Apache Solr for Indexing Data

By Sachin Handiekar, Anshul Johri

Apache Solr for Indexing Data By Sachin Handiekar, Anshul Johri

Enhance your Solr indexing experience with advanced techniques and the built-in functionalities available in Apache Solr

About This Book

- Learn about distributed indexing and real-time optimization to change index data on fly
- Index data from various sources and web crawlers using built-in analyzers and tokenizers
- This step-by-step guide is packed with real-life examples on indexing data

Who This Book Is For

This book is for developers who want to increase their experience of indexing in Solr by learning about the various index handlers, analyzers, and methods available in Solr. Beginner level Solr development skills are expected.

What You Will Learn

- Get to know the basic features of Solr indexing and the analyzers/tokenizers available
- Index XML/JSON data in Solr using the HTTP Post tool and CURL command
- Work with Data Import Handler to index data from a database
- Use Apache Tika with Solr to index word documents, PDFs, and much more
- Utilize Apache Nutch and Solr integration to index crawled data from web pages
- Update indexes in real-time data feeds
- Discover techniques to index multi-language and distributed data in Solr
- Combine the various indexing techniques into a real-life working example of an online shopping web application

In Detail

Apache Solr is a widely used, open source enterprise search server that delivers

powerful indexing and searching features. These features help fetch relevant information from various sources and documentation. Solr also combines with other open source tools such as Apache Tika and Apache Nutch to provide more powerful features.

This fast-paced guide starts by helping you set up Solr and get acquainted with its basic building blocks, to give you a better understanding of Solr indexing. You'll quickly move on to indexing text and boosting the indexing time. Next, you'll focus on basic indexing techniques, various index handlers designed to modify documents, and indexing a structured data source through Data Import Handler.

Moving on, you will learn techniques to perform real-time indexing and atomic updates, as well as more advanced indexing techniques such as de-duplication. Later on, we'll help you set up a cluster of Solr servers that combine fault tolerance and high availability. You will also gain insights into working scenarios of different aspects of Solr and how to use Solr with e-commerce data.

By the end of the book, you will be competent and confident working with indexing and will have a good knowledge base to efficiently program elements.

Style and approach

This fast-paced guide is packed with examples that are written in an easy-to-follow style, and are accompanied by detailed explanation. Working examples are included to help you get better results for your applications.

Download Apache Solr for Indexing Data ...pdf

Read Online Apache Solr for Indexing Data ...pdf

Apache Solr for Indexing Data

By Sachin Handiekar, Anshul Johri

Apache Solr for Indexing Data By Sachin Handiekar, Anshul Johri

Enhance your Solr indexing experience with advanced techniques and the built-in functionalities available in Apache Solr

About This Book

- Learn about distributed indexing and real-time optimization to change index data on fly
- Index data from various sources and web crawlers using built-in analyzers and tokenizers
- This step-by-step guide is packed with real-life examples on indexing data

Who This Book Is For

This book is for developers who want to increase their experience of indexing in Solr by learning about the various index handlers, analyzers, and methods available in Solr. Beginner level Solr development skills are expected.

What You Will Learn

- Get to know the basic features of Solr indexing and the analyzers/tokenizers available
- Index XML/JSON data in Solr using the HTTP Post tool and CURL command
- Work with Data Import Handler to index data from a database
- Use Apache Tika with Solr to index word documents, PDFs, and much more
- Utilize Apache Nutch and Solr integration to index crawled data from web pages
- Update indexes in real-time data feeds
- Discover techniques to index multi-language and distributed data in Solr
- Combine the various indexing techniques into a real-life working example of an online shopping web application

In Detail

Apache Solr is a widely used, open source enterprise search server that delivers powerful indexing and searching features. These features help fetch relevant information from various sources and documentation. Solr also combines with other open source tools such as Apache Tika and Apache Nutch to provide more powerful features.

This fast-paced guide starts by helping you set up Solr and get acquainted with its basic building blocks, to give you a better understanding of Solr indexing. You'll quickly move on to indexing text and boosting the indexing time. Next, you'll focus on basic indexing techniques, various index handlers designed to modify documents, and indexing a structured data source through Data Import Handler.

Moving on, you will learn techniques to perform real-time indexing and atomic updates, as well as more advanced indexing techniques such as de-duplication. Later on, we'll help you set up a cluster of Solr servers

that combine fault tolerance and high availability. You will also gain insights into working scenarios of different aspects of Solr and how to use Solr with e-commerce data.

By the end of the book, you will be competent and confident working with indexing and will have a good knowledge base to efficiently program elements.

Style and approach

This fast-paced guide is packed with examples that are written in an easy-to-follow style, and are accompanied by detailed explanation. Working examples are included to help you get better results for your applications.

Apache Solr for Indexing Data By Sachin Handiekar, Anshul Johri Bibliography

- Rank: #1477263 in eBooks
- Published on: 2015-12-28
- Released on: 2015-12-28
- Format: Kindle eBook

Download Apache Solr for Indexing Data ...pdf

Read Online Apache Solr for Indexing Data ...pdf

Download and Read Free Online Apache Solr for Indexing Data By Sachin Handiekar, Anshul Johri

Editorial Review

About the Author

Sachin Handiekar

Sachin Handiekar is a senior software developer with over 5 years of experience in Java EE development. He graduated in computer science from the University of Greenwich, London, and currently works for a global consulting company, developing enterprise applications using various open source technologies, such as Apache Camel, ServiceMix, ActiveMQ, and ZooKeeper. He has a lot of interest in open source projects and has contributed code to Apache Camel and developed plugins for the Spring Social, which can be found on GitHub at https://github.com/sachin-handiekar. He also actively writes about enterprise application development on his blog (http://www.sachinhandiekar.com/).

Anshul Johri

Anshul Johri has more than 10 years of technical experience in software engineering. He did his masters in computer science from the computer science department in the University of Pune. Anshul has always been a start-up mindset guy, working on fast-paced development using cutting-edge technologies and doing multiple things at a time. His core strength has always been search technology, whereby Solr plays an important role in his career. Anshul started using Solr around 9 years ago, and since then, he has never looked back. He did better and better with Solr, whether using it or contributing to the open source search community. He has used Solr extensively in all his organizations across various projects. As mentioned earlier, Anshul has always been a start-up mindset guy. Because of that, he has worked with many start-ups in his career so far, which includes early-age and mid-size start-ups as well. To name a few, they are Ibibo.com, Asklaila.com, Bookadda.com, and so on. His last company was Amazon, where he spent around 2 years building scalable systems for Amazon Prime (a global product). Anshul recently started his own company in India with another friend from Amazon and founded http://www.rentomo.com/, a unique concept of a peer-to-peer sharing platform in a trusted community. He heads the technology and other core pillars of his own start-up. Anshul did the technical review of the book Indexing with Solr, published by Packt Publishing.

Users Review

From reader reviews:

Carolyn Livingston:

Often the book Apache Solr for Indexing Data will bring that you the new experience of reading some sort of book. The author style to elucidate the idea is very unique. When you try to find new book to read, this book very ideal to you. The book Apache Solr for Indexing Data is much recommended to you you just read. You can also get the e-book through the official web site, so you can quicker to read the book.

Eric Ray:

Apache Solr for Indexing Data can be one of your nice books that are good idea. We all recommend that straight away because this guide has good vocabulary that may increase your knowledge in words, easy to understand, bit entertaining but delivering the information. The writer giving his/her effort to get every word into enjoyment arrangement in writing Apache Solr for Indexing Data nevertheless doesn't forget the main place, giving the reader the hottest in addition to based confirm resource data that maybe you can be considered one of it. This great information can certainly drawn you into completely new stage of crucial imagining.

Angela Drew:

Do you really one of the book lovers? If yes, do you ever feeling doubt when you are in the book store? Make an effort to pick one book that you never know the inside because don't assess book by its deal with may doesn't work at this point is difficult job because you are afraid that the inside maybe not as fantastic as in the outside seem likes. Maybe you answer can be Apache Solr for Indexing Data why because the excellent cover that make you consider about the content will not disappoint an individual. The inside or content is definitely fantastic as the outside or cover. Your reading sixth sense will directly make suggestions to pick up this book.

Ray Davis:

As a university student exactly feel bored for you to reading. If their teacher asked them to go to the library or even make summary for some publication, they are complained. Just little students that has reading's heart and soul or real their leisure activity. They just do what the trainer want, like asked to go to the library. They go to there but nothing reading really. Any students feel that reading is not important, boring in addition to can't see colorful photographs on there. Yeah, it is for being complicated. Book is very important to suit your needs. As we know that on this age, many ways to get whatever we want. Likewise word says, many ways to reach Chinese's country. Therefore this Apache Solr for Indexing Data can make you feel more interested to read.

Download and Read Online Apache Solr for Indexing Data By Sachin Handiekar, Anshul Johri #GV7AW1H0L5C

Read Apache Solr for Indexing Data By Sachin Handiekar, Anshul Johri for online ebook

Apache Solr for Indexing Data By Sachin Handiekar, Anshul Johri Free PDF d0wnl0ad, audio books, books to read, good books to read, cheap books, good books, online books, books online, book reviews epub, read books online, books to read online, online library, greatbooks to read, PDF best books to read, top books to read Apache Solr for Indexing Data By Sachin Handiekar, Anshul Johri books to read online.

Online Apache Solr for Indexing Data By Sachin Handiekar, Anshul Johri ebook PDF download

Apache Solr for Indexing Data By Sachin Handiekar, Anshul Johri Doc

Apache Solr for Indexing Data By Sachin Handiekar, Anshul Johri Mobipocket

Apache Solr for Indexing Data By Sachin Handiekar, Anshul Johri EPub

GV7AW1H0L5C: Apache Solr for Indexing Data By Sachin Handiekar, Anshul Johri