



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

From reader reviews:

Christopher Clarke:

The event that you get from Tika in Action may be the more deep you digging the information that hide within the words the more you get interested in reading it. It does not mean that this book is hard to comprehend but Tika in Action giving you excitement feeling of reading. The article writer conveys their point in specific way that can be understood by simply anyone who read the idea because the author of this publication is well-known enough. This book also makes your personal vocabulary increase well. Making it easy to understand then can go with you, both in printed or e-book style are available. We advise you for having this kind of Tika in Action instantly.

Detra Satterwhite:

The book untitled Tika in Action is the publication that recommended to you to read. You can see the quality of the reserve content that will be shown to an individual. The language that creator use to explained their ideas are easily to understand. The article author was did a lot of exploration when write the book, therefore the information that they share for your requirements is absolutely accurate. You also might get the e-book of Tika in Action from the publisher to make you far more enjoy free time.

Jennifer Ruiz:

This Tika in Action is great book for you because the content that is certainly full of information for you who also always deal with world and have to make decision every minute. This book reveal it information accurately using great coordinate word or we can state no rambling sentences inside it. So if you are read the item hurriedly you can have whole information in it. Doesn't mean it only gives you straight forward sentences but difficult core information with wonderful delivering sentences. Having Tika in Action in your hand like finding the world in your arm, info in it is not ridiculous one particular. We can say that no book that offer you world inside ten or fifteen tiny right but this book already do that. So , it is good reading book. Heya Mr. and Mrs. active do you still doubt that?

Dwight Hancock:

As we know that book is vital thing to add our information for everything. By a guide we can know everything you want. A book is a list of written, printed, illustrated or blank sheet. Every year has been exactly added. This guide Tika in Action was filled about science. Spend your time to add your knowledge about your scientific research competence. Some people has several feel when they reading a book. If you know how big selling point of a book, you can sense enjoy to read a guide. In the modern era like right now, many ways to get book you wanted.

**Download and Read Online Tika in Action Chris Mattmann, Jukka
Zitting #FTY4072VXCA**

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