



# Taming Text: How to Find, Organize, and Manipulate It

By Grant S. Ingersoll, Thomas S. Morton, Drew Farris

Download now

Read Online ➔

**Taming Text: How to Find, Organize, and Manipulate It** By Grant S. Ingersoll, Thomas S. Morton, Drew Farris

## Summary

*Taming Text*, winner of the 2013 Jolt Awards for Productivity, is a hands-on, example-driven guide to working with unstructured text in the context of real-world applications. This book explores how to automatically organize text using approaches such as full-text search, proper name recognition, clustering, tagging, information extraction, and summarization. The book guides you through examples illustrating each of these topics, as well as the foundations upon which they are built.

## About this Book

There is so much text in our lives, we are practically drowning in it. Fortunately, there are innovative tools and techniques for managing unstructured information that can throw the smart developer a much-needed lifeline. You'll find them in this book.

*Taming Text* is a practical, example-driven guide to working with text in real applications. This book introduces you to useful techniques like full-text search, proper name recognition, clustering, tagging, information extraction, and summarization. You'll explore real use cases as you systematically absorb the foundations upon which they are built. Written in a clear and concise style, this book avoids jargon, explaining the subject in terms you can understand without a background in statistics or natural language processing. Examples are in Java, but the concepts can be applied in any language.

Written for Java developers, the book requires no prior knowledge of GWT.

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.

Winner of 2013 Jolt Awards: The Best Books—one of five notable books every serious programmer should read.

## What's Inside

- When to use text-taming techniques
- Important open-source libraries like Solr and Mahout
- How to build text-processing applications

## About the Authors

**Grant Ingersoll** is an engineer, speaker, and trainer, a Lucenecommitter, and a cofounder of the Mahout machine-learning project. **Thomas Morton** is the primary developer of OpenNLP and Maximum Entropy. **Drew Farris** is a technology consultant, software developer, and contributor to Mahout, Lucene, and Solr.

*"Takes the mystery out of very complex processes."*—From the Foreword by Liz Liddy, Dean, iSchool, Syracuse University

## Table of Contents

1. Getting started taming text
2. Foundations of taming text
3. Searching
4. Fuzzy string matching
5. Identifying people, places, and things
6. Clustering text
7. Classification, categorization, and tagging
8. Building an example question answering system
9. Untamed text: exploring the next frontier

 [Download Taming Text: How to Find, Organize, and Manipulate ...pdf](#)

 [Read Online Taming Text: How to Find, Organize, and Manipula ...pdf](#)

# Taming Text: How to Find, Organize, and Manipulate It

*By Grant S. Ingersoll, Thomas S. Morton, Drew Farris*

**Taming Text: How to Find, Organize, and Manipulate It** By Grant S. Ingersoll, Thomas S. Morton, Drew Farris

## Summary

*Taming Text*, winner of the 2013 Jolt Awards for Productivity, is a hands-on, example-driven guide to working with unstructured text in the context of real-world applications. This book explores how to automatically organize text using approaches such as full-text search, proper name recognition, clustering, tagging, information extraction, and summarization. The book guides you through examples illustrating each of these topics, as well as the foundations upon which they are built.

## About this Book

There is so much text in our lives, we are practically drowning in it. Fortunately, there are innovative tools and techniques for managing unstructured information that can throw the smart developer a much-needed lifeline. You'll find them in this book.

*Taming Text* is a practical, example-driven guide to working with text in real applications. This book introduces you to useful techniques like full-text search, proper name recognition, clustering, tagging, information extraction, and summarization. You'll explore real use cases as you systematically absorb the foundations upon which they are built. Written in a clear and concise style, this book avoids jargon, explaining the subject in terms you can understand without a background in statistics or natural language processing. Examples are in Java, but the concepts can be applied in any language.

Written for Java developers, the book requires no prior knowledge of GWT.

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.

Winner of 2013 Jolt Awards: The Best Books—one of five notable books every serious programmer should read.

## What's Inside

- When to use text-taming techniques
- Important open-source libraries like Solr and Mahout
- How to build text-processing applications

## About the Authors

**Grant Ingersoll** is an engineer, speaker, and trainer, a Lucene committer, and a cofounder of the Mahout machine-learning project. **Thomas Morton** is the primary developer of OpenNLP and Maximum Entropy. **Drew Farris** is a technology consultant, software developer, and contributor to Mahout, Lucene, and Solr.

*"Takes the mystery out of very complex processes."*—From the Foreword by Liz Liddy, Dean, iSchool,

## Table of Contents

1. Getting started taming text
2. Foundations of taming text
3. Searching
4. Fuzzy string matching
5. Identifying people, places, and things
6. Clustering text
7. Classification, categorization, and tagging
8. Building an example question answering system
9. Untamed text: exploring the next frontier

## **Taming Text: How to Find, Organize, and Manipulate It By Grant S. Ingersoll, Thomas S. Morton, Drew Farris Bibliography**

- Sales Rank: #268407 in Books
- Brand: Brand: Manning Publications
- Published on: 2013-01-24
- Original language: English
- Number of items: 1
- Dimensions: 9.25" h x .69" w x 7.38" l, 1.18 pounds
- Binding: Paperback
- 320 pages

 [Download Taming Text: How to Find, Organize, and Manipulate ...pdf](#)

 [Read Online Taming Text: How to Find, Organize, and Manipula ...pdf](#)

## **Editorial Review**

### **Users Review**

#### **From reader reviews:**

##### **Bobby House:**

What do you think of book? It is just for students since they're still students or the idea for all people in the world, what the best subject for that? Just simply you can be answered for that issue above. Every person has several personality and hobby for every single other. Don't to be pressured someone or something that they don't need do that. You must know how great and important the book Taming Text: How to Find, Organize, and Manipulate It. All type of book is it possible to see on many options. You can look for the internet resources or other social media.

##### **Jennifer Tomasini:**

Taming Text: How to Find, Organize, and Manipulate It can be one of your basic books that are good idea. All of us recommend that straight away because this guide has good vocabulary which could increase your knowledge in words, easy to understand, bit entertaining but delivering the information. The author giving his/her effort to set every word into joy arrangement in writing Taming Text: How to Find, Organize, and Manipulate It however doesn't forget the main level, giving the reader the hottest and based confirm resource details that maybe you can be certainly one of it. This great information may drawn you into brand new stage of crucial imagining.

##### **Randal Revilla:**

Beside this kind of Taming Text: How to Find, Organize, and Manipulate It in your phone, it could possibly give you a way to get closer to the new knowledge or details. The information and the knowledge you can got here is fresh in the oven so don't be worry if you feel like an previous people live in narrow town. It is good thing to have Taming Text: How to Find, Organize, and Manipulate It because this book offers for you readable information. Do you occasionally have book but you don't get what it's facts concerning. Oh come on, that would not happen if you have this in your hand. The Enjoyable option here cannot be questionable, such as treasuring beautiful island. Use you still want to miss this? Find this book and read it from currently!

##### **Chrissy Stallings:**

Is it you actually who having spare time and then spend it whole day by simply watching television programs or just resting on the bed? Do you need something totally new? This Taming Text: How to Find, Organize, and Manipulate It can be the response, oh how comes? The new book you know. You are consequently out of date, spending your free time by reading in this brand new era is common not a nerd activity. So what

these textbooks have than the others?

**Download and Read Online Taming Text: How to Find, Organize,  
and Manipulate It By Grant S. Ingersoll, Thomas S. Morton, Drew  
Farris #IOD7XPMY1JR**

# **Read Taming Text: How to Find, Organize, and Manipulate It By Grant S. Ingersoll, Thomas S. Morton, Drew Farris for online ebook**

Taming Text: How to Find, Organize, and Manipulate It By Grant S. Ingersoll, Thomas S. Morton, Drew Farris 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 Taming Text: How to Find, Organize, and Manipulate It By Grant S. Ingersoll, Thomas S. Morton, Drew Farris books to read online.

## **Online Taming Text: How to Find, Organize, and Manipulate It By Grant S. Ingersoll, Thomas S. Morton, Drew Farris ebook PDF download**

**Taming Text: How to Find, Organize, and Manipulate It By Grant S. Ingersoll, Thomas S. Morton, Drew Farris Doc**

**Taming Text: How to Find, Organize, and Manipulate It By Grant S. Ingersoll, Thomas S. Morton, Drew Farris Mobipocket**

**Taming Text: How to Find, Organize, and Manipulate It By Grant S. Ingersoll, Thomas S. Morton, Drew Farris EPub**