

Updated 05/02/2025

[Sign up](#)

ElasticSearch training

2 days (14 hours)

Presentation

[ElasticSearch](#) is a next-generation open source search and indexing engine. Its distributed nature and ability to be resilient and highly available have already won over major players such as Wikipedia, LinkedIn, Netflix, Ebay and WordPress.

ElasticSearch has been specially designed to index very large volumes of data, while ensuring high-performance scalability and fault tolerance.

This course introduces the concepts of search engines, before detailing the basic features of Elasticsearch.

This course provides all the knowledge you need to use and exploit Elastic Search effectively, and solve the most common problems encountered.

Our Elastic Search training course will use the latest stable version of the project ([Elasticsearch 8.17](#) to date).

Objectives

- Learn how to use Elasticsearch, from installation to configuration and integration
- Use Elasticsearch's Java API and client to index and search documents
- Monitor Elasticsearch and understand usage statistics
- Identify the challenges and use cases of a search engine
- Situating Elastic Search in a Big Data environment
- Understanding how Elastic Search works
- Index large volumes of data
- Understand system administration and monitoring to ensure system availability

Target audience

- Architect
- Developer
- Project Manager
- Big Data

Prerequisites

- Knowledge of Java & Linux
- Ideally have taken [our ElasticStack training course](#)
- [Test My Knowledge](#)

ElasticSearch training program

Introduction to search engines

- General
- Overview of Open Source search solutions
- Integrating a search engine into an application
- Challenges and keys to success
- Overview of new features in [versions 6, 7 & 8](#)

Introduction to Elasticsearch

- Project history
- Apache Lucene
- What Elastisearch brings to the table compared with Lucene

Elasticsearch basics

- Clustering principles
- Installation
- Configuration
- Notion of Node, Index and Type
- Data partitioning
- Rest API overview

Observability at Elastic

- The pillars of observability
 - Monitoring
 - Logs
 - Metrics
 - API traces
- Sending data from the Elastic server

- Metric shipping data
- Shipping log data
- Application performance monitoring (APM)
- Observability applications

Document indexing

- Index and document design
- Index or delete documents with the Rest API
- Mass indexing
- Version
- Weighting
- Other functions (routing, consistency, child document, etc.)

Mapping

- Definition and role of mapping
- Field type
- Predefined fields
- Index metadata

Text analysis and extraction

- Basics of text extraction and analysis
 - Analysts
 - Char filters
 - Tokenizers
 - Token Filters
- Use cases
- Configuring and using predefined or custom Analyzers
- Multilingual text analysis
- Word deletion
- Extraction of email addresses and urls
- Removing HTML tags from text
- Setting up a spelling correction system
- Indexing binary files (using Apache Tika)

Document search

- Searching documents with the Rest API
- Results management

- Query types
 - match_all query
 - Query type query_string/simple_query_string
 - Match queries (and derivatives)
 - Term and terms queries
 - Wildcard request
 - Range query
 - Fuzzy query
 - Boolean query
 - Existing/missing request
 - and, or and not queries
 - Type and id queries
- Differentiate queries and filters
- Filter types
- Efficiently combining filters
- Relevance

Advanced search functions

- Relevance and score management
- Suggestions
- Autocomplete
- Highlighting
- Search Objects / Nested / Parent-Child
- More Like This
- Geospatial research

Elasticsearch and Java API

- Available customer types
- Integrating elasticsearch into a Java application
 - Embedded
 - Customer node
 - Customer transport
- Using the Java API
- Index and search documents
- Managing indexes
- Managing mappings

Cloud and Clustering

- How an elasticsearch cluster works
- Prevent split brains
- Setting up an elasticsearch cluster
- Scalability and data volume
- Backing up and restoring an elasticsearch cluster

- Monitoring an elasticsearch cluster :
 - API health
 - API state
 - API stats
 - pending_task API
 - API _nodes alias node info
 - hot_threads API
 - the _cat API
 - Monitoring plugins
 - Logs

Advanced features

- Other types of elasticsearch plugins
- River: definition
- Scripting
- Pre-heating your cluster
- Percolation
- Node maintenance
- Distribution of indexes on different nodes
- Queries and statistics

ADDITIONAL MODULE IN ENGLISH ON REQUEST (+2 DAYS)

- Training language: English
- Course level : Beginner to intermediate

This training course enables you to master the basic concepts of Elasticsearch and explore its full range of functions. It provides the knowledge needed to use Elasticsearch effectively, based on real-life use cases. The course covers best practices and frequently encountered problems.

Theory: 60% Practical: 40%

Audience :

- Software Developers
- Data Engineers
- Architects

Prerequisites :

- Knowledge of REST/HTTP, Json, Yaml are appreciated
- No knowledge required

Getting Started

- Elasticsearch Overview
- Keys Features
- Basic Concepts
- Install Elasticsearch
- CRUD Operations
- First steps on Search API

Mapping and Analysis

- Introduction
- Data Types
- Main parameters
- API mapping
- Analysis and Inverted Index
- Custom Analyzer
- Multi-Fields

Querying

- Search API Overview
- Terms Search
- Full Text Search
- Compound Queries

Aggregations

- Aggregations Overview
- Metrics, Aggregations
- Buckets Aggregations
- Pipelines Aggregations

Modeling data

- Denormalization
- Object relationship
- Nested relationship
- Parent/Child relationship

Dynamic mapping and Templates

- Dynamic field mapping
- Dynamic template
- Template API

Nodes and Cluster Management

- Multi-node cluster
- Node Types
- Cluster settings

Ingest API

- Ingest Node
- API pipeline
- Scripting

Overview

- Script API
- Painless language

More Features

- Alias
- Watcher
- Highlighting
- Suggestion
- Reindex Data
- Update by Query and Delete by Query

Complementary module: Preparing for Elastic Engineer certification (+2 days)

Companies concerned

This course is aimed at both individuals and companies, large or small, wishing to train their teams in a new advanced computer technology, or to acquire specific business knowledge or modern methods.

Positioning on entry to training

Positioning at the start of training complies with Qualiopi quality criteria. As soon as registration is finalized, the learner receives a self-assessment questionnaire which enables us to assess his or her estimated level of proficiency in different types of technology, as well as his or her expectations and personal objectives for the training to come, within the limits imposed by the selected format. This questionnaire also enables us to anticipate any connection or security difficulties within the company (intra-company or virtual classroom) which could be problematic for the follow-up and smooth running of the training session.

Teaching methods

Practical course: 60% Practical, 40% Theory. Training material distributed in digital format to all participants.

Organization

The course alternates theoretical input from the trainer, supported by examples, brainstorming sessions and group work.

Validation

At the end of the session, a multiple-choice questionnaire verifies the correct acquisition of skills.

Sanction

A certificate will be issued to each trainee who completes the course.