

Apache Solr



Overview

Accelerate your enterprise search engine and bring relevancy in your search analytics

Description

Key Features

- A practical guide in building expertise with Indexing, Faceting, Clustering and Pagination
- Master the management and administration of Enterprise Search Applications and services seamlessly
- Handle multiple data inputs such as JSON, xml, pdf, doc, xls,ppt, csv and much more.

What You Will Learn

- Design schema using schema API to access data in the database
- Advance querying and fine-tuning techniques for better performance
- Get to grips with indexing using Client API
- Set up a fault tolerant and highly available server with newer distributed capabilities, SolrCloud
- Explore Apache Tika to upload data with Solr Cell
- Understand different data operations that can be done while indexing
- Master advanced querying through Velocity Search UI, faceting and Query Re-ranking, pagination and spatial search
- Learn to use JavaScript, Python, SolrJ and Ruby for interacting with Solr

By the end of this course, you will gain advanced proficiency to build out-of-box smart search solutions for your enterprise demands.

Note

- Lab 1 is theoretical Lab.
- Lab 4, 5, 6, 7 and 8 do not have separate code files.
- Lab 2,3 and 9 files are provided. Open "/root/Desktop/apache-solr/lab_files" directory in VScode to view files.
- Most of the code is command-line. Some of the files comes in-built with the setup, we need to change the content of the files and run them.

Labs

Labs for this course are listed below:

Lab Guides: `Lab_1 - Lab_9`

Lab Files Directory: `lab_files`

1. Introduction to Solr
2. Getting Started
3. Designing Schemas
4. Mastering Text Analysis Methodologies
5. Data Indexing and Operations
6. Advanced Queries -- Part I
7. Advanced Queries -- Part II
8. Managing and Fine-Tuning Solr
9. Client APIs -- An Overview

Labs Part 2

Lab Files Directory: `exercise_files`

1. Apache Solr - Core
2. Apache Solr - Indexing Data
3. Apache Solr - Adding Docs (XML)
4. Apache Solr - Updating Data
5. Apache Solr - Deleting Documents
6. Apache Solr - Querying Data
7. Apache Solr - Faceting

Labs Part 3

1. Solr Tutorial

Extra**About**

Apache Solr is the only standalone enterprise search server with a REST-like application interface. providing highly scalable, distributed search and index replication for many of the world's largest internet sites.

To begin with, you would be introduced to how you perform full text search, multiple filter search, perform dynamic clustering and so on helping you to brush up the basics of Apache Solr. You will also explore the new features and advanced options released in Apache Solr 7.x which will get you numerous performance aspects and making data investigation simpler, easier and powerful. You will learn to build complex queries, extensive filters and how are they compiled in your system to bring relevance in your search tools. You will learn to carry out Solr scoring, elements affecting the document score and how you can optimize or tune the score for the application at hand. You will learn to extract features of documents, writing complex queries in re-ranking the documents. You will also learn advanced options helping you to know what content is indexed and how the extracted content is indexed. Throughout the course, you would go through complex problems with solutions along with varied approaches to tackle your business needs.