem as well as other concerns around security and privacy are addressed by the state-of-the-art managed search service providers.

Organizations which desired comprehensive enterprise search experiences had to build on-premises custom solutions. This task was often outsourced to enterprise search providers who would allocate resources (time, money, and people) to implement enterprise search for the client in the latter's private datacenters. After the advent of enterprise-scale cloud providers and the gradual buildup of trust in such providers, the general apprehension over security and privacy concerns of storing sensitive enterprise information on the cloud has eased. A study in 2017 [2] roughly three-quarters of the I.T. professionals surveyed reporting that they are using some or the kind of cloud managed service. Some of the popular cloud services include Web hosting (75%), productivity solutions (55%), email hosting (54%) and content filtering (45%). Despite the initial fears around security, recent studies have shown that cloud service workloads are suffering at least 60% fewer security incidents than those in traditional data centers [3]. With

the increased maturity and adoption of the cloud, organizations feel safer to leverage cloud-based search solutions.

It is safe to say that web search services are cloud-first experiences. Web search providers like Google and Microsoft are first-party customers of their own cloud offerings ensuring that benefits of the cloud are automatically applied to the web search space. On similar lines, search-as-a-service models developed by search service providers cannot be truly envisioned nor realized without the cloud as their backbone. Thus, search-as-a-service is often implicitly recognized as cloud-powered-search. Not only does the cloud provide the needed abstraction through well- defined interfaces and reduce the overall efforts and timeframes, but also ensures the high availability, scalability, and reliability of such services.

Search-as-a-service is designed to be leveraged for building rich enterprise search experiences. In other words, it is used to enable search across enterprise's data and scenarios. This search experience can be enabled for both internal and external stakeholders. This differentiates search-as-a-service from web search services. The migration journey of enterprise search from on-premises to the cloud involves choosing a trusted search-as-a-service provider as a first step.

Although both web search and enterprise search have effectively moved to the cloud, the implications of the latter from an organization's standpoint are far more significant. As such this paper shall limit its discussion to enterprise search systems. This paper shall first discuss the basic and changing landscapes of enterprise search. It shall then deep-dive into how search-as-a-service models can help build complex enterprise search solution in a quick and timely fashion. Finally, it highlights a few industry-leading cloud-search offerings that are based on the search-as-a-service model.

2. LITERATURE REVIEW

There is ample amount of literature surrounding the world of enterprise search. Martin White in their book "Enterprise Search" [4] details out a step-by-step guidance on planning and implementation of a managed search system that meets the requirements of any organization's business as well needs of its employees – it describes 12 critical factor that help gauge an organization's search needs, illustrates how to make a business case for search, helps set quality guidelines for organizational content and metadata, provides a comprehensive list of open source and commercial search technology and at the same time instructs how to choose vendors based on individual requirements and not merely going by the trend. A.Z. Broder et al. in their work titled "Towards the next generation of enterprise search technology" [5] clarifies the differences between search systems for the Web and those for enterprises and deep dives in building modern enterprise search systems based on the Unstructured Information Management Architecture (UIMA) that it proposes. Rajat Mukherjee et al. in "Enterprise Search: Tough Stuff: Why is it that searching an intranet is so much harder than searching the Web?" [6] reiterate why web search models cannot be adapted as is to the enterprise search scenarios and provides suggestions on how to improve spidering and indexing, data filtering and search relevance in the context of enterprise search. Pavel A. Dmitriev et al. in their work on "Using annotations in enterprise search" [7] describe how user annotations can be utilized to improve the quality of intranet search and propose two ways to obtain user annotations (implicit and explicit feedback) and illustrate how they can be integrated into a search engine. David Hawking in their paper titled "Challenges in Enterprise Search" [8] characterizes enterprise search, highlights its economic magnitude, and states some of the unresolved research questions in the field of enterprise search. Tony Russell-Rose et al. scrutinize the needs and behaviors of individuals across a broad range of discovery and search scenarios within different types of enterprises and present a taxonomy of "discovery modes" for enterprise search in their paper "A Taxonomy of Enterprise Search" [9]. Luanne Fruend et al. in their study "Enterprise search behavior of software engineers" [10] investigate how a group of software engineers use a workplace search system and identify patterns of search behavior following it up with design recommendations for search systems that will better adapt to the requirements of software engineers. Gianluca Demartini in their research on "Leveraging semantic technologies for enterprise search" [11] present a plan that aims to use Information Retrieval, Semantic Web, and User Modelling techniques to cope with the challenges that arise in enterprise search scenarios. PengFei Li et al. report the findings of an experiment that simulates heterogeneous enterprise retrieval and compares different merging algorithms, in their paper on "Merging algorithms for enterprise search" [12].

So far, the works that have been described lay the foundation for understanding enterprise search. This section shall now focus on literature around search-as-a-service and how it can be leveraged to build meaningful and secure cloud-powered enterprise search solutions. Predrag Dašić et al. in their paper "Service Models for Cloud Computing: Search as a Service (SaaS)" [13] provide an overview of the present state and technological advances of search-as-a-service as well as highlighting key security issues on current internet service platforms. "Many-objective visual analytics: In search of search-as-a-service" [14], a study by Matthew Woodruff et al. evaluates five leading multi-objective evolutionary algorithms (MOEAs) in the context of many-objective visual analytics (MOVA) for optimizing search-as-a-service, based on their effectiveness, efficiency, reliability, and controllability on four different formulations of a benchmark theoretical design problem. Based on their findings, the Borg MOEA is called out as a leading algorithm for search-as-a-service. Jooyoung Lee et al. present how to build a unique search-

as-a-service offering customized for the digital forensic scenarios in their work titled "Digital forensics as a service: A case study of forensic indexed search" [15]. Work carried out in "The EBI search engine: EBI search as a service—making biological data accessible for all" [16] by Young M. Park et al. propose the EBI search-as-a-service offering that provides out-of-the-box search capabilities that can be integrated into third-party portal via RESTful APIs to retrieve search results from several biological data sources (such as PDBe, ChEBI, ChEMBL, UniProtKb, RNAcentral, etc.) relating to transcriptomics, genomics, metabolomics, proteomics, and so on. "Integrating Azure Search" [17] by Thurupathan Vijayakumar, describes the Azure search-as-a-service model and how developers can focus solely on the search functionality and let go of the management of search resources by integrating Azure search in their applications.

The following literature focus on securing search-as-a-service, as that remains the biggest obstacle in organization's decision to implement enterprise search on the cloud. Aameek Singh et al. in their study "Search-as-a-service: Outsources search over outsourced storage" [18] present a novel filesystem-based search-as-a-service that integrates indexing/search mechanisms and access control into a unified framework to support access control aware search. Their approach performs indexing within the secure enterprise domain and uses an access control barrel (ACB) primitive to encapsulate access control within these indices following which these indices are systematically encrypted and shipped to the search provider for hosting. All these activities are managed and performed by the search provider ensuring that the client enterprise expends only minimal effort to integrate enterprise search in their applications and at the same time not having to fully trust the search provider for data confidentiality. "A Secure Cloud Data Sharing Protocol for Enterprise Supporting Hierarchical Keyword Search" [19] by Hongbo Li eat al. describes the process of securing enterprise data stored on the cloud and still providing a rich search experience by using public key encryption with keyword search (PEKS) which enables keyword search without decryption in public key encryptions settings. Additionally, to there is need to provision a hierarchy-based monitoring where a higher layer employee should have privilege over a lower layer employee's data to check if there are any violations without letting employees be aware of the same. The authors describe a variant of PEKS named Hierarchical Public Key Encryption with Keyword Search (HPEKS) which supports this monitoring function without authorization from the sender. SM Zobaed et al. in their work "SAED: Edge-Based Intelligence for Privacy-Preserving Enterprise Search on the Cloud" [20] help organizations overcome the privacy challenges of using a fully cloud-based enterprise search services by leveraging the power of the edge to separate the intelligent aspect of search from its pattern matching aspect. They propose Smartness at Edge (SAED) that provides intelligence in the form of personalized and semantic search at the edge while preserving privacy of the search on the cloud tier. SAED is available as an out-out-of-the-box component and can be integrated into existing enterprise search systems without demanding any change on the latter. "Privacy-Preserving Clustering of Unstructured Big Data for Cloud-Based Enterprise Search Solutions" [21] by SM Zobaed et al., describes the hindrances on data processing caused by client-side encryption solutions that have been employed by many enterprises to protect privacy while utilizing cloud-based enterprise search services. They identified that clustering encrypted data (which is critical while dealing with different forms of big data) was the major bottleneck during the data processing. To preserve data privacy, they propose privacy-preserving clustering schemes which function based on statistical characteristics of the data to determine both the appropriate number of clusters and relevant content for each cluster.

There is also literature on how to build these search-as-service models and package them as a meaningful offering to attract enterprises in an increasingly growing cloud-first market. For instance, Bukhary Ikhwan Ismail et al. in their work titled "A Proposed Framework for Search as a Service on Private Cloud" [22] discuss the challenges and requirements to host a search engine as a service. Their work is particularly useful for search providers who are looking to provision a managed search service on their own private cloud instead of delivering it to popular public clouds.

3. ENTERPRISE SEARCH SYSTEMS

Enterprise search systems are essential software used to search for information inside a corporate organization. Such systems identify and enable indexing, searching and display of specific content to authorized users across the enterprise. The searchable data can be present anywhere inside the company and it can be in any format – in databases, document management systems, e-mail servers, address books, on paper and so on. Some key components of an Enterprise search system are illustrated in Figure 1.

Enterprise search is super critical for work productivity, be it for internal or external users of the organization's information. Several studies highlight the high costs of information search time in organizations which have not implemented a robust enterprise search system. A survey by SearchYourCloud [23] discovered that "It takes workers up to eight searches to find the right document and information, according to 80 percent of respondents". Further, a study by McKinsey [24] discovered that "The average interaction worker spends an estimated 28 percent of the workweek managing e-mail and nearly 20 percent looking for internal information or tracking down colleagues who can help with specific tasks.". That means an average employee spends nearly 1.8 hours every day for searching information.

The generic working of enterprise search systems closely resembles the following phases (Refer Figure 2):

1. Content submission: This involves the task of ensuring that the information to be searched is available to the search system. There are 2 basic models – push and pull. In a push model, source or the content repository pushes new content to the search system, for instance through an API (programmed to do so periodically, or on manual prompts). The other approach is pull based, where real-time indexing is crucial and the search system itself pulls and gathers the content from various sources using specialized connectors.

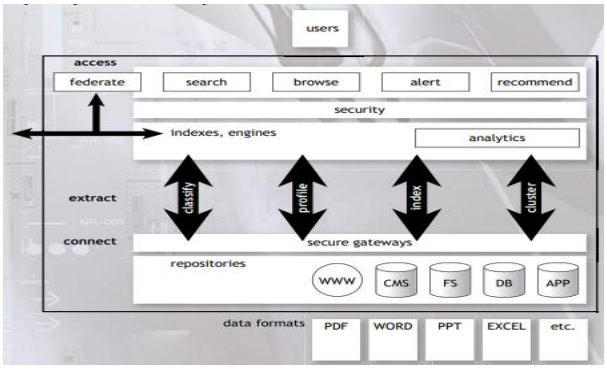

Fig-1. Key Components of an Enterprise Search System

2. Pre-processing: The content retrieved from different sources might be in disparate formats or document types. This phase processes incoming documents are converted to plain text using document filters so that they can be search efficiently. This phase also includes tokenization.

3. Indexing: After processing the contents, they are stored in an index. The index is a structure that enables rapid real-time responses to search queries. Without an index, the search system would scan every document available, which would require considerable time and computing power.

4. Query processing: Using a specific API or any other interface, queries are submitted to the search system. The system then uses the phrase to search the index and return an appropriate result set.

5. Ranking: Finally, the search system determines the sequence in which search results should be presented to the user. Ranking is done on a variety of parameters like relative position of query terms in documents (whether they are present in the title, in the file name, etc.), age of the document or other metadata, etc.

Security regarding access to documents in the result set may be handled in two ways by the search system:

1. Early binding which ensures that authorization and access levels are checked at the indexing phase. This may lead to faster query times but are less up to date (less accurate) if access levels were changed after the indexing phase.

2. Late binding ensures authorization and access levels are checked at query-processing phase, meaning that the result-set is first computed after which user's permissions are validated on each document. Despite the slowness evident in this method, it maintains up to date security.

The above-described system is considered quite "naïve" in present-day conversations revolving search. Despite engineering such a system, enterprises may still identify severe ineffectiveness in delivering high value and meaningful search experiences to a broad set of potential stakeholders. Although the "naïve" system helps users find available content, it is not enough to improve user satisfaction as there is no guarantee that the users find what they might "actually" require – considering that fact that their needs change in different contexts over a period. For instance, two different users might be looking for vastly different things when searching for say "watch battery". One might be looking for information on how long a battery will last while the other might be interested in purchasing a new one. The "naïve" system would return the same result set to both users – the result set is ranked as being most relevant to the query "watch battery" rather than to the context and intent of the individual user who executed that query.

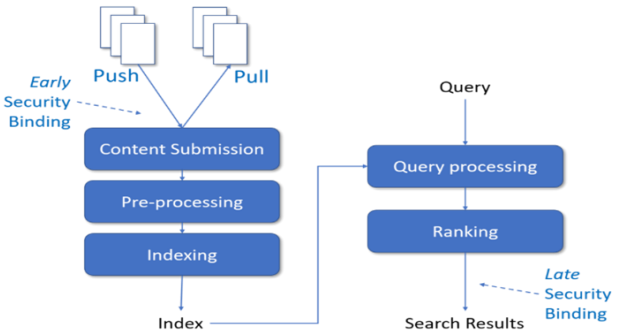

Fig-2. Generic working of an enterprise search system

Individual relevance and personalized results offer the true value-add potential for search, and that becomes a guarantee only when the "naïve" enterprise search is transformed into intelligent search. This involves adding machine learning (powered by usage analytics) to the various stages of the "naïve" enterprise search system. For instance, providing summarization, concept extraction, terminology extraction, collocation extraction and hyperparameter optimization to the pre-processing phase, or adding word sense disambiguation, statistical spelling and vocabulary correction, automatic language detection to the query processing phase, or incorporating question answering, semantic search and query suggestions to the ranking phase. Of course, all of these can be developed without machine learning and usage analytics, but it would demand that administrators do so such tasks manually by either adding synonyms, or creating boosting rules, promoting documents, and so on. The fact that

relevance is a constantly changing landscape – documents that was most relevant last week might be extremely irrelevant today - it is challenging for administrators to keep pace with the rate of change. Machine learning enables transformation of highly manual and complex enterprise search into intelligent, self-learning and self-tuning search where all the afore mentioned enhancement techniques are automated in a fashion that is geared towards improving search relevance and hence driving personalized high-value user experience.

For the smallest of organizations, even implementing a "naïve" enterprise search system on-premises from scratch can be a herculean task involving high up-front investment and greater development effort. This may ultimately lead to poor operational efficiency and high customer churn rates. This is where the power of managed search services or search-as-a-service really comes to light. As mentioned, earlier search-as-a-service offerings are cloud-first and most of the complex infrastructure, platform and software capabilities are hosted and provided for by a cloud-search vendor. Such vendors not only provide storage and compute mechanisms for building large-scale search systems but also expose several interfaces (such as APIs) for common industry-specific search scenarios and features that can be leveraged out-of-the-box. Moreover, most cloud-search providers are adopting intelligent search by design which means machine learning is baked into their offerings.

4. SEARCH-AS-A-SERVICE

Search-as-a-service is a cloud-first service model that is designed and built by various cloud-search providers to considerably ease any organization's effort in handling the latter's enterprise search scenarios. The organization using such a service is considered a tenant or client and is said to have established a service contract with the provider. Cloud-search providers deliver the search specific infrastructure, platform and even software interfaces to the tenant enabling the latter to rapidly build a complex enterprise search solution without having to worry about the technicalities behind the leveraged features.

The general life cycle of a typical provider-tenant interaction begins with importing data from the tenant's data sources to the service provider's datastores. This is useful only if the contents are static failing which it would require periodic data refresh to maintain consistency (data can also be updated directly at the provider's end). To handle complex dynamic scenarios, many search providers even allow to establish always-on links to the tenant's data sources to ensure that the provider sees near real-time data. The provider ensures efficient indexing of both the stored and of data pulled in at real-time. The next step involves configuring that index for precise ranking and relevance of the search results. This configuration can be managed by the tenant (they can opt for the defaults offered by the provider as well) in a search-as-a-service control center that is built by the provider to enable a rich visibility-and-control experience to the tenants. Finally, the tenant proceeds to implement a search interface in their application that connects to various search capabilities offered by the provided through easy-to-use APIs. A typical search-as-a-service offering can be outlined as illustrated in Figure 3.

Some of the features of a modern search-as-a-service offering include relevance and ranking, result categorization, spell check and correction, filtering, pagination, location-awareness, multi-lingual support, security, analytics, natural language processing, personalized-relevance, search-as-you-type, reverse image search, search within videos and images, etc. Search-as-a-service enhances the meaning of "search" to beyond merely returning a set of results based off a query. It showcases how visualization, mapping, filtering, anomaly detection, forecasting can all be integrated into a comprehensive state-of-the-art enterprise search solution.

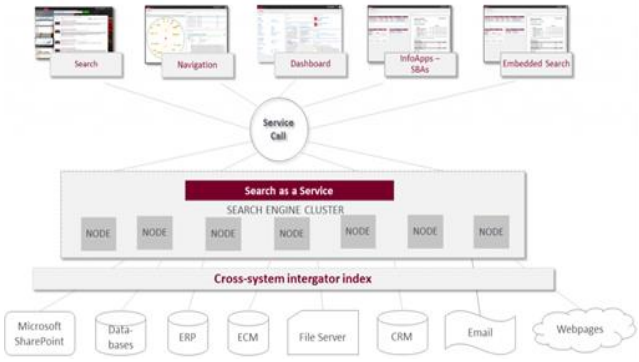

Fig-3. An outline of search-as-a-service

Data security has traditionally been a constant roadblock for organizations looking to leverage search-as-a-service or any cloud offering in general. However, over time reputed cloud providers have built up trust by investing and innovating in the security and privacy space. As of today, trusted cloud providers offer more secure or at least a comparable security to on-premises datacenters. This is primarily driven by the fact that providers are incorporating security and privacy by design adopting an always encrypted approach for data be it at rest or in-transit, certificate-based authentication for tenant applications communicating with the data, configurable role-based access controls and policies, strong isolation of data and applications of different tenants, state-of-the-art network security practices at each of their datacenters, continuous backup and quick restores.

Search-as-a-service ultimately improves the layer of abstraction between the cloud-search provider and the enterprise tenants, thereby enabling enterprises to focus more on business values and outcomes and let the providers take ownership of the underlying complex

technicalities. Although on first sight it seems that organizations must give up complete control over the architecture, most cloud-search providers are become more flexible in their offerings by adopting the hybrid mindset. They do so by granting a finer grain of control to organizations such that the latter can pick and choose only the components they want to reuse and implement the rest in their desired proprietary manner. The tenants have complete visibility and control over the consumed resources – out-of-the-box services, hosted applications, and data. This is done through monitoring and management solutions baked into a rich control center experience built by most industry-leading cloud-search providers.

5. POPULAR CLOUD-SEARCH PROVIDERS

Cloud-search providers implement the search-as-service model to deliver pre-designed but configurable high quality search related services to organizations. Organizations can find such providers on the cloud marketplaces. Cloud marketplaces are available in most public cloud platforms such as Microsoft Azure, Amazon Web Services and Google Cloud Platform. The search providers can be direct first-party in a sense they are offering services on their own cloud platform (and hence infrastructure, such as Azure Cognitive Search) or third-party in a sense that they are providing services on a non-native cloud platform built on top of the latter's platform (such as Elastic on Azure). Most cloud-search providers also offer additional consultative tech support, training, and consulting services. Trust in cloud-search providers is protected through contracts between the tenants and such providers are backed by service-level-agreements (SLAs) that transform availability, reliability, scalability, and security guarantees made by the provider into legal terms. The following sections describe some popular industry-leading cloud-search providers and the capabilities of their search-as-a-service offerings. The generic workflow of utilizing a cloud search service is illustrated in Figure 4.

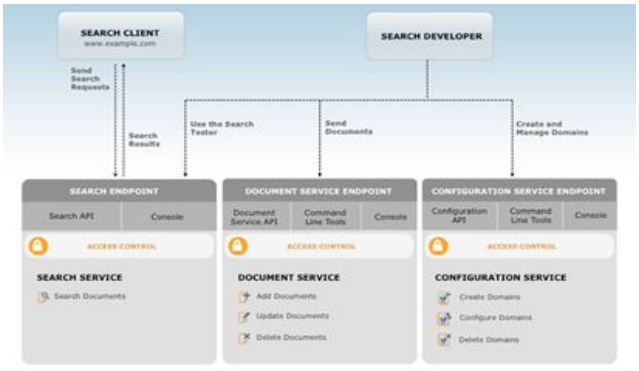

Fig-4. Generic workflow of a cloud search service

One of the popular cloud search service is Azure Cognitive Search Service. It empowers developers with an architecture, tools, and APIs that enable building a rich search experience over enterprise-wide private and heterogeneous content across web, desktop, and mobile environments. Common enterprise scenarios covered by this service includes e-commerce site search, catalog or document search, or knowledge mining for data science. The following capabilities are available to be consumed upon creating an instance of this service: a full text search enabled search engine, persistent storage of private content inside a search index, rich indexing that allows text analysis and optional A.I. functionalities for content extraction and transformation, variety of query capabilities such as simple syntax, full Lucene syntax and typeahead search, client libraries and REST APIs for programmability support across .NET, Python, Java and JavaScript, in-built Azure integration at the machine learning and data layer along with connectivity to other applied AI-powered cognitive services. To add on to these, it also allows easy implementation of search-related features such as relevance tuning, faceted navigation, filters (including geo-spatial search), autocomplete, and synonym mapping. This search service can seamlessly integrate with other Azure resources and services utilizing the notion of indexers whose job is to automate data ingestion/retrieval from data sources, and skillsets that incorporate consumable A.I. from Azure Cognitive Services (such as image and text analysis), or any custom A.I. that is created in Azure Machine Learning or wrapped inside Azure Functions. The search service runs two major workloads: indexing and querying. The mechanisms that drive their operations are irrelevant to a client as they can easily consume the functionality that is exposed through a simple REST API or a .NET SDK that abstracts the inherent complexity of information retrieval.

Similarly, Amazon CloudSearch is offered as a managed search service in the AWS Cloud. It promises the same ease of provisioning, setup, and maintenance as other cloud search providers. AWS Management console is an easy-to-use tool that allows one to create a search domain and upload data that is to be made searchable, and Amazon CloudSearch would do the rest automatically – such as provisioning the required resources and deploying a highly customized and tuned search index. Search parameters can be easily changed to meet the requirements of the clients, search relevance can be fine-tuned, and new settings can be applied at any time with zero disruption or downtime to the service. Most of the features provided by Azure Cognitive Search Service are available in Amazon CloudSearch.

Google Cloud Search is an offering within the Google Cloud Platform, which allows clients to add managed search capabilities across applications Google Workplaces such as Drive, Gmail, Sites, Calendar, Docs, Contacts and more, thus enabling a rich integration of Google Workplaces with applications built by enterprises who use Google Workplaces as their mainstay collaboration platform. In addition to surfacing work-related documents, it also acts

as a directory service allowing users to seek out colleagues' contact information and view what events and files they have in common. Though it started off as a Google Workplaces solution only, it has expanded today to be a stand-alone offering catering to clients who may not have Google Workplaces.

Elastic has been a popular brand in the enterprise search space ever since its inception. Elastic had traditionally been a services provider, but it did not provide management, hosting or infrastructure capabilities. It offered a rich search stack which had to implemented by its clients on the latter's private datacenters. Apart from the initial set-up, it also required any additional management to be handled at the client's end itself. This meant that enterprises had to undergo extra training and hire people who had sufficient expertise in developing for, deploying, and managing Elastic on-premises. Elastic quickly realized these shortcomings and rode the wave of the cloud to launch Elastic Cloud by partnering with Amazon, Microsoft, and Google. As a result, Elastic now offers its cloud managed service to enterprises who are using either the Amazon Web Services, Microsoft Azure or the Google Cloud Platform. In other words, managed Elastic is available as a marketplace offering in these clouds. Behind the scenes, the Elastic engineering team spent considerable effort into ensuring that Elastic can be deployed on these clouds, and it continues to provide improvements as well as SLA-backed support to customers on each of these three clouds. This offering is considerably different than those described before as this is a third-party service available for consumption through marketplaces whereas the prior ones were first-party services provided by the cloud owners themselves.

There are a variety of other cloud search offerings that are available as a first party managed service by itself, or as a third-party marketplace offering on the three popular clouds – Amazon Web Services, Microsoft Azure, and Google Cloud Platform, or both. Algolia for instance, provides its hosted and managed search service on its own cloud (dedicated to its primary business – search) as well as versions of the same search service ready-made for deployment on popular clouds. Other known cloud search providers include Coveo Search, HubStor, Bonsai, IndexDen, SearchBlox, Searchify, and SwiftType, each of which offer their own unique capabilities that solve equally unique enterprise search scenarios. For example, Coveo Search is extremely useful to add search capabilities to Customer Relationship Management (CRM) applications. Thereby, it is critical to evaluate these cloud search services in the context of the enterprise search solution once desires to build rather than on its generic popularity or abundance of features.

6. CONCLUSIONS

Information is the currency of today's digital first economy. But information is often hidden in seemingly endless document(s) and in disparate heterogeneous data sources. Searching these extensive and varied documents and data sources (databases, file servers, emails, etc.) remains a pain-staking and time-consuming activity to an average user. This can negatively impact any enterprise in the long run as difficulties experienced during the search process can lead to loss of a potential customer in the case of external applications as well as leave employees disgruntled and hampering their productivity in the context of internal applications. Good enterprise search solutions are inherently complex, and organizations often tend to compromise or overlook their value when it comes to conversations around potential areas of investments. There is a growing need for a cost-effective, time and effort saving solution to the problem of enterprise search. Search-as-a-service offerings (or managed search services) come to the rescue and do the heavy-lifting of providing a single-point solution for handling most of the complexities of search in an out-of-the-box fashion. Moreover, search-as-a-service is designed for the cloud and allows providers to offer infrastructure, hosting, maintenance, and support apart from the actual search capabilities. Increasingly, these search capabilities have machine learning at their core, and if not at least are being enhanced by machine learning. All this ensures that organizations always have the state-of-the-art enterprise search solutions with no initial or future effort on their part. The cloud takes care of the scalability, availability and reliability and the enterprises effectively need to pay only for what they use – which means it is no longer a hurdle to convince investors on providing better enterprise search solutions. There are various cloud search providers in the market either offering first-party services or third-party versions of the same on other clouds.

New discoveries in the machine learning space such as ongoing work on GPT-3 can completely revolutionize the meaning of search – search would not be limited to merely finding existing content in the linked data sources based on a query string but rather act as a question-answer agent by generating complete and concise answers on the fly that attempt to answer the question inferred from a user search query, and doing so by paying attention to context, intent and identity of the individual user. The future of search is as complex as it is fascinating, and hence it is imperative that organizations move to cloud search offerings and let the search providers handle the technicalities of adapting to innovations that are yet to come.

REFERENCES

[1] S. Mulpuru, C. Johnson, L. Naparstek, "Must-Have eCommerce Features", The Retail E-Commerce Playbook, Forrestor Reports, Oct 2020, Accessed on: May 20th, 2021. [Online]. Available doi: https://www.forrester.com/report/MustHave+eCommerce+Features/-/E-RES89561

[2] "Trend Report: Why Businesses are Moving to the Cloud", Insight Blogs, An Insight and Microsoft Office Collaboration, Feb 2017. Accessed on: May 15th 2021, [Online]. Available: https://www.uk.insight.com/content/dam/insight/EMEA/blog/2017/05/Trend%20Report%20-%20Why%20Businesses%20are%20Moving%20to%20the%20Cloud.pdf

[3] D.M. Smith, "Cloud Strategy Leadership", Mar 2017. Accessed on: May 15th 2021, [Online]. Avaialable: https://www.gartner.com/imagesrv/books/cloud/cloud_strategy_leadership.pdf

[4] M. White, " Enterprise Search: Enhancing Business Performance". 2nd Edition, O'Reilly Media, 2015. ISBN: 9781491915493. [Online]. Available: https://www.oreilly.com/library/view/enterprise-search-2nd/9781491915523/

[5] A. Z. Broder and A. C. Ciccolo, "Towards the next generation of enterprise search technology," in IBM Systems Journal, vol. 43, no. 3, pp. 451-454, 2004, doi: https://doi.org/10.1147/sj.433.0451

[6] R. Mukherjee and J. Mao, "Enterprise Search: Tough Stuff: Why is it that searching an intranet is so much harder than searching the Web?", Association for Computing Machinery, vol. 2, no. 2, pp. 36-46, 2004, doi: https://doi.org/10.1145/988392.988406

[7] P.A. Dmitriev, N. Eiron, M. Fontoura, and E. Shekita, "Using annotations in enterprise search", in proceedings of the 15th international conference on World Wide Web (WWW '06), Association for Computing Machinery, pp. 811–817, 2006, doi: https://doi.org/10.1145/1135777.1135900

[8] D. Hawking, "Challenges in Enterprise Search", Proceedings of the 15th Australasian Database Conference, Australian Computer Society, Inc., pp. 15 – 24, doi: https://dl.acm.org/doi/10.5555/1012294.1012297

[9] T. Russell-Rose, J. Lamantia, M. Burrell, "A Taxonomy of Enterprise Search", In EuroHCIR, pp. 15-18. 2011, doi: http://ceur-ws.org/Vol-763/paper4.pdf

[10] L. Freund, E.G. Toms, "Enterprise search behaviour of software engineers", in proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval, Association for Computing Machinery, pp. 645–646, 2006, doi: https://doi.org/10.1145/1148170.1148297

[11] G. Demartini, "Leveraging semantic technologies for enterprise search", in proceedings of the ACM first Ph.D. workshop in CIKM , Association for Computing Machinery, pp. 25–32, 2007, doi: https://doi.org/10.1145/1316874.1316879

[12] P. Li, P. Thomas, D. Hawking, "Merging algorithms for enterprise search", in proceedings of the 18th Australasian Document Computing Symposium, Association for Computing Machinery, pp. 42–49. 2013, doi: https://doi.org/10.1145/2537734.2537750

[13] P. Dašić, J. Dašić, B. Crvenković, "Service Models for Cloud Computing: Search as a Service (SaaS)", International Journal of Engineering and Technology,vol. 8, pp. 2366-2373, 2016, doi: https://doi.org/10.21817/ijet/2016/v8i5/160805034

[14] M.J Woodruff, T. Simpson, P. Reed, "Many Objective Visual Analytics: In Search of Search-as-a-Service", AIAA AVIATION 2014 -15th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, 2014, doi: https://doi.org/10.2514/6.2014-2729

[15] J. Lee and S. Un, "Digital forensics as a service: A case study of forensic indexed search," 2012 International Conference on ICT Convergence (ICTC), 2012, pp. 499-503, doi: https://doi.org/10.1109/ICTC.2012.6387185

[16] Y.M Park, S. Squizzato, N. Buso, T. Gur, R. Lopez, "The EBI search engine: EBI search as a service—making biological data accessible for all.", Nucleic acids research, vol. 45, no. W1, pp. W545-W549, 2017, doi: https://doi.org/10.1093/nar/gkx359

[17] T. Vijayakumar, "Integrating Azure Search. In: Practical Azure Application Development", Apress, pp. 157-174, 2017, doi: https://doi.org/10.1007/978-1-4842-2817-3_7

[18] A. Singh, M. Srivatsa, L. Liu. "Search-as-a-service: Outsourced search over outsourced storage." ACM Transactions on the Web (TWEB), vol. 3, no. 4, pp. 1-33, 2009, doi: https://doi.org/10.1145/1594173.1594175

[19] H. Li, Q. Huang and W. Susilo, "A Secure Cloud Data Sharing Protocol for Enterprise Supporting Hierarchical Keyword Search," in IEEE Transactions on Dependable and Secure Computing, doi: https://doi.org/10.1109/TDSC.2020.3027611

[20] S.M. Zobaed, M.A. Salehi, R. Buyya, "SAED: Edge-Based Intelligence for Privacy-Preserving Enterprise Search on the Cloud." arXiv preprint arXiv:2102.13367, 2021, doi: https://arxiv.org/abs/2102.13367

[21] S.M. Zobaed, R. Buyya , M.A. Salehi, "Privacy-Preserving Clustering of Unstructured Big Data for Cloud-Based Enterprise Search Solutions", arXiv:2005.11317, 2020, Available:https://arxiv.org/2005.11317

[22] B. I. Ismail et al., "A Proposed Framework for Search as a Service on Private Cloud," 2018 IEEE Conference on Open Systems (ICOS), 2018, pp. 7-12, doi: https://doi.org/10.1109/ICOS.2018.8632814

[23] P. Bernstein, "SearchYourCloud Survey, It Takes up to 8 Attempts to Find an Accurate Search Result", Nov. 5, 2013. [Online]. Accessed on May 13th 2021,Available: https://www.techzone360.com/topics/techzone/articles/2013/11/05/359192-searchyourcloud-survey-it-takes-up-8-attempts-find.htm

[23] [24] M. Chui, J. Manyika, J. Bughin, R. Dobbs, C. Roxburgh, H. Sarrazin, G. Sands, M. Westergren, "The social economy: Unlocking value and productivity through social technologies", Jul 1st 2012. [Online]. Accessed on May 14th, 2021. Available: https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-social-economy#