



Taming Text

Grant S. Ingersoll , Thomas S. Morton , Andrew L. Farris

[Download now](#)

[Read Online](#) 

Taming Text

Grant S. Ingersoll , Thomas S. Morton , Andrew L. Farris

Taming Text Grant S. Ingersoll , Thomas S. Morton , Andrew L. Farris

It is no secret that the world is drowning in text and data. This causes real problems for everyday users who need to make sense of all the information available, and for software engineers who want to make their text-based applications more useful and user-friendly. Whether building a search engine for a corporate website, automatically organizing email, or extracting important nuggets of information from the news, dealing with unstructured text can be daunting.

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

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.

Taming Text Details

Date : Published January 21st 2013 by Manning Publications (first published July 2011)

ISBN : 9781933988382

Author : Grant S. Ingersoll , Thomas S. Morton , Andrew L. Farris

Format : Paperback 322 pages

Genre : Computer Science, Programming, Science, Technology, Reference, Computers

 [Download Taming Text ...pdf](#)

 [Read Online Taming Text ...pdf](#)

Download and Read Free Online Taming Text Grant S. Ingersoll , Thomas S. Morton , Andrew L. Farris

From Reader Review Taming Text for online ebook

Zac says

This book was alright. Many of the theoretical bits are things I've already encountered. The practical bits weren't so relevant because they were too tightly coupled to specific Java libraries and other pieces of software. Nothing too earth shattering here, but might be just right for some readers.

Ashish says

This book is like a hands on guide to Text Analytics and Processing as it talks about the Open Source projects related to this topic. Other than that it is a very good introduction to basics of Text Analysis and how one can use Open Source Solutions like Lucene, Solr, OpenNLP etc to do the same.

The last chapter about Untamed text: the next frontier is very good. For someone interested in Text Processing and Analytics that last chapter is a very good read with a lots of ideas regarding what could be done next in this field.

Vuk Trifkovic says

Bit scattergun, a bit basic, but sound introduction to text handling.

Deane Barker says

The book alternates between a great overview of the subject, and getting down-and-dirty with the code. I don't know a good solution to this, but I was less concerned with the code and more concerned with the overhead view.

It's a good discussion of how to manage text: how to tokenize it, search it, cluster it, and classify it. Towards the later chapters, it bogs way, way down -- Chapter 7, on classification, is probably a quarter of the entire book in length. In many cases they "dropped to code" immediately without discussion the theory much.

In the end, I got what I came for -- I understand the basic concepts of text processing. I now know what I don't know, and that's often half the battle.

Alex Ott says

Good overview of different systems (Solr, OpenNLP & Mahout), approaches & algorithms for working with (unstructured) text - search, analyze, cluster, etc.

The book itself is completely practical with references to articles & books for people interested in more detailed/theoretical information.

Dgg32 says

A nice introduction to the field natural language programming. May come in handy for some of my projects.

Akshay Ratan says

Good for understanding content analysis using efficient techniques. Apache Solr and Lucene explained with concepts and codes.
