



# elasticsearch

**POR: ANDRÉS FELIPE MARTÍNEZ H.  
UNIVERSIDAD CENTRAL  
MS. ANALÍTICA DE DATOS**



<https://elastic.co/es/elasticsearch>

# Contenido

- 01** | ¿QUÉ ES?
- 02** | HISTORIA DE ELASTICSEARCH
- 03** | CARACTERISTICAS
- 04** | CASOS DE USO
- 05** | VENTAJAS Y DESVENTAJAS



elasticsearch

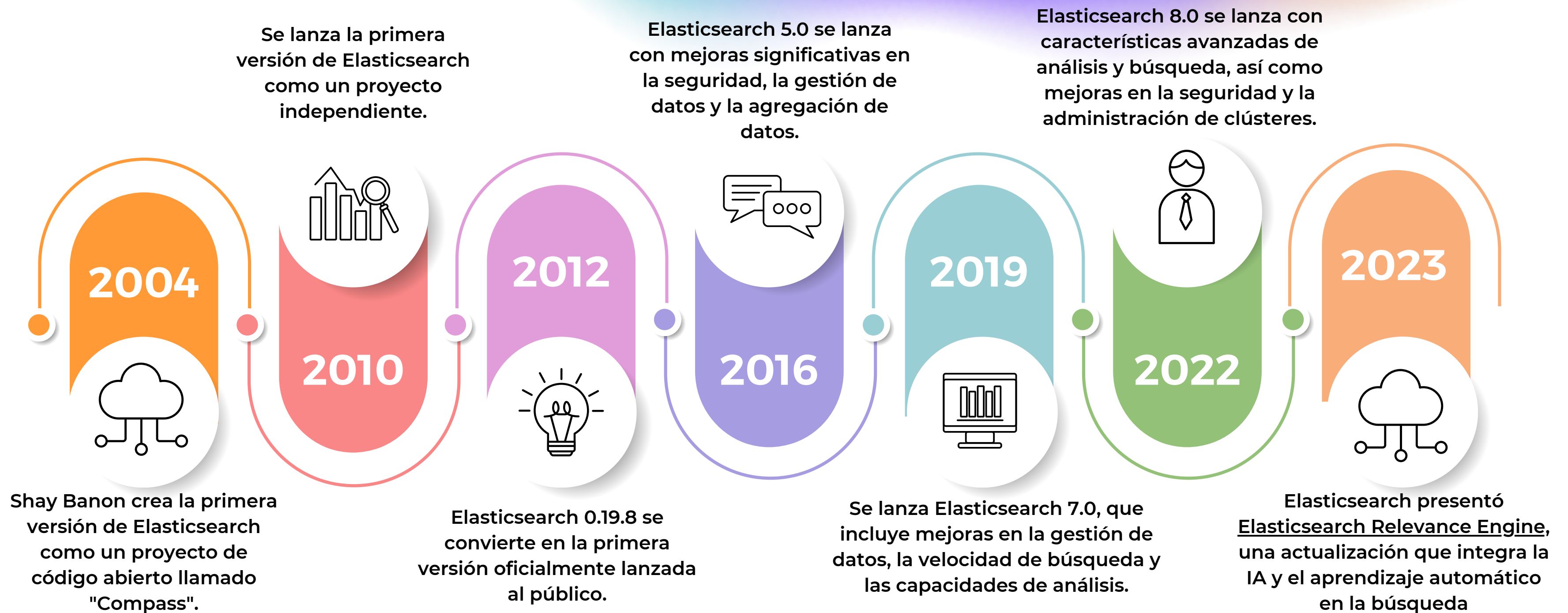
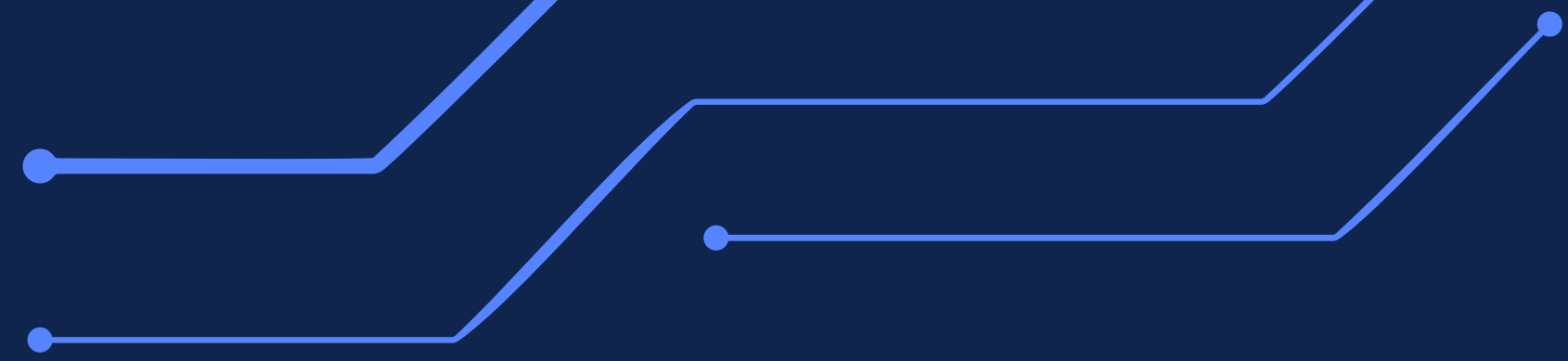
# ¿Qué es Elasticsearch?

Elasticsearch es un motor avanzado de búsqueda y análisis de código abierto que también funciona como una base de datos NoSQL. Está desarrollado en Java y construido sobre la biblioteca Apache Lucene.

Está diseñado para almacenar, buscar y analizar grandes volúmenes de datos de manera rápida y eficiente.



# Historia de Elasticsearch



# Características de Elasticsearch

Dentro de las principales características de Elasticsearch encontramos.



## BÚSQUEDAS DE ALTA VELOCIDAD

A diferencia de las bases de datos que funcionan a través de tablas y esquemas, Elasticsearch emplea una estructura basada en índices.

## INTERACCIÓN CON EL USUARIO

Interactúa a través de API REST integrales, procesando y devolviendo formatos tipo JSON.

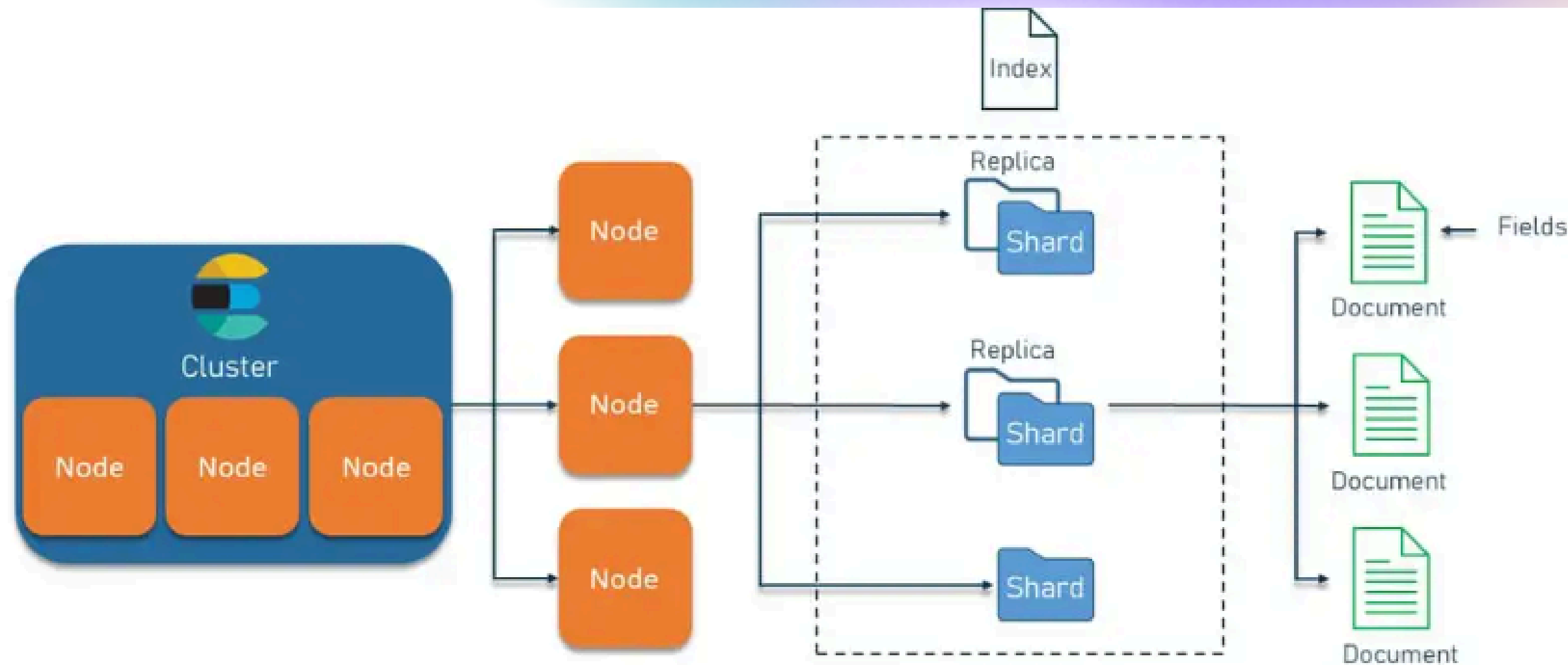
## DISTRIBUCIÓN Y ESCALABILIDAD

Está diseñada para escalar horizontalmente. Puede manejar grandes volúmenes de datos distribuidos en múltiples nodos.

## ANÁLISIS DE TEXTO COMPLETO

Utiliza algoritmos de análisis de lenguaje natural para indexar y buscar texto completo, permitiendo filtrado, búsquedas avanzadas y análisis de contenido.

# ¿Cómo funciona?



Su arquitectura está diseñada para ser distribuida. Sus capacidades, almacenamiento, ejecución de búsquedas y análisis no se limitan a una sola máquina o servidor, sino que se distribuyen entre varios nodos de un clúster.

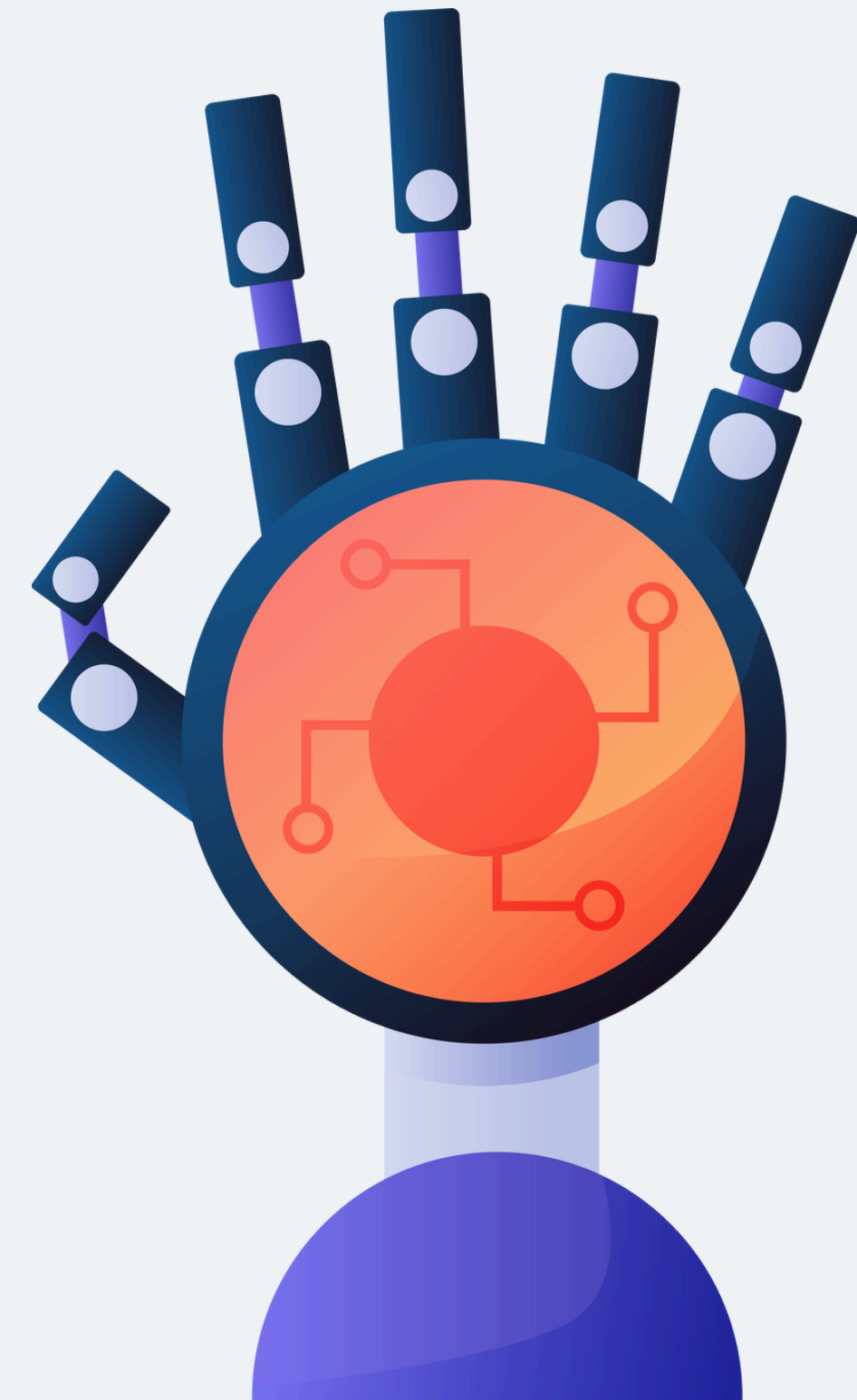


# Conceptos básicos

**Nodo Maestro:** Es el nodo de control en un clúster. Es responsable de crear o eliminar índices, reasignar particiones cuando los nodos abandonan o ingresan al cluster y hacer seguimiento de los nodos.

**Nodo de datos:** Almacena datos y realiza operaciones relacionadas con ellos, como la búsqueda y la agregación. Son la base del funcionamiento de Elasticsearch.

**Nodo cliente:** actúa como un equilibrador de carga inteligente que reenvía las solicitudes a los nodos adecuados en cada clúster.



**Clúster:** entidad formada por uno o varios nodos que trabajan juntos para satisfacer las solicitudes de indexación, búsqueda y análisis de datos.

# Más conceptos básicos

**Fragmentos:** posibilitan la escalabilidad horizontal al dividir índices en unidades independientes. Esta partición permite distribuirlos entre varios servidores.

**Replicas:** Las particiones de réplica son copias de las particiones principales y tienen dos propósitos principales: tolerancia a errores y equilibrio de carga.

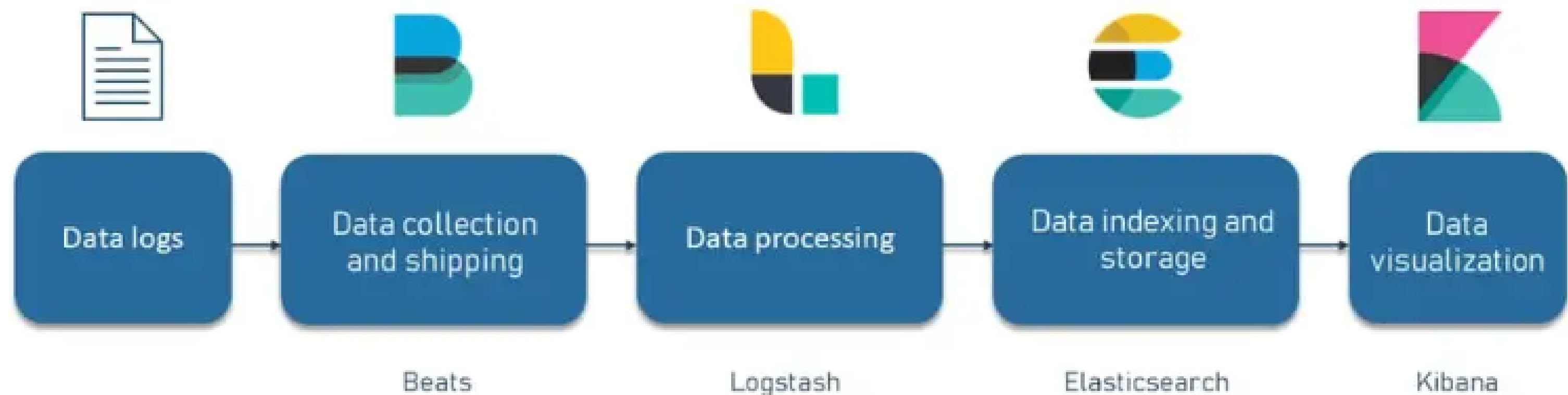
**Campos:** Son la unidad de datos más pequeña y funciona como pares "clave-valor" dentro de un documento JSON que se almacena en un índice.



**Índice invertido:** para cada término único que se encuentra en el documento, el índice invertido enumera los documentos en los que aparece el término.



# Pilas elásticas



Elasticsearch hace parte del flujo de trabajo o “pila elástica” ofrecida por Elastic, junto con otras herramientas de análisis de datos.

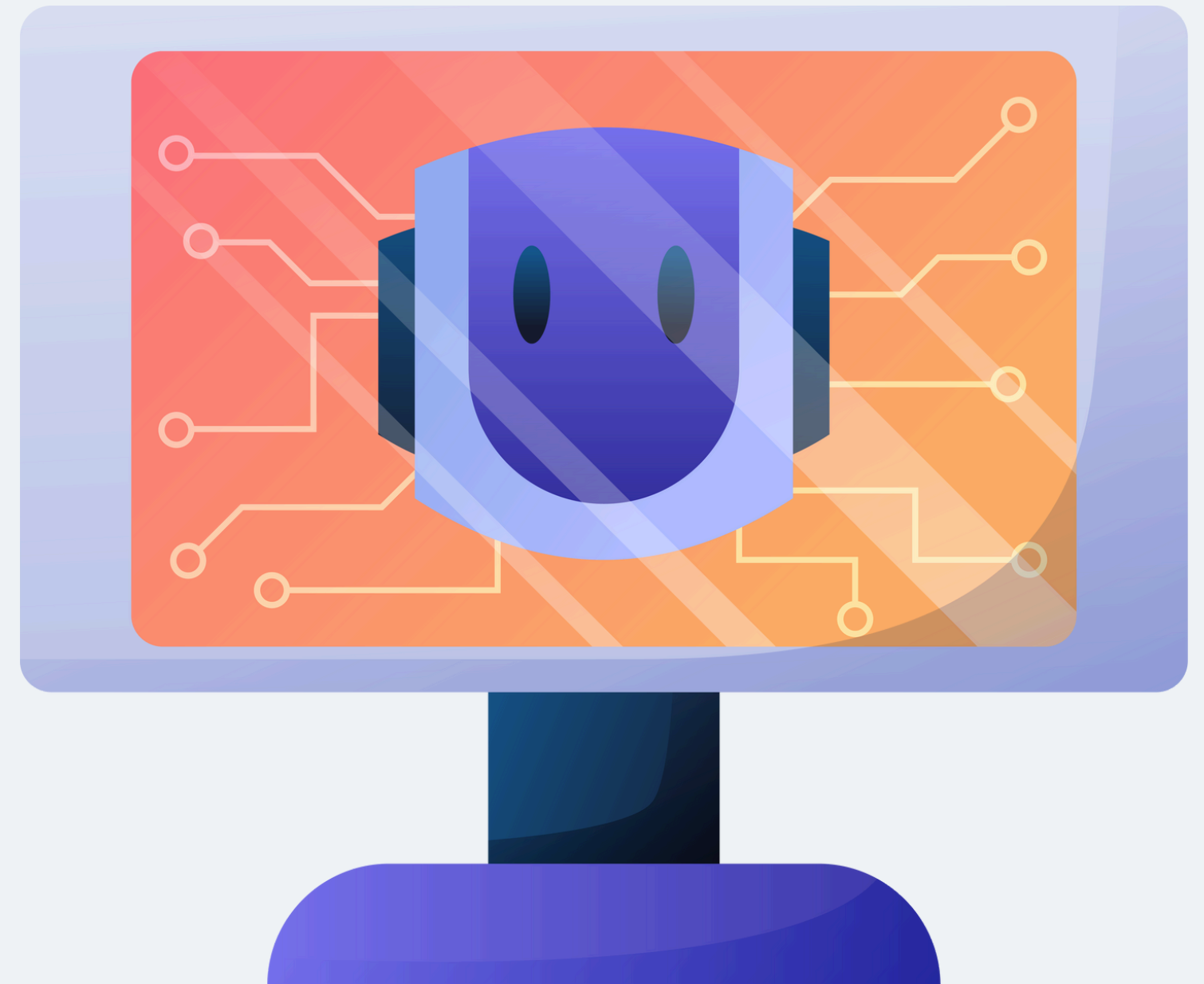
# Casos de uso

**Función de búsqueda de aplicaciones y sitios web.**

**Análisis de registros, métricas y eventos de seguridad**

**Modelado de comportamiento en tiempo real con machine learning**

**Automatización de flujos de trabajo empresariales complejos**



**elasticsearch**

# Ventajas de Elasticsearch

Dentro de las principales ventajas de Elasticsearch encontramos.



## ESCALABILIDAD

Es altamente escalable y puede manejar grandes volúmenes de datos distribuidos en múltiples nodos en un clúster.

## VELOCIDAD Y RENDIMIENTO

Está diseñado para proporcionar resultados de búsqueda casi instantáneos, incluso en grandes conjuntos de datos.

## FLEXIBILIDAD DE CONSULTAS Y BUSQUEDA

No requiere un esquema rígido para indexar datos, lo que permite trabajar con datos semi-estructurados y no estructurados.

## TOLERANCIA A FALLOS Y RECUPERACIÓN

Replica automáticamente los datos. Cada documento de un índice pertenece a una partición principal, y las particiones de réplica actúan como sus copias,

# Desventajas de Elasticsearch

Por otra parte, entre las principales desventajas de Elasticsearch encontramos.

## ALTO CONSUMO DE RECURSOS

Para implementaciones a gran escala, puede ser necesario invertir en hardware y recursos de infraestructura significativos para garantizar un rendimiento óptimo.

## CURVA DÍFICIL DE APRENDIZAJE

Puede requerir un tiempo de aprendizaje significativo, especialmente para configuraciones avanzadas y optimización de consultas.

## DOCUMENTACIÓN INCOMPLETA

Evidencia una falta de tutoriales completos, e incluso los desarrolladores experimentados notan la ausencia de ejemplos detallados para funcionalidades avanzadas.

## LICENCIA COMPLEJA Y ALTO COSTO

Aunque Elasticsearch es de código abierto, las empresas que requieren soporte premium o funcionalidades avanzadas pueden incurrir en costos significativos.



# ¡Gracias!



**UNIVERSIDAD  
CENTRAL**

Vigilada Mineducación