

1. Record Nr.	UNINA9910786778203321
Autore	Karambelkar Hrishikesh Vijay
Titolo	Scaling apache solr : optimize your searches using high-performance enterprise search repositories with apache solr / / Hrishikesh Vijay Karambelkar
Pubbl/distr/stampa	Bradford, England : , : Packt Publishing, , 2014 ©2014
ISBN	1-78398-175-X
Edizione	[1st edition]
Descrizione fisica	1 online resource (298 p.)
Collana	Community experience distilled
Disciplina	005.758
Soggetti	Search engines - Programming
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Includes index.
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	Cover; Copyright; Credits; About the Author; About the Reviewers; www.PacktPub.com; Table of Contents; Preface; Chapter 1: Understanding Apache Solr; Challenges in enterprise search; Understanding Apache Solr; Features of Apache Solr; Solr for end users; Powerful full text search; Search through rich information; Results ranking, pagination, and sorting; Facets for better browsing experience; Advanced search capabilities; Administration; Apache Solr architecture; Storage; Solr application; Integration; Client APIs and SolrJ client; Other interfaces; Practical use cases for Apache Solr Enterprise search for a job search agencyProblem statement; Approach; Enterprise search for energy industry; Problem statement; Approach; Summary; Chapter 2: Getting Started with Apache Solr; Setting up Apache Solr; Prerequisites; Running Solr on Jetty; Running Solr on Tomcat; Solr administration; What's next?; Common problems and solution; Understanding the Solr structure; The Solr home directory structure; Solr navigation; Configuring the Apache Solr for enterprise; Defining a Solr schema; Solr fields; Dynamic Fields in Solr; Copying the fields; Field types Other important elements in the Solr schemaConfiguring Solr parameters; solr.xml and Solr core; solrconfig.xml; The Solr plugin; Other configurations; Understanding SolrJ; Summary; Chapter 3: Analyzing Data with Apache Solr; Understanding enterprise data;

Categorizing by characteristics; Categorizing by access pattern; Categorizing by data formats; Loading data using native handlers; Quick and simple data loading - post tool; Working with JSON, XML, and CSV; Handling JSON data; Working with CSV data; Working with XML data; Working with rich documents; Understanding Apache Tika Using Solr Cell (ExtractingRequestHandler) Adding metadata to your rich documents; Importing structured data from the database; Configuring the data source; Importing data in Solr; Full import; Delta import; Loading RDBMS tables in Solr; Advanced topics with Solr; Deduplication; Extracting information from scanned documents; Searching through images using LIRE; Summary; Chapter 4: Designing Enterprise Search; Designing aspects for enterprise search; Identifying requirements; Matching user expectations through relevance; Access to searched entities and user interface
Improving search performance and ensuring instance scalability Working with applications through federated search; Other differentiators - mobiles, linguistic search, and security; Enterprise search data-processing patterns; Standalone search engine server; Distributed enterprise search pattern; The replicated enterprise search pattern; Distributed and replicated; Data integrating pattern for search; Data import by enterprise search; Applications pushing data; Middleware-based integration; Case study - designing an enterprise knowledge repository search for software IT services
Gathering requirements

Sommario/riassunto

This book is a step-by-step guide for readers who would like to learn how to build complete enterprise search solutions, with ample real-world examples and case studies. If you are a developer, designer, or architect who would like to build enterprise search solutions for your customers or organization, but have no prior knowledge of Apache Solr/Lucene technologies, this is the book for you.
