



Universidad Nacional de Costa Rica

Sede Regional Brunca

Campus Pérez Zeledón

Curso: Administración de Base de Datos

Profesor: Master Carlos Carranza Blanco

II Ciclo 2025

Fecha de entrega:

23 de septiembre del 2025

Estudiantes:

Leon Marin Alejandro

Fonseca Hidalgo Kendall

Elasticsearch

I. Introducción	3
II. Desarrollo	4
A. Antecedentes	4
B. Ventajas y desventajas	6
a. Ventajas	6
b. Desventajas	8
C. Futuro de la tecnología	10
D. Análisis del Tema	14
III. Conclusiones y Recomendaciones	17
A. Conclusiones	17
B. Recomendaciones	18
Bibliografía	21

I. Introducción

En el panorama actual de la informática, el crecimiento exponencial de los datos generados por aplicaciones, sistemas empresariales y dispositivos conectados ha impulsado la necesidad de herramientas capaces de procesar, buscar y analizar información en tiempo real. Dentro del contexto del curso Administración de Base de Datos de la Sede Regional Brunca, Campus Pérez Zeledón, se nos ha encomendado la investigación de Elasticsearch, una de las tecnologías más relevantes para la administración y explotación de datos, que ofrece un motor de búsqueda distribuido, escalable y de alto rendimiento, diseñado para trabajar con grandes volúmenes de información de manera eficiente.

Este documento tiene como objetivo profundizar en el estudio de Elasticsearch, abordando sus antecedentes, características técnicas y funcionales, así como sus ventajas y desventajas frente a otras soluciones de búsqueda y análisis de datos. Asimismo, se analizarán sus posibles aplicaciones futuras en un entorno cada vez más orientado hacia la Big Data, la observabilidad y la integración con inteligencia artificial, aspectos que marcan la evolución de la gestión y el análisis de la información en la actualidad.

La investigación se estructura siguiendo los lineamientos del curso, presentando una visión integral que combina los aspectos históricos, técnicos y prácticos de la herramienta. Se incluirán ejemplos de uso reales que ilustren la implementación de Elasticsearch en entornos con grandes volúmenes de datos, con el fin de demostrar sus capacidades y potencial de aplicación. Además, se reflexionará sobre su impacto en la optimización de procesos de búsqueda y análisis, así como en la toma de decisiones estratégicas basadas en datos, destacando su papel como un recurso clave tanto en el ámbito profesional como en el académico.

II. Desarrollo

A. Antecedentes

Elasticsearch es un motor de búsqueda y análisis de código abierto desarrollado sobre la biblioteca Apache Lucene, ampliamente reconocido por su capacidad para manejar grandes volúmenes de información de forma rápida y eficiente. De acuerdo con IBM (2021), esta tecnología permite a los desarrolladores incorporar funciones de búsqueda extremadamente escalables en sus aplicaciones, adaptándose a diferentes necesidades empresariales y técnicas. Su arquitectura distribuida fue diseñada para realizar búsquedas de texto completo en documentos JSON, eliminando la necesidad de esquemas estrictos y ofreciendo compatibilidad con múltiples tipos de bases de datos. Una de sus principales fortalezas radica en la posibilidad de recibir datos no estructurados provenientes de diversas fuentes, procesarlos y almacenarlos en formatos especializados que optimizan la búsqueda basada en lenguaje natural.

El proyecto vio la luz en el año 2010 gracias a la empresa Elastic, con el objetivo inicial de convertirse en una solución flexible y distribuida que permitiera la indexación y recuperación eficiente de información textual. Con el tiempo, su popularidad se incrementó de forma exponencial, hasta convertirse en una pieza central del ecosistema conocido como ELK Stack. Este crecimiento se consolidó en 2013, cuando Elastic decidió integrar su motor de búsqueda con Logstash, una herramienta de ingesta y procesamiento de datos, y Kibana, una plataforma de visualización interactiva. Según Banon S. y Kren P. (2025), esta integración transformó radicalmente el papel de Elasticsearch, que pasó de ser una simple herramienta de indexación de texto a un sistema integral capaz de ofrecer almacenamiento, búsqueda y análisis de grandes volúmenes de datos en tiempo real.

La creación del ELK Stack permitió que esta tecnología fuera ampliamente adoptada en áreas como el análisis de registros de aplicaciones (application logging), la supervisión de métricas operativas, la monitorización de infraestructura y el análisis de seguridad. Este enfoque integral resultó especialmente útil para organizaciones que requerían un sistema centralizado capaz de recibir datos en tiempo real desde múltiples fuentes, procesarlos de forma simultánea y generar visualizaciones y reportes útiles para la toma de decisiones estratégicas.

En cuanto a su evolución técnica, Elasticsearch ha atravesado una serie de hitos importantes que han marcado su madurez como plataforma:

- 2014 – Versión estable inicial: Se incorporaron mejoras clave como el sistema de snapshots y recuperación, agregaciones avanzadas, circuit breakers para evitar sobrecargas y la API _cat, que facilitó la administración del clúster.
- 2015 – Seguridad y nube: Se introdujo Shield, un plugin de seguridad que inicialmente era de pago, así como la adquisición de found.no, lo que permitió lanzar el servicio administrado Elastic Cloud. Además, se añadieron pipelines y se fortaleció la seguridad mediante Java Security Manager.
- 2016 – Salto a la versión 5: Elastic decidió unificar sus productos bajo el nombre Elastic Stack, integrando Beats como agentes ligeros para la recolección de datos, nodos de ingesta y el lenguaje de scripting Painless, optimizado para operaciones internas.
- 2017 – Versión 6: Se eliminaron los “tipos” de documentos para simplificar el modelado de datos y se optimizó el ordenamiento de índices.
- 2019 – Versión 7: Se introdujo el nuevo coordinador de clúster Zen2, más resiliente y escalable, y se ofreció seguridad gratuita a partir de las versiones 6.8 y 7.1, marcando un cambio importante en su política de acceso a funciones críticas.
- 2020 – Optimización de series temporales: Se implementaron ILM (Index Lifecycle Management), data tiers y searchable snapshots, lo que permitió gestionar datos históricos de manera más rentable y eficiente.
- 2024 – Retorno al código abierto: Tras haber abandonado el modelo abierto en 2014, Elastic reincorporó la licencia AGPL, devolviendo a la comunidad la posibilidad de usar y modificar el software sin restricciones propietarias. En este mismo año se simplificó el despliegue local de Elasticsearch y Kibana, y se lanzó LogsDB, una modalidad de indexación que reduce el uso de almacenamiento en logs hasta en un 65 %.

A lo largo de su historia, Elasticsearch ha demostrado una notable capacidad de adaptación a las demandas del mercado tecnológico. Desde sus inicios centrados en la búsqueda de texto, ha evolucionado hacia un sistema robusto para el análisis de datos, la monitorización y la seguridad, incorporando funciones antes exclusivas de soluciones propietarias. La apertura de la seguridad de pago en 2019, junto con la reducción de costos de almacenamiento y la inclusión de nuevas capacidades como los vectores de búsqueda en la versión 8, refuerzan su posicionamiento como una herramienta versátil y preparada para escenarios de alto rendimiento.

En la actualidad, Elasticsearch es utilizado por empresas de todos los tamaños y sectores para resolver problemas críticos de búsqueda y análisis de datos. Su flexibilidad, escalabilidad y constante innovación lo han convertido en un estándar de facto en soluciones de observabilidad, análisis de logs y motores de búsqueda internos. Su evolución, tanto tecnológica como estratégica, evidencia un compromiso continuo con su visión original: proporcionar una plataforma de búsqueda potente, eficiente y accesible, capaz de adaptarse a las necesidades cambiantes del entorno digital global.

B. Ventajas y desventajas

a. Ventajas

Como cualquier tecnología en la actualidad, Elasticsearch presenta ventajas y desventajas. En esta sección se abordarán sus principales beneficios, los cuales han contribuido a que este motor de búsqueda sea ampliamente adoptado por la industria. Una de las más relevantes es la escalabilidad, como lo indica AltexSoft (2023): “Elasticsearch is built for scalability, enabling businesses to grow without worrying about data management constraints”. Esto significa que la plataforma fue diseñada para crecer de manera flexible, permitiendo a las empresas expandirse sin preocuparse por limitaciones en la gestión de datos. De forma complementaria, IBM (2021) señala que Elasticsearch también ofrece escalabilidad horizontal mediante la replicación de clusters, indicando que “esto le brinda a su organización la capacidad de utilizar todas las características de Elasticsearch mientras reduce las latencias para los usuarios y garantiza una alta disponibilidad de los servicios”. En otras palabras, la replicación de clusters no sólo mejora la disponibilidad y la velocidad de respuesta, sino que también permite a las organizaciones aprovechar al máximo las capacidades de la herramienta.

Otra ventaja significativa es su amplia compatibilidad con múltiples lenguajes de programación. Para los desarrolladores, resulta fundamental trabajar con tecnologías que puedan integrarse con diversos entornos sin requerir herramientas adicionales por falta de compatibilidad. En este sentido, AltexSoft (2023) afirma que “Elasticsearch provides multiple client libraries in various programming languages, including Java, JavaScript, PHP, C#, Ruby, and Python”. Esto quiere decir que existen bibliotecas oficiales para estos lenguajes, facilitando la integración con aplicaciones ya desarrolladas o en desarrollo. Incluso IBM (2021) agrega que también cuenta con soporte para entornos como .NET y otros lenguajes adicionales, lo que amplía aún más su alcance.

Una tercera ventaja clave es su comunidad activa y en constante crecimiento, lo que se traduce en una amplia disponibilidad de documentación, foros y soluciones para problemas comunes en entornos de desarrollo. AltexSoft (2023) destaca que “The rapid response time, with questions often answered in approximately an hour, testifies to the high level of the community's engagement”. Esto implica que, en promedio, una consulta en los canales oficiales recibe respuesta en aproximadamente una hora, lo cual agiliza la resolución de problemas. La comunidad cuenta con un foro de discusión con más de 129,375 usuarios (dato de 2023), un canal de Slack para intercambiar información de forma más directa, alrededor de 58,061 preguntas en Stack Overflow, y un repositorio oficial en GitHub con cerca de 65 mil estrellas y 1,800 contribuidores activos. Además, Elasticsearch mantiene un programa de contribución que premia y reconoce a los miembros más activos, fomentando así la colaboración y el crecimiento del ecosistema.

Continuaremos con las más importantes como lo es la flexibilidad avanzada de sus consultas y la analítica en tiempo real. Según AltexSoft (2023), “Elasticsearch offers a highly flexible Query API that integrates both search and real-time analytics”. Esto significa que su API permite combinar en una sola operación funciones de búsqueda y análisis de datos sin necesidad de procesos separados, lo que agiliza el trabajo y reduce la complejidad técnica. Esta API está basada en RESTful y utiliza JSON, lo que facilita la indexación de datos y la ejecución de consultas complejas de forma sencilla. Entre las funcionalidades que pueden realizarse en una sola consulta se incluyen filtrado, ordenamiento, paginación y agregaciones. Además, su motor analítico soporta análisis detallados, como lo describe AltexSoft (2023): “nested aggregations that can group users by multiple criteria like city and gender and then calculate the average age for each group”. En otras palabras, es posible agrupar datos por

distintos criterios al mismo tiempo (por ejemplo, ciudad y género) y luego obtener métricas específicas como la edad promedio de cada subgrupo. Este nivel de detalle en el análisis convierte a Elasticsearch en una herramienta ideal para aplicaciones con alto consumo y procesamiento de datos.

En cuanto a la velocidad y rendimiento, AltexSoft (2023) afirma que “Speed and performance are foundational to Elasticsearch's appeal, setting it apart from traditional data storage and retrieval systems”. Esto quiere decir que la rapidez de respuesta y la capacidad de procesamiento son elementos centrales que lo diferencian de otros sistemas más convencionales. Una de las tecnologías clave detrás de este rendimiento es el uso de un índice invertido, el cual registra todas las palabras únicas de los documentos junto con su ubicación, permitiendo localizar información de forma casi instantánea. Gracias a su arquitectura distribuida, que permite la ejecución paralela de consultas en múltiples shards, Elasticsearch logra tiempos de respuesta muy bajos. De hecho, como indica AltexSoft (2023), “in real-world applications with more than 2 billion documents indexed, retrieval speeds have been reported to remain consistently under one second”. Esto implica que incluso con volúmenes masivos de datos, el sistema mantiene tiempos de búsqueda menores a un segundo, superando ampliamente a motores de procesamiento por lotes como Hadoop y habilitando análisis en tiempo real.

b. Desventajas

Si bien las ventajas descritas posicionan a Elasticsearch como una de las soluciones más potentes y versátiles en el ámbito de la búsqueda y el análisis de datos, es importante reconocer que, como cualquier tecnología, también presenta limitaciones que deben ser consideradas antes de su adopción. Estas desventajas no necesariamente restan valor a la herramienta, pero sí representan factores que pueden influir en su implementación y uso efectivo.

Una de ellas es su pronunciada curva de aprendizaje. AltexSoft (2023) indica que “Elasticsearch offers a rich set of features and functionalities, but tapping into its full potential can require overcoming a steep learning curve”. Esto quiere decir que, aunque la herramienta cuenta con un amplio conjunto de capacidades, aprovecharlas al máximo puede requerir un tiempo considerable de estudio y adaptación. Según reseñas de usuarios citadas por AltexSoft (2023), “the query mechanism can be complex to grasp, especially for those

not well-versed in SQL or database concepts”, lo que significa que el sistema de consultas puede resultar complicado para quienes no tienen experiencia previa en SQL o bases de datos. Además, dominar su lenguaje de consultas especializado implica comprender a fondo elementos como analyzers y tokenizers, lo que puede ser intimidante para principiantes y requerir tiempo de adaptación incluso para desarrolladores experimentados.

Otra limitación es la documentación inconsistente e incompleta. AltexSoft (2023) menciona que “its documentation often falls short of expectations”, refiriéndose a que la documentación oficial no siempre satisface las necesidades de los usuarios, especialmente en lo relacionado con buenas prácticas y ejemplos para funcionalidades avanzadas. Asimismo, “documentation gaps, including hardware configuration and capacity planning at scale, exacerbate the product's steep learning curve” (AltexSoft, 2023), lo que significa que la ausencia de guías claras sobre configuración de hardware y planificación de capacidad a gran escala aumenta la dificultad de aprendizaje y optimización del sistema.

No obstante, esta carencia se ve en parte mitigada por una de las ventajas previamente mencionadas: su comunidad activa y en constante crecimiento. Tal como señala AltexSoft (2023), “The rapid response time, with questions often answered in approximately an hour, testifies to the high level of the community's engagement”. Esto implica que, aunque la documentación oficial pueda resultar insuficiente, es posible acceder a soluciones y buenas prácticas a través de foros, Slack, Stack Overflow o el repositorio de GitHub. De esta forma, la comunidad actúa como un soporte complementario que ayuda a cubrir los vacíos de la documentación y a resolver dudas en tiempos reducidos.

También se debe considerar la complejidad en el licenciamiento y la falta de claridad en los costos. Como explica AltexSoft (2023), “the intricacies of Elasticsearch's licensing model can present challenges for organizations”, lo que implica que el modelo de licencias puede generar confusión, especialmente para las empresas que desean integrarlo como una función embebida en sus productos. Además, “the varied pricing models and additional features can render the total cost less transparent” (AltexSoft, 2023), es decir, que la existencia de distintos planes y funcionalidades extra puede dificultar la estimación del costo real, requiriendo un análisis cuidadoso para evitar gastos imprevistos.

Por último, es relevante mencionar que Elasticsearch puede ser una herramienta con alta demanda de recursos. Aunque destaca por su velocidad y rendimiento, AltexSoft (2023)

advierte que “it can be CPU-intensive, especially when handling multiple tasks like indexing, searching, and aggregating data concurrently”. Esto significa que, al procesar operaciones simultáneas como indexación, búsqueda y agregación de datos, su consumo de CPU puede ser considerable. Por lo tanto, se requiere una planificación adecuada de recursos, asegurando un número suficiente de núcleos de procesamiento para mantener un funcionamiento eficiente, el cual dependerá de la carga de trabajo y el caso de uso específico.

C. Futuro de la tecnología

A lo largo de esta investigación se han revisado los antecedentes de Elasticsearch, así como sus principales ventajas y desventajas. En esta sección se abordará cuál es el futuro de esta tecnología, que para algunos usuarios resulta altamente atractiva y para otros plantea ciertas reservas. Desde el inicio se mencionó la influencia que tiene este motor de búsqueda en grandes tendencias como la búsqueda en tiempo real, la Big Data y la integración con la Inteligencia Artificial. Si bien Elastic, la compañía desarrolladora de Elasticsearch, no suele ser completamente abierta respecto a sus planes futuros, a través de la investigación ha sido posible identificar algunas direcciones estratégicas hacia las cuales se encamina esta tecnología.

Iniciando con una de las transformaciones más importantes en la evolución de esta tecnología, en el apartado de antecedentes mencionamos que en el año 2024 Elasticsearch adoptó la licencia AGPL y se convirtió en open source. Este paso no solo representó un cambio en el modelo de licenciamiento, sino también un mensaje claro del compromiso de la compañía con el desarrollo abierto y colaborativo. Tal como se señaló en la sección de ventajas, Elasticsearch cuenta con una comunidad activa y en constante crecimiento; sin embargo, la decisión de abrir su código fortaleció aún más este ecosistema, ya que permitió una mayor participación de desarrolladores de todo el mundo, quienes han contribuido con mejoras, correcciones y sugerencias que enriquecen de manera significativa la plataforma. La comunidad y la academia han reconocido que, a partir de la transición a open source, las contribuciones se han incrementado tanto en cantidad como en calidad, ampliando el horizonte de posibilidades de la tecnología. De acuerdo con la propia página oficial de Elastic, hasta la fecha se contabilizan más de 5 billones de descargas, más de 103,000 estrellas en GitHub y alrededor de 183,000 pull requests, cifras que reflejan no solo la

popularidad del proyecto, sino también el nivel de compromiso de la comunidad global con su mejora y expansión.

Otra dirección clave en el futuro de Elasticsearch es su apertura hacia un ecosistema de bases de datos más amplio y diverso. Según ObjectRocket (2019), Elasticsearch no fue diseñado originalmente para ser una base de datos principal como MySQL o MongoDB, sino que fue concebido como un motor de búsqueda flexible y escalable, capaz de complementar y potenciar otras tecnologías. Esta visión ha llevado a Elastic a expandir sus capacidades para integrarse con diferentes entornos de almacenamiento, con el objetivo de atraer y consolidar una comunidad de desarrolladores más diversa. Un ejemplo claro de esta estrategia es la integración con Hadoop, donde, a través de un complemento (add-on), Elasticsearch habilitó la creación de índices que mejoran significativamente las funciones de búsqueda del ecosistema Hadoop, el cual era criticado por sus limitadas capacidades en este campo. Tal como menciona ObjectRocket (2019), esta extensión no solo optimizó la búsqueda dentro de Hadoop, sino que también abrió la puerta a nuevas soluciones de análisis de datos en entornos de Big Data, ampliando el alcance y relevancia de Elasticsearch en el mercado.

La inclusión de Kibana, previamente mencionada en los antecedentes, también representó un salto cualitativo en el desarrollo del ecosistema de Elastic. Esta herramienta permitió transformar a Elasticsearch en algo más que un motor de búsqueda, al dotarlo de la capacidad de generar gráficos y visualizaciones interactivas de datos, lo que resultó especialmente valioso en contextos donde el análisis en tiempo real es fundamental. A diferencia de motores como MongoDB o MySQL, que carecen de mecanismos nativos de visualización, Kibana hizo posible que los usuarios pudieran interpretar de manera visual y accesible los resultados de sus consultas. ObjectRocket (2019) subraya esta diferencia, destacando que ni Mongo ni MySQL ofrecen una solución comparable dentro de sus sistemas principales. Lo que hoy conocemos como ELK Stack (Elasticsearch, Logstash y Kibana) consolidó así una propuesta única y de gran valor para las organizaciones, no solo por su potencia técnica, sino porque es completamente gratuito, lo que permite a empresas, instituciones académicas y desarrolladores independientes acceder a herramientas de visualización de alto nivel sin incurrir en costos adicionales.

Mencionamos al inicio cómo Elasticsearch sigue de cerca las grandes tendencias tecnológicas, y en particular no se queda atrás en el ámbito de la Inteligencia Artificial. Shay Banon and Philipp Krenn (2025) reconocen que, a pesar de más de una década de avances, aún persisten retos clásicos en el manejo de datos: “15 years of query tuning, and we’re still just throwing boost: 10 at the problem. 15 years of making your logs searchable while you still have no idea what’s happening in production. Still the best at finding that one log line... if you remember how you indexed it.” Este planteamiento deja en evidencia que, aunque Elasticsearch ha logrado consolidarse como líder en búsqueda, la complejidad de sus consultas y el entendimiento de los datos en entornos productivos continúan siendo un desafío.

No obstante, la incorporación de la Inteligencia Artificial se perfila como el elemento clave para redefinir el futuro de esta tecnología. Como explican Banon y Krenn (2025), “AI is redefining what’s possible — from turning raw logs into actionable insights with the AI Assistant for observability and security, to more relevant search with semantic understanding and intelligent re-ranking”, lo que significa que la IA no solo enriquece la búsqueda haciéndola más semántica y relevante, sino que también permite transformar registros en bruto en información accionable, especialmente en áreas críticas como la observabilidad y la seguridad.

Esta evolución supone un cambio profundo: Elasticsearch ya no se limita a ofrecer resultados de búsqueda, sino que avanza hacia un modelo capaz de comprender y contextualizar los datos. Y lo más relevante es que este proceso apenas comienza, tal como advierten Banon y Krenn (2025), “This is only the beginning. More AI-powered features are on the horizon — bringing smarter search, enhanced observability, and stronger security. The future of Elasticsearch isn’t just about finding data; it’s about understanding it.” De este modo, el futuro de Elasticsearch no se concibe únicamente como una herramienta de indexación y consulta de información, sino como una plataforma que, mediante la Inteligencia Artificial, permitirá una búsqueda más inteligente, una observación más detallada de los sistemas y una seguridad reforzada, consolidando su papel como tecnología fundamental en la nueva era del análisis de datos.

Otro de los ejes hacia los cuales apunta el futuro de Elasticsearch es la evolución hacia entornos serverless, eliminando gran parte de la complejidad operativa que históricamente ha acompañado a este motor de búsqueda. Como señalan Banon y Krenn

(2025): “No shards, nodes, or versions. Elasticsearch Serverless — which is GA on AWS and just entered technical preview on Azure — takes care of the operational issues you might have experienced in the past”. Esto significa que con Elasticsearch Serverless desaparecen tareas como la configuración de shards, la gestión de nodos o las actualizaciones de versiones, aspectos que por años han representado dolores de cabeza para los administradores. No es casual que los propios autores hagan un recuento irónico de estas dificultades: “15 years in, and someone is still setting `number_of_shards: 100` for no reason. 15 years, and we’re still debating `refresh_interval: 1s` vs `30s` like it’s a life-or-death decision. 15 years of major versions, minor heart attacks, and the thrill of migrating to the latest version.” Con este reconocimiento se enfatiza que los problemas de configuración y migración, tan comunes en la historia de Elasticsearch, tienden a quedar superados en el modelo serverless, donde la infraestructura pasa a ser manejada automáticamente en la nube.

Actualmente, Elasticsearch Serverless ya está disponible de manera general en AWS y se encuentra en preview técnico en Azure, lo que demuestra el compromiso de Elastic por simplificar la experiencia de uso y liberar a los equipos de operaciones de tareas repetitivas y críticas. A la par de este cambio arquitectónico, otra innovación destacada es el Elasticsearch Piped Query Language (ES|QL), que busca transformar la manera en que los usuarios interactúan con los datos. Como celebran Banon y Krenn (2025): “Cheers to 15 years of Elasticsearch — where the Query DSL is still the most complex part of your day. But it doesn’t have to be.”

Con esta afirmación se reconoce que el lenguaje de consultas tradicional, Query DSL, ha sido durante años un obstáculo por su complejidad, pero con ES|QL se introduce una sintaxis mucho más simple y legible, acompañada de un motor de cómputo diseñado con el rendimiento como prioridad. Aunque su desarrollo continúa y nuevas funcionalidades se encuentran en proceso, ya es posible utilizar ES|QL en la actualidad, ofreciendo una alternativa más accesible para quienes buscan extraer información de manera ágil sin renunciar a la potencia analítica de Elasticsearch. Como aseguran los autores, “Don’t worry; the Query DSL will understand”, destacando que ambas opciones convivirán, permitiendo a los usuarios aprovechar lo mejor de la simplicidad de ES|QL sin abandonar la compatibilidad con el lenguaje de consultas tradicional.

En conclusión, el futuro de Elasticsearch se vislumbra como una etapa de consolidación y expansión estratégica. Tal como señala ObjectRocket (2019), “Elasticsearch

is on an exciting growth journey and the super wide approach appears to suit their long-term strategy well”, lo que refleja que la compañía mantiene una visión de crecimiento sostenido y diversificación en su uso. Este enfoque implica que, si bien Elastic no abandonará las funcionalidades que lo posicionaron como un referente en el mundo de las bases de datos, su estrategia se centra cada vez más en abordar los casos límite que las soluciones tradicionales no logran cubrir. De esta forma, Elasticsearch no pretende sustituir a sistemas como MySQL o MongoDB, sino complementarlos con un alcance más amplio, flexible y adaptable. ObjectRocket (2019) destaca que esta orientación responde a una estrategia de largo plazo que busca no solo conservar la relevancia tecnológica, sino también abrir nuevas oportunidades en campos donde las bases de datos convencionales muestran limitaciones. Así, Elasticsearch se perfila como una herramienta que evoluciona de un motor de búsqueda a un ecosistema integral capaz de responder a desafíos complejos de almacenamiento, búsqueda y análisis de datos.

D. Análisis del Tema

Elasticsearch se ha consolidado como un motor de búsqueda y análisis que ocupa un lugar particular dentro de la arquitectura moderna de datos. No debe entenderse como un sustituto de las bases de datos relacionales o transaccionales tradicionales, sino como un complemento especializado que aporta rapidez y eficiencia en la búsqueda textual, en la indexación distribuida de documentos JSON y en la analítica en tiempo real sobre grandes volúmenes de información. Esta capacidad lo ha convertido en una herramienta clave en entornos donde las consultas textuales, la observabilidad y el análisis de registros de aplicaciones requieren respuestas con baja latencia y alta escalabilidad.

En los últimos años, la evolución de Elasticsearch ha estado marcada por cambios estratégicos que redefinen su propuesta de valor. La adopción de la licencia AGPL en 2024 devolvió a la comunidad el acceso a un modelo abierto, al tiempo que impuso nuevas consideraciones legales para las organizaciones que lo integran en productos comerciales. Este hecho fortaleció el ecosistema académico y de código abierto, aunque obligó a las empresas a evaluar cuidadosamente sus implicaciones de cumplimiento. De forma paralela, la introducción de LogsDB permitió reducir significativamente los costos de almacenamiento de registros, una de las principales limitaciones de la plataforma, mientras que la disponibilidad de Elasticsearch Serverless en la nube simplificó los procesos operativos al eliminar tareas

históricamente complejas como la configuración de shards, la gestión de nodos o las actualizaciones de versiones.

Las ventajas de esta tecnología son evidentes. Su velocidad de respuesta, basada en el uso de índices invertidos y en la ejecución paralela de consultas distribuidas, la coloca por encima de sistemas convencionales para búsquedas masivas, logrando tiempos de consulta de menos de un segundo incluso con miles de millones de documentos. A esto se suma su flexibilidad, ya que combina búsqueda y analítica en una misma API, y la potencia de su ecosistema de herramientas como Kibana y Logstash, que facilitan la recolección, transformación y visualización de datos en tiempo real. El respaldo de una comunidad activa, con foros, repositorios y contribuciones constantes, refuerza además la madurez de la plataforma.

No obstante, estas fortalezas conviven con desafíos que no deben minimizarse. La curva de aprendizaje de Elasticsearch es pronunciada: comprender su lenguaje de consultas, el diseño de índices o el manejo de analyzers y tokenizers exige un nivel de conocimiento que puede resultar abrumador para equipos sin experiencia previa. La documentación oficial, aunque extensa, suele presentar vacíos en aspectos prácticos como la configuración de hardware o la planificación de capacidad a gran escala, lo que obliga a los usuarios a depender de la comunidad para resolver problemas. A ello se suma la complejidad de su modelo de licenciamiento y la falta de transparencia en algunos esquemas de costos, factores que pueden generar incertidumbre en proyectos empresariales de largo plazo. Finalmente, su alta demanda de recursos en escenarios de indexación, búsqueda y agregación simultáneas obliga a planificar con cuidado la infraestructura necesaria, lo que incrementa los requerimientos de hardware y, en consecuencia, el costo total de propiedad.

Al compararlo con otras alternativas, se observa que Elasticsearch no busca sustituir a bases de datos relacionales como MySQL o PostgreSQL, ni a sistemas de análisis masivo como Hadoop, sino complementarlos. Frente a motores transaccionales, carece de integridad relacional y de mecanismos ACID robustos, pero aporta una flexibilidad y rapidez que estos no pueden ofrecer en búsquedas textuales complejas. Frente a alternativas como OpenSearch, su principal competidor, mantiene una ventaja en innovación y soporte corporativo, aunque OpenSearch ofrece un modelo de licencia más permisivo que puede resultar atractivo en ciertos escenarios.

En síntesis, la pertinencia de adoptar Elasticsearch depende del tipo de proyecto y de los recursos disponibles. Su uso resulta especialmente recomendable cuando se trabaja con volúmenes masivos de datos semiestructurados, cuando la velocidad de búsqueda y el análisis en tiempo real son críticos, y cuando se dispone de la infraestructura y del personal capacitado para administrarlo correctamente. Por el contrario, en contextos de datos pequeños, consultas simples o equipos con poca experiencia, puede ser más apropiado optar por soluciones tradicionales más sencillas y económicas.

En el ámbito académico, Elasticsearch se presenta como una plataforma ideal para experimentar con conceptos de Big Data, observabilidad y analítica distribuida, gracias a su naturaleza abierta y a las facilidades que hoy ofrecen las versiones serverless en la nube. En el mundo profesional, en cambio, la decisión de adoptarlo debe equilibrar su innegable capacidad técnica con los costos operativos, las implicaciones legales de su modelo de licencia y la curva de aprendizaje que exige. Todo esto confirma que Elasticsearch no es una solución universal, sino una herramienta estratégica cuyo verdadero valor se alcanza cuando se implementa de forma planificada, consciente y alineada con las necesidades específicas de cada organización.

III. Conclusiones y Recomendaciones

A. Conclusiones

A lo largo de este documento se ha realizado un análisis exhaustivo de todo lo que ofrece Elasticsearch como tecnología. Hemos observado cómo, a lo largo de más de quince años de existencia, este motor de búsqueda se ha consolidado como una herramienta clave para la exploración y el análisis de grandes volúmenes de datos en tiempo real. Desde sus orígenes en el año 2010, cuando fue concebido con un enfoque hacia la búsqueda de texto libre, hasta su integración dentro del conocido ELK Stack (Elasticsearch, Logstash y Kibana), hemos sido testigos de su evolución constante y de cómo ha sabido adaptarse a las nuevas necesidades del mercado tecnológico. Aún en la actualidad continúa incorporando funcionalidades que le permiten mantenerse vigente y competitivo.

El estudio desarrollado incluyó no solo sus antecedentes, sino también una revisión detallada de sus principales ventajas. Entre ellas destacan su alto rendimiento, la amplia compatibilidad con diversos lenguajes de programación, así como el respaldo de una gran comunidad activa que contribuye a su crecimiento y mejora continua. Todas estas características explican por qué Elasticsearch se ha convertido en una de las tecnologías más utilizadas en el mundo, no solo en el ámbito empresarial, sino también en sectores como la seguridad informática, la salud, la banca, la educación y los servicios digitales. Tal y como menciona AltexSoft (2023), la popularidad de este motor radica en su capacidad para responder de manera eficiente a los retos de búsqueda y analítica que enfrentan organizaciones de cualquier tamaño.

Por otra parte, también se identificaron los desafíos y desventajas que acompañan a esta tecnología. Estos aspectos no significan que Elasticsearch sea un mal motor de búsqueda, pero sí recalcan la importancia de que su adopción se realice con plena conciencia de sus implicaciones. Entre sus principales limitaciones se encuentran la complejidad en su administración y mantenimiento, la alta curva de aprendizaje que puede dificultar su adopción inicial, y el elevado consumo de recursos tanto a nivel de hardware como de infraestructura. Adicionalmente, se destacó la incertidumbre que generan los constantes cambios entre las versiones de pago y la versión de código abierto, ya que este aspecto puede impactar en la estabilidad y en la planificación a largo plazo de las organizaciones que lo utilizan.

No obstante, cabe resaltar que ninguna de estas limitaciones puede considerarse catastrófica. Al contrario, se trata de factores que, si se gestionan de manera adecuada, no impiden aprovechar las múltiples ventajas de Elasticsearch. La clave está en evaluar correctamente los requerimientos de cada proyecto y asegurarse de que el uso de este motor responde a necesidades concretas, evitando así frustraciones técnicas o costos inesperados.

En conclusión, los hallazgos de este análisis permiten afirmar que Elasticsearch constituye una solución robusta, madura y ampliamente validada por el mercado, capaz de responder a los requerimientos de búsqueda y análisis de datos a gran escala. Su consolidación como estándar en la industria de la búsqueda, el monitoreo y el registro de datos (logging) lo posiciona como una de las tecnologías de referencia en la actualidad. Además, su futuro apunta hacia una mayor integración con la inteligencia artificial, la automatización y la simplificación operativa, elementos que fortalecerán aún más su papel dentro del ecosistema tecnológico.

No obstante, es importante reiterar que su adopción debe ir acompañada de buenas prácticas de implementación, monitoreo constante y una planificación estratégica clara que considere sus fortalezas y debilidades. Solo así será posible obtener el máximo provecho de este motor y garantizar que los resultados estén a la altura de las expectativas. En definitiva, Elasticsearch no solo ha demostrado ser una tecnología de gran impacto en el presente, sino que también se proyecta como una herramienta con gran potencial de crecimiento en los escenarios futuros de la analítica y la gestión de datos.

B. Recomendaciones

Debido a los análisis realizados a lo largo de esta investigación, se presentan una serie de recomendaciones que consideramos fundamentales para quienes decidan implementar Elasticsearch en sus empresas, organizaciones o proyectos de desarrollo. Estas sugerencias buscan servir como una guía práctica que facilite una adopción consciente, eficiente y alineada con las necesidades reales de cada entorno.

En primer lugar, es imprescindible evaluar cuidadosamente la pertinencia del caso de uso antes de optar por Elasticsearch. Aunque este motor se distingue por su gran capacidad en búsquedas de texto, análisis de datos complejos y respuestas en tiempo real, no en todos los contextos representa la opción más adecuada. Si el proyecto a desarrollar implica un volumen

de datos modesto o consultas relativamente simples, es probable que una base de datos tradicional sea suficiente para cubrir los requerimientos. Por lo tanto, Elasticsearch debe adoptarse únicamente cuando se requieran sus ventajas diferenciales: escalabilidad, velocidad y capacidad para manejar grandes volúmenes de información no estructurada. Esta reflexión inicial asegura que la implementación aporte un valor tangible frente a soluciones más simples y menos costosas.

Una segunda recomendación esencial consiste en la planificación estratégica de la infraestructura y los recursos necesarios. Tal como se discutió en el análisis, Elasticsearch es una tecnología demandante en términos de consumo de CPU, memoria y almacenamiento. Por ello, resulta prioritario dimensionar correctamente los recursos desde un inicio. Se aconseja implementar un entorno de pruebas para identificar los requisitos concretos del sistema y, a partir de ello, definir una infraestructura adecuada. La experiencia señala que el uso de equipos con suficiente memoria RAM y unidades de almacenamiento SSD de alto rendimiento optimiza significativamente la experiencia. Asimismo, es recomendable monitorear de manera continua el consumo de recursos y diseñar una estrategia de escalabilidad horizontal, incorporando nodos adicionales conforme la carga lo exija. De esta forma se evitan cuellos de botella y se garantiza un rendimiento sostenido en el tiempo.

De igual manera, es crucial invertir en la capacitación del personal técnico. Elasticsearch presenta una curva de aprendizaje pronunciada, lo cual convierte la formación en un factor determinante para el éxito del proyecto. Se recomienda que tanto desarrolladores como administradores reciban entrenamiento formal, estudien la documentación oficial y se familiaricen con las mejores prácticas de diseño de consultas, optimización de índices y administración de clústeres. Implementar buenas prácticas desde el inicio permite anticipar y prevenir problemas de mayor magnitud en etapas posteriores. A la par de ello, resulta fundamental adoptar herramientas de gestión de datos, como el Index Lifecycle Management (ILM), que automatiza la rotación y eliminación de índices, contribuyendo así a reducir costos y simplificar la operación. En este mismo sentido, puede ser conveniente analizar la posibilidad de implementar Elasticsearch como servicio mediante plataformas como Elastic Cloud o Amazon OpenSearch Service, que ofrecen una alternativa práctica para gestionar escalabilidad y reducir la carga operativa asociada al mantenimiento.

Otra recomendación relevante se centra en la actualización continua del sistema. Elasticsearch es un proyecto en evolución constante, en el que periódicamente se liberan

versiones que incluyen mejoras de seguridad, optimización de rendimiento y nuevas funcionalidades. Mantener los clústeres, ya sean físicos o virtuales, siempre actualizados es esencial para garantizar tanto la protección como la eficiencia del entorno de trabajo. En paralelo, resulta altamente recomendable fomentar la participación en la comunidad de Elasticsearch, ya que foros, conferencias, blogs especializados y grupos de usuarios constituyen espacios donde se comparten soluciones, experiencias y complementos de gran valor. La interacción con esta comunidad brinda acceso al conocimiento colectivo y permite adoptar prácticas probadas que reducen significativamente los tiempos de resolución de problemas.

En conclusión, Elasticsearch es una herramienta con un potencial inmenso, pero para aprovecharlo plenamente se requiere una implementación planificada, capacitada y sostenida en buenas prácticas. La adopción consciente de estas recomendaciones no solo permitirá mitigar los riesgos y desafíos que conlleva el uso de este motor, sino que también maximizará sus beneficios. A lo largo de esta investigación hemos constatado que Elasticsearch se ha convertido en una tecnología madura, potente y ampliamente adoptada, capaz de transformar la manera en que las organizaciones gestionan y analizan datos. No obstante, es la preparación, el monitoreo constante y el compromiso con la mejora continua lo que convierte a este motor de búsqueda en un verdadero aliado estratégico para alcanzar altos niveles de eficiencia en la exploración y análisis de información.

Bibliografía

- Shannon. B, Krenn. P . (2025, Enero 30). *Elasticsearch history: 15 years of search innovation*. Elastic
<https://www.elastic.co/search-labs/blog/elasticsearch-history-15-years>.
- IBM. (2021, Septiembre 29.). *¿Qué es Elasticsearch?* IBM. Recuperado el 16 de agosto de 2025 de <https://www.ibm.com/es-es/think/topics/elasticsearch>.
- AltexSoft. (2023, Octubre 12). *Elasticsearch: pros and cons of using this search engine*. AltexSoft. <https://www.altexsoft.com/blog/elasticsearch-pros-cons/>.
- Sankaran. P. (2021, Febrero 9). *Elastic's return to open source*. Revenera. <https://www.revenera.com/blog/software-composition-analysis/elastics-return-to-open-source/>.
- ObjectRocket. (2019, Julio 11). *The future of Elasticsearch*. ObjectRocket. <https://www.objectrocket.com/resource/the-future-of-elasticsearch/>.
- Elastic. (s. f.). *Open source at Elastic*. Elastic. Recuperado el 16 de agosto de 2025 de <https://www.elastic.co/about/open-source>.