



Tika in Action

Chris Mattmann, Jukka Zitting

Download now

[Click here](#) if your download doesn't start automatically

Tika in Action

Chris Mattmann, Jukka Zitting

Tika in Action Chris Mattmann, Jukka Zitting

Summary

Tika in Action is a hands-on guide to content mining with Apache Tika. The book's many examples and case studies offer real-world experience from domains ranging from search engines to digital asset management and scientific data processing.

About the Technology

Tika is an Apache toolkit that has built into it everything you and your app need to know about file formats. Using Tika, your applications can discover and extract content from digital documents in almost any format, including exotic ones.

About this Book

Tika in Action is the ultimate guide to content mining using Apache Tika. You'll learn how to pull usable information from otherwise inaccessible sources, including internet media and file archives. This example-rich book teaches you to build and extend applications based on real-world experience with search engines, digital asset management, and scientific data processing. In addition to architectural overviews, you'll find detailed chapters on features like metadata extraction, automatic language detection, and custom parser development.

This book is written for developers who are new to both Scala and Lift and covers just enough Scala to get you started.

Purchase of the print book comes with an offer of a free PDF, ePub, and Kindle eBook from Manning. Also available is all code from the book.

What's Inside

- Crack MS Word, PDF, HTML, and ZIP
- Integrate with search engines, CMS, and other data sources
- Learn through experimentation
- Many examples

This book requires no previous knowledge of Tika or text mining techniques. It assumes a working knowledge of Java.

=====?

Table of Contents

1. The case for the digital Babel fish
2. Getting started with Tika
3. The information landscape PART 2 TIKA IN DETAIL
4. Document type detection

5. Content extraction
6. Understanding metadata
7. Language detection
8. What's in a file? PART 3 INTEGRATION AND ADVANCED USE
9. The big picture
10. Tika and the Lucene search stack
11. Extending Tika PART 4 CASE STUDIES
12. Powering NASA science data systems
13. Content management with Apache Jackrabbit
14. Curating cancer research data with Tika
15. The classic search engine example

 [Download Tika in Action ...pdf](#)

 [Read Online Tika in Action ...pdf](#)

Download and Read Free Online Tika in Action Chris Mattmann, Jukka Zitting

Download and Read Free Online Tika in Action Chris Mattmann, Jukka Zitting

From reader reviews:

Herman Ovalle:

Why don't make it to be your habit? Right now, try to prepare your time to do the important work, like looking for your favorite e-book and reading a e-book. Beside you can solve your long lasting problem; you can add your knowledge by the e-book entitled Tika in Action. Try to face the book Tika in Action as your good friend. It means that it can to be your friend when you sense alone and beside associated with course make you smarter than previously. Yeah, it is very fortunated for you personally. The book makes you more confidence because you can know everything by the book. So , let me make new experience in addition to knowledge with this book.

Joyce Johnson:

In this 21st century, people become competitive in every single way. By being competitive at this point, people have do something to make them survives, being in the middle of typically the crowded place and notice by surrounding. One thing that at times many people have underestimated the idea for a while is reading. Yep, by reading a book your ability to survive improve then having chance to endure than other is high. For you personally who want to start reading a book, we give you this kind of Tika in Action book as beginning and daily reading book. Why, because this book is usually more than just a book.

Lila Johnson:

The book Tika in Action has a lot info on it. So when you check out this book you can get a lot of profit. The book was written by the very famous author. Mcdougal makes some research previous to write this book. This kind of book very easy to read you will get the point easily after reading this book.

John Rivera:

What is your hobby? Have you heard in which question when you got pupils? We believe that that problem was given by teacher to their students. Many kinds of hobby, Every individual has different hobby. So you know that little person just like reading or as looking at become their hobby. You need to know that reading is very important as well as book as to be the issue. Book is important thing to incorporate you knowledge, except your own teacher or lecturer. You will find good news or update regarding something by book. A substantial number of sorts of books that can you go onto be your object. One of them is actually Tika in Action.

Download and Read Online Tika in Action Chris Mattmann, Jukka

Zitting #VQFA9IYEC4S

Read Tika in Action by Chris Mattmann, Jukka Zitting for online ebook

Tika in Action by Chris Mattmann, Jukka Zitting 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 Tika in Action by Chris Mattmann, Jukka Zitting books to read online.

Online Tika in Action by Chris Mattmann, Jukka Zitting ebook PDF download

Tika in Action by Chris Mattmann, Jukka Zitting Doc

Tika in Action by Chris Mattmann, Jukka Zitting Mobipocket

Tika in Action by Chris Mattmann, Jukka Zitting EPub