

1. Record Nr.	UNINA9910464134903321
Autore	Kuc Rafa
Titolo	Solr cookbook : solve real-time problems related to Apache Solr 4.x and 5.0 effectively with the help of over 100 easy-to-follow recipes // Rafa Kuc
Pubbl/distr/stampa	Birmingham, England : , : Packt Publishing, , 2015 ©2015
ISBN	1-78355-316-2
Edizione	[Third edition.]
Descrizione fisica	1 online resource (356 p.)
Disciplina	005.758
Soggetti	Search engines - Programming Web search engines Electronic books.
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Includes index. "Quick Answers To Common Problems"--Cover.
Nota di contenuto	Cover; Copyright; Credits; About the Author; Acknowledgments; About the Reviewers; www.PacktPub.com; Table of Contents; Preface; Chapter 1: Apache Solr Configuration; Introduction; Running Solr on a standalone Jetty; Installing ZooKeeper for SolrCloud; Migrating configuration from master-slave to SolrCloud; Choosing the proper directory configuration; Configuring the Solr spellchecker; Using Solr in a schemaless mode; Limiting I/O usage; Using core discovery; Configuring SolrCloud for NRT use cases; Configuring SolrCloud for high-indexing use cases Configuring SolrCloud for high-querying use cases Configuring the Solr heartbeat mechanism; Changing similarity; Chapter 2: Indexing Your Data; Introduction; Indexing PDF files; Counting the number of fields; Using parsing update processors to parse data; Using scripting update processors to modify documents; Indexing data from a database using Data Import Handler; Incremental imports with DIH; Transforming data when using DIH; Indexing multiple geographical points; Updating document fields; Detecting the document language during indexation; Optimizing the primary key indexation

Handling multiple currenciesChapter 3: Analyzing Your Text Data; Introduction; Using the enumeration type; Removing HTML tags during indexing; Storing data outside of Solr index; Using synonyms; Stemming different languages; Using nonaggressive stemmers; Using the n-gram approach to do performant trailing wildcard searches; Using position increment to divide sentences; Using patterns to replace tokens; Chapter 4: Querying Solr; Introduction; Understanding and using the Lucene query language; Using position aware queries; Using boosting with autocomplete; Phrase queries with shingles Handling user queries without errorsHandling hierarchies with nested documents; Sorting data on the basis of function value; Controlling the number of terms needed to match; Affecting document score using function queries; Using simple nested queries; Using the Solr document query join functionality; Handling typos with n-grams; Rescoring query results; Chapter 5: Faceting; Introduction; Getting the number of documents with the same field value; Getting the number of documents with the same value range; Getting the number of documents matching the query and subquery Removing filters from faceting resultsUsing decision tree facetting; Calculating facetting for relevant documents in groups; Improving facetting performance for low cardinality fields; Chapter 6: Improving Solr Performance; Introduction; Handling deep paging efficiently; Configuring the document cache; Configuring the query result cache; Configuring the filter cache; Improving Solr query performance after the start and commit operations; Lowering the memory consumption of facetting and sorting; Speeding up indexing with Solr segment merge tuning Avoiding caching of rare filters to improve the performance

Sommario/riassunto

This book is for intermediate Solr Developers who are willing to learn and implement Pro-level practices, techniques, and solutions. This edition will specifically appeal to developers who wish to quickly get to grips with the changes and new features of Apache Solr 5.
