

# Elasticsearch developer



Elasticsearch is a distributed, open-source search and analytics engine for all types of data, including textual, numerical, geospatial, structured, and unstructured.

Elasticsearch is built on Apache Lucene and was first released in 2010 by Elasticsearch N.V. (now known as Elastic). Known for its simple REST APIs, distributed nature, speed, and scalability, Elasticsearch is the central component of the Elastic Stack, a set of open-source tools for data ingestion, enrichment, storage, analysis, and visualization.

Commonly referred to as the ELK Stack (after Elasticsearch, Logstash, and Kibana), the Elastic Stack now includes a rich collection of lightweight shipping agents known as Beats for sending data to Elasticsearch.

Table of Content:

Starting with some basics

- Basics of Full text search and Information Retrieval
- Overview of the Elastic stack
- Elasticsearch and the REST API
- Using Elasticsearch from your favourite programming language
- Search and the various query types
- Hands-on experience with indexing and searching texts

The Analysis Chain and Index Mappings

- The inverted index and full-text search
- Term normalization with Analysers, Tokenizers and TokenFilters
- Understanding and poking into the analysis chain
- Creating and using a custom analyser
- Using Index Mappings to control analysis and other index features

The Search API

- Pagination and Sorting
- Precision and Recall
- Understanding scoring and how it is applied
- Building smart queries that can influence scoring correctly
- Scripting
- Query explanation and profiling
- Results highlighting
- Various power query tools and a lot of good advice

Elasticsearch must-knows

- Document oriented design and why it's crucial to do right with Elasticsearch
- Suggesters
- Record linkage via MoreLikeThis
- Geo-spatial search
- Multi-lingual search

- Anomaly detection methods
- The percolator

The aggregations framework, Logstash, Beats and Kibana

- Real-time data analysis and reporting
- The Aggregations Framework: Metric and Bucket aggregations
- Pipeline aggregations
- Various powerful aggregations tricks
- Using Kibana as a powerful Web UI on top of the aggregation's framework
- Timelion
- Logstash and Beats

