

1. Record Nr.	UNINA9911006754203321
Autore	Kuc Rafa
Titolo	Apache Solr 4 cookbook : over 100 recipes to make Apache Solr faster, more reliable, and return better results // Rafa Kuc
Pubbl/distr/stampa	Birmingham, : Packt Pub., 2013
ISBN	1-62198-883-X 1-78216-133-3 1-299-19858-9
Edizione	[2nd ed.]
Descrizione fisica	1 online resource (328 p.)
Collana	Community experience distilled
Disciplina	025.04
Soggetti	Search engines - Programming Open source software Web search engines
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Includes index.
Nota di contenuto	Cover; Copyright; Credits; About the Author; Acknowledgement; About the Reviewers; www.PacktPub.com; Table of Contents; Preface; Chapter 1: Apache Solr Configuration; Introduction; Running Solr on Jetty; Running Solr on Apache Tomcat; Installing a standalone ZooKeeper; Clustering your data; Choosing the right directory implementation; Configuring spellchecker to not use its own index; Solr cache configuration; How to fetch and index web pages; How to set up the extracting request handler; Changing the default similarity implementation; Chapter 2: Indexing Your Data; Introduction Indexing PDF filesGenerating unique fields automatically; Extracting metadata from binary files; How to properly configure Data Import Handler with JDBC; Indexing data from a database using Data Import Handler; How to import data using Data Import Handler and delta query; How to use Data Import Handler with the URL data source; How to modify data while importing with Data Import Handler; Updating a single field of your document; Handling multiple currencies; Detecting the document's language; Optimizing your primary key field indexing; Chapter 3: Analyzing Your Text Data; Introduction Storing additional information using payloadsEliminating XML and

HTML tags from text; Copying the contents of one field to another; Changing words to other words; Splitting text by CamelCase; Splitting text by whitespace only; Making plural words singular without stemming; Lowercasing the whole string; Storing geographical points in the index; Stemming your data; Preparing text to perform an efficient trailing wildcard search; Splitting text by numbers and non-whitespace characters; Using Hunspell as a stemmer; Using your own stemming dictionary; Protecting words from being stemmed

Chapter 4: Querying SolrIntroduction; Asking for a particular field value; Sorting results by a field value; How to search for a phrase, not a single word; Boosting phrases over words; Positioning some documents over others on a query; Positioning documents with words closer to each other first; Sorting results by a distance from a point; Getting documents with only a partial match; Affecting scoring with functions; Nesting queries; Modifying returned documents; Using parent-child relationships; Ignoring typos in terms of performance; Detecting and omitting duplicate documents

Using field aliasesReturning a value of a function in the results; Chapter 5: Using the Faceting Mechanism; 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 results; Sorting faceting results in alphabetical order; Implementing the autosuggest feature using faceting; Getting the number of documents that don't have a value in the field; Having two different facet limits for two different fields in the same query

Using decision tree faceting

Sommario/riassunto

Over 100 practical recipes to make Apache Solr faster, more reliable and return better results
