

Manning Taming Text

Summary Hadoop in Practice, Second Edition provides over 100 tested, instantly useful techniques that will help you conquer big data, using Hadoop. This revised new edition covers changes and new features in the Hadoop core architecture, including MapReduce 2. Brand new chapters cover YARN and integrating Kafka, Impala, and Spark SQL with Hadoop. You'll also get new and updated techniques for Flume, Sqoop, and Mahout, all of which have seen major new versions recently. In short, this is the most practical, up-to-date coverage of Hadoop available anywhere. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Book It's always a good time to upgrade your Hadoop skills! Hadoop in Practice, Second Edition provides a collection of 104 tested, instantly useful techniques for analyzing real-time streams, moving data securely, machine learning, managing large-scale clusters, and taming big data using Hadoop. This completely revised edition covers changes and new features in Hadoop core, including MapReduce 2 and YARN. You'll pick up hands-on best practices for integrating Spark, Kafka, and Impala with Hadoop, and get new and updated techniques for the latest versions of Flume, Sqoop, and Mahout. In short, this is the most practical, up-to-date coverage of Hadoop available. Readers need to know a programming language like Java and have basic familiarity with Hadoop. What's Inside Thoroughly updated for Hadoop 2 How to write YARN applications Integrate real-time technologies like Storm, Impala, and Spark Predictive analytics using Mahout and RR Readers need to know a programming language like Java and have basic familiarity with Hadoop. About the Author Alex Holmes works on tough big-data problems. He is a software engineer, author, speaker, and blogger specializing in large-scale Hadoop projects. Table of Contents PART 1 BACKGROUND AND FUNDAMENTALS Hadoop in a heartbeat Introduction to YARN PART 2 DATA LOGISTICS Data serialization—working with text and beyond Organizing and optimizing data in HDFS Moving data into and out of Hadoop PART 3 BIG DATA PATTERNS Applying MapReduce patterns to big data Utilizing data structures and algorithms at scale Tuning, debugging, and testing PART 4 BEYOND MAPREDUCE SQL on Hadoop Writing a YARN application

This book constitutes the refereed proceedings of the 8th International Workshop on Learning Technology for Education Challenges, LTEC 2019, held in Zamora, Spain, in July 2019. The 41 revised full papers presented were carefully reviewed and selected from 83 submissions. The papers are organized in the following topical sections: learning technologies; learning tools and environment; e-learning and MOOCs; learning practices; social media learning tools; machine learning and evaluation support programs. LTEC 2019 examines how these technologies and pedagogical advances can be used to change the way teachers teach and students learn, while giving special emphasis to the pedagogically effective ways we can harness these new technologies in education.

This book constitutes the thoroughly refereed post-conference proceedings of the First International Workshop on Personal Analytics and Privacy, PAP 2017, held in Skopje, Macedonia, in September 2017. The 14 papers presented together with 2 invited talks in this volume were carefully reviewed and selected for inclusion in this book and handle topics such as personal analytics, personal data mining and privacy in the context where real individual data are used for developing a data-driven service, for realizing a social study aimed at understanding nowadays society, and for publication purposes.

A comprehensive guide to using the web application, including such topics as text analysis, faceted search, result grouping, multilingual search, advanced geospatial and data operations, and relevancy tuning.

Master predictive analytics, from start to finish Start with strategy and management Master methods and build models Transform your models into highly-effective code—in both Python and R This one-of-a-kind book will help you use predictive analytics, Python, and R to solve real business problems and drive real competitive advantage. You ' ll master predictive analytics through realistic case studies, intuitive data visualizations, and up-to-date code for both Python and R—not complex math. Step by step, you ' ll walk through defining problems, identifying data, crafting and optimizing models, writing effective Python and R code, interpreting results, and more. Each chapter focuses on one of today ' s key applications for predictive analytics, delivering skills and knowledge to put models to work—and maximize their value. Thomas W. Miller, leader of Northwestern University ' s pioneering program in predictive analytics, addresses everything you need to succeed: strategy and management, methods and models, and technology and code. If you ' re new to predictive analytics, you ' ll gain a strong foundation for achieving accurate, actionable results. If you ' re already working in the field, you ' ll master powerful new skills. If you ' re familiar with either Python or R, you ' ll discover how these languages complement each other, enabling you to do even more. All data sets, extensive Python and R code, and additional examples available for download at <http://www.ftpress.com/miller/> Python and R offer immense power in predictive analytics, data science, and big data. This book will help you leverage that power to solve real business problems, and drive real competitive advantage. Thomas W. Miller ' s unique balanced approach combines business context and quantitative tools, illuminating each technique with carefully explained code for the latest versions of Python and R. If you ' re new to predictive analytics, Miller gives you a strong foundation for achieving accurate, actionable results. If you ' re already a modeler, programmer, or manager, you ' ll learn crucial skills you don ' t already have. Using Python and R, Miller addresses multiple business challenges, including segmentation, brand positioning, product choice modeling, pricing research, finance, sports, text analytics, sentiment analysis, and social network analysis. He illuminates the use of cross-sectional data, time series, spatial, and spatio-temporal data. You ' ll learn why each problem matters, what data are relevant, and how to explore the data you ' ve identified. Miller guides you through conceptually modeling each data set with words and figures; and then modeling it again with realistic code that delivers actionable insights. You ' ll walk through model construction, explanatory variable subset selection, and validation, mastering best practices for improving out-of-sample predictive performance. Miller employs data visualization and statistical graphics to help you explore data, present models, and evaluate performance. Appendices include five complete case studies, and a detailed primer on modern data science methods. Use Python and R to gain powerful, actionable, profitable insights about: Advertising and promotion Consumer preference and choice Market

baskets and related purchases Economic forecasting Operations management Unstructured text and language Customer sentiment Brand and price Sports team performance And much more Over the past few years, e-government has been rapidly changing the way governmental services are provided to citizens and businesses. These services improve business and government exchange capability, provide a new way to discover and share information, and play a part in the evolution of future technologies. The Handbook of Research on Democratic Strategies and Citizen-Centered E-Government Services seeks to address which services in e-government should be provided to users and how. This premier reference work gives an overview of the latest achievements in the field of e-government services, provides in-depth analysis of and research on the development and deployment of cutting-edge applications, and provides insight into future trends for researchers, teachers, students, government workers, and IT professionals.

Pijnlijk nauwkeurig beschrijft een dierenliefhebber de mentale strijd die hij voert met zijn roofvogel. Hij doet dat in een sterk beeldende taal, vol hilarische hersenkronkels. Dit dagboek, want dat is het eigenlijk, verslaat een tot mislukken gedoemde queeste: een man is vergeefs op zoek naar een verondersteld verleden waarin mens en dier nog é é n waren. De Britse auteur Terence White (1906-1964) werd beroemd met zijn roman over koning Arthur en de ridders van de Ronde Tafel, The Once and Future King (1958). Maar daarv ó ó r al schreef hij dit meesterwerkje over zijn pogingen om een wilde havik te temmen. Directe aanleiding voor zijn onderneming is de depressie waarin de naderende Tweede Wereldoorlog hem heeft gedompeld. White zoekt zijn heil in een romantisch kluizenaarsbestaan op het Engelse platteland, maar dat ontardt in een ware uitputtingslag. De zeventiende-eeuwse valkerijhandboeken die hij raadpleegt zijn hem nauwelijks tot steun. Veel van zijn ervaringen met havik Gos heeft White later verwerkt in The Once and Future King. Voor J.K. Rowling, auteur van de Harry Potterboeken, was Whites oeuvre een grote inspiratiebron. De jonge Arthur noemde zij zelfs ‘ de spirituele voorvader ’ van Harry. Het personage Arthur werd op zijn beurt geboren in die kluizenaarshut te Buckinghamshire, als alter ego van de zoekende White zelf.

Language—that is, oral or written content that references abstract concepts in subtle ways—is what sets us apart as a species, and in an age defined by such content, language has become both the fuel and the currency of our modern information society. This has posed a vexing new challenge for linguists and engineers working in the field of language-processing: how do we parse and process not just language itself, but language in vast, overwhelming quantities? Modern Computational Models of Semantic Discovery in Natural Language compiles and reviews the most prominent linguistic theories into a single source that serves as an essential reference for future solutions to one of the most important challenges of our age. This comprehensive publication benefits an audience of students and professionals, researchers, and practitioners of linguistics and language discovery. This book includes a comprehensive range of topics and chapters covering digital media, social interaction in online environments, text and data mining, language processing and translation, and contextual documentation, among others.

[Nooit Meer Mister Nice Guy](#)

[Modern Computational Models of Semantic Discovery in Natural Language](#)

[Understanding, analyzing, and generating text with Python](#)

[First International Workshop, PAP 2017, Held in Conjunction with ECML PKDD 2017, Skopje, Macedonia, September 18, 2017, Revised Selected Papers](#)

[Business Problems and Solutions with R, Revised and Expanded Edition](#)

[Artificial Intelligence for Customer Relationship Management](#)

[Interpreting Text and Unstructured Data for Business Intelligence](#)

[Modeling Techniques in Predictive Analytics](#)

[Modeling Techniques in Predictive Analytics with Python and R](#)

[Maestra](#)

[With Examples Using Elasticsearch and Solr](#)

AI

NLP

()

Keras TensorFlow

NLP()

NLP

NLP

1853. Caroline Fletcher groeit op in Richmond, in het zuiden van de Verenigde Staten. Ze is jong, blank en rijk, en wordt dag en nacht omringd door zwarte slaven die haar verzorgen en bedienen. Voor Caroline is dat de normaalste zaak van de wereld. Dat verandert wanneer ze beseft hoe wreed en onrechtvaardig slaven behandeld worden. Ze gaat zich inzetten voor afschaffing van de slavernij. Wat ze nooit had voorzien, gebeurt: ze komt tegenover haar familie, vrienden, en zelfs haar verloofde te staan. Dan breekt de burgeroorlog uit tussen de zuidelijke en de noordelijke staten, waar de slavernij al is afgeschaft. Caroline wordt verscheurd door tweestrijd. Ze is loyaal aan haar land, maar ze weet ook dat er in het zuiden alleen een einde zal komen aan de slavernij als het noorden de oorlog wint. Voor alles wil ze doen wat God van haar verlangt. Maar moet ze dan haar land en liefde verraden? Is haar overtuiging zo'n hoge prijs waard?

Social media is an invaluable source of time-critical information during a crisis. However, emergency response and humanitarian relief organizations that would like to use this information

struggle with an avalanche of social media messages that exceeds human capacity to process. Emergency managers, decision makers, and affected communities can make sense of social media through a combination of machine computation and human compassion - expressed by thousands of digital volunteers who publish, process, and summarize potentially life-saving information. This book brings together computational methods from many disciplines: natural language processing, semantic technologies, data mining, machine learning, network analysis, human-computer interaction, and information visualization, focusing on methods that are commonly used for processing social media messages under time-critical constraints, and offering more than 500 references to in-depth information.

op een septemberavond in 1785 krijgt koopman Hancock urgent bezoek van een van zijn kapiteins: hij heeft diens schip verkocht in ruil voor een meermin, dat onweerstaanbaar mysterieuze wezen van de zee. Geruchten verspreiden zich razendsnel door de haven, koffi ehuizen, salons en bordelen: iedereen wil het wonder aanschouwen! Door de meermin dringt Hancock zelfs door tot de fin de siècle van Londen en op een opulent feest ontmoet hij Angelica Neal, oh, de meest begerenswaardige aller vrouwen... en een zeer vaardige courtisane. Hun onwaarschijnlijke liaison stuurt hun leven een riskante kant op. In deze verrukkelijke, uiterst geslaagde debuutroman met veel oog voor schitterende details onderzoekt Imogen Hermes Gowar wat klasse, familie en vrouwzijn betekende in het Londen van de achttiende eeuw.

This handbook offers a thorough treatment of the science of linguistic annotation. Leaders in the field guide the reader through the process of modeling, creating an annotation language, building a corpus and evaluating it for correctness. Essential reading for both computer scientists and linguistic researchers. Linguistic annotation is an increasingly important activity in the field of computational linguistics because of its critical role in the development of language models for natural language processing applications. Part one of this book covers all phases of the linguistic annotation process, from annotation scheme design and choice of representation format through both the manual and automatic annotation process, evaluation, and iterative improvement of annotation accuracy. The second part of the book includes case studies of annotation projects across the spectrum of linguistic annotation types, including morpho-syntactic tagging, syntactic analyses, a range of semantic analyses (semantic roles, named entities, sentiment and opinion), time and event and spatial analyses, and discourse level analyses including discourse structure, co-reference, etc. Each case study addresses the various phases and processes discussed in the chapters of part one.

In 1930 verscheen in Engeland Her privates we van Frederic Manning, die zich presenteerde onder het pseudoniem Soldaat 19022. De critici, onder wie grote namen als Ezra Pound, Ernest Hemingway, T.E. Lawrence en E.M. Forster, waren lovend over dit boek dat het dagelijkse leven achter de frontlinies van de oorlog van '14-'18 beschrijft. Hoofdpersoon is soldaat Bourne. Hij deelt het lot van honderdduizenden die ploeterden in hitte en koude, die werden uitgestuurd op onmogelijke missies en vaak niet terugkeerden. Bourne is Mannings alter ego; hij deelt zijn voorkeur voor kortgeknipt haar, zijn liefde voor paarden en zijn verslaving aan sigaret en alcohol. Manning volgt Bourne en zijn kameraden, ze trekken steeds dichterbij elkaar toe in de wetenschap dat ze eenmaal 'over de top', in het niemandsland tussen de linies, uiteindelijk moederziel alleen zullen zijn. 'Het mooiste en indrukwekkendste boek dat ik ooit heb gelezen over mannen in oorlog.' - Ernest Hemingway

A concise, hands-on guide with many practical examples and a detailed treatise on inference and social science research that will help you in mining data in the real world. Whether you are an undergraduate who wishes to get hands-on experience working with social data from the Web, a practitioner wishing to expand your competencies and learn unsupervised sentiment analysis, or you are simply interested in social data analysis, this book will prove to be an essential asset. No previous experience with R or statistics is required, though having knowledge of both will enrich your experience.

Judith Rashleigh werkt overdag als assistent in een gerenommeerd Londens veilinghuis en om rond te kunnen komen is ze 's nachts hostess in een chique club in de West End-buurt. Als ze een vervalsing ontdekt, wordt ze bij het veilinghuis ontslagen en vlucht ze met een welwillende en rijke klant van de nachtclub naar de Franse Riviera. Als ze daar, zich lavend tussen de rijken and famous, op de fraudezaak stuit die haar uit Londen heeft doen vluchten, zint ze op zoete wraak...

[Data Science for Economics and Finance](#)

[The Answer Machine](#)

[R: Mining spatial, text, web, and social media data](#)

[Handbook of Research on Big Data Storage and Visualization Techniques](#)

[De meermin en de courtisane](#)

[Solr in Action](#)

[Social Media Mining with R](#)

[Text Mining and Visualization](#)

[The Diabolic](#)

[Taming complex software with functional thinking](#)

[Big Data Analytics](#)

[Handbook of Research on Democratic Strategies and Citizen-Centered E-Government Services](#)

'Als wij schurken waren' van M. L. Rio vertelt het verhaal van Oliver Marks. Tien jaar heeft hij in de gevangenis gezeten - voor een moord die hij wel of niet heeft gepleegd. Dát is wat detective Colborne zo graag nog van hem zou horen nu hij met pensioen gaat. Tien jaar eerder: zeven jonge, briljante en gepassioneerde studenten op een arts college, spelen en bestuderen Shakespeare's werk. Ze zijn vol van Shakespeare, vol van zijn personages, en zetten hun rol ook buiten het toneel voort. Totdat de vaste

rolverdeling volledig wordt omgegooid en het verbroken evenwicht zijn consequenties krijgt in het echte leven. Een van hen wordt dood aangetroffen, de andere zes zien zich voor de uitdaging gesteld de politie, en met name zichzelf, te overtuigen van hun onschuld.

Nemesis lijkt een doodgewoon meisje, maar niets is minder waar. Ze is een Diabolic, gemaakt om te doden. Op jonge leeftijd wordt ze gekocht door een adellijke familie om hun dochter Sydonia te beschermen - zo nodig met haar leven. Als Sydonia zich moet melden aan het hof van de keizer, die duidelijk kwaad in de zin heeft, neemt Nemesis haar plaats in. Niemand mag ontdekken wie en wat ze werkelijk is. Maar dan ontmoet ze Tyrus, de toekomstige opvolger van de keizer. Is het wel waar wat haar altijd is verteld? Hebben Diabolics dan toch gevoel, ergens diep vanbinnen?

The Answer Machine is a practical, non-technical guide to the technologies behind information seeking and analysis. It introduces search and content analytics to software buyers, knowledge managers, and searchers who want to understand and design effective online environments. The book describes how search evolved from an expert-only to an end user tool. It provides an overview of search engines, categorization and clustering, natural language processing, content analytics, and visualization technologies. Detailed profiles for Web search, eCommerce search, eDiscovery, and enterprise search contrast the types of users, uses, tasks, technologies, and interaction designs for each. These variables shape each application, although the underlying technologies are the same. Types of information tasks and the trade-offs between precision and recall, time, volume and precision, and privacy vs. personalization are discussed within this context. The book examines trends toward convenient, context-aware computing, big data and analytics technologies, conversational systems, and answer machines. The Answer Machine explores IBM Watson's DeepQA technology and describes how it is used to answer health care and Jeopardy questions. The book concludes by discussing the implications of these advances: how they will change the way we run our businesses, practice medicine, govern, or conduct our lives in the digital age. Table of Contents: Introduction / The Query Process and Barriers to Finding Information Online / Online Search: An Evolution / Search and Discovery Technologies: An Overview / Information Access: A Spectrum of Needs and Uses / Future Tense: The Next Era in Information Access and Discovery / Answer Machines

To succeed with predictive analytics, you must understand it on three levels: Strategy and management Methods and models Technology and code This up-to-the-minute reference thoroughly covers all three categories. Now fully updated, this uniquely accessible book will help you use predictive analytics to solve real business problems and drive real competitive advantage. If you're new to the discipline, it will give you the strong foundation you need to get accurate, actionable results. If you're already a modeler, programmer, or manager, it will teach you crucial skills you don't yet have. Unlike competitive books, this guide illuminates the discipline through realistic vignettes and intuitive data visualizations-not complex math. Thomas W. Miller, leader of Northwestern University's pioneering program in predictive analytics, guides you through defining problems, identifying data, crafting and optimizing models, writing effective R code, interpreting results, and more. Every chapter focuses on one of today's key applications for predictive analytics, delivering skills and knowledge to put models to work-and maximize their value. Reflecting extensive student and instructor feedback, this edition adds five classroom-tested case studies, updates all code for new versions of R, explains code behavior more clearly and completely, and covers modern data science methods even more effectively. All data sets, extensive R code, and additional examples available for download at <http://www.ftpress.com/miller> If you want to make the most of predictive analytics, data science, and big data, this is the book for you. Thomas W. Miller's unique balanced approach combines business context and quantitative tools, appealing to managers, analysts, programmers, and students alike. Miller addresses multiple business cases and challenges, including segmentation, brand positioning, product choice modeling, pricing research, finance, sports, text analytics, sentiment analysis, and social network analysis. He illuminates the use of cross-sectional data, time series, spatial, and spatio-temporal data. You'll learn why each problem matters, what data are relevant, and how to explore the data you've identified. Miller guides you through conceptually modeling each data set with words and figures; and then modeling it again with realistic R programs that deliver actionable insights. You'll walk through model construction, explanatory variable subset selection, and validation, mastering best practices for improving out-of-sample predictive performance. Throughout, Miller employs data visualization and statistical graphics to help you explore data, present

models, and evaluate performance. This edition adds five new case studies, updates all code for the newest versions of R, adds more commenting to clarify how the code works, and offers a more detailed and up-to-date primer on data science methods. Gain powerful, actionable, profitable insights about: Advertising and promotion Consumer preference and choice Market baskets and related purchases Economic forecasting Operations management Unstructured text and language Customer sentiment Brand and price Sports team performance And much more

Summary Natural Language Processing in Action is your guide to creating machines that understand human language using the power of Python with its ecosystem of packages dedicated to NLP and AI. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology Recent advances in deep learning empower applications to understand text and speech with extreme accuracy. The result? Chatbots that can imitate real people, meaningful resume-to-job matches, superb predictive search, and automatically generated document summaries—all at a low cost. New techniques, along with accessible tools like Keras and TensorFlow, make professional-quality NLP easier than ever before. About the Book Natural Language Processing in Action is your guide to building machines that can read and interpret human language. In it, you'll use readily available Python packages to capture the meaning in text and react accordingly. The book expands traditional NLP approaches to include neural networks, modern deep learning algorithms, and generative techniques as you tackle real-world problems like extracting dates and names, composing text, and answering free-form questions. What's inside Some sentences in this book were written by NLP! Can you guess which ones? Working with Keras, TensorFlow, gensim, and scikit-learn Rule-based and data-based NLP Scalable pipelines About the Reader This book requires a basic understanding of deep learning and intermediate Python skills. About the Author Hobson Lane, Cole Howard, and Hannes Max Hapke are experienced NLP engineers who use these techniques in production. Table of Contents PART 1 - WORDY MACHINES Packets of thought (NLP overview) Build your vocabulary (word tokenization) Math with words (TF-IDF vectors) Finding meaning in word counts (semantic analysis) PART 2 - DEEPER LEARNING (NEURAL NETWORKS) Baby steps with neural networks (perceptrons and backpropagation) Reasoning with word vectors (Word2vec) Getting words in order with convolutional neural networks (CNNs) Loopy (recurrent) neural networks (RNNs) Improving retention with long short-term memory networks Sequence-to-sequence models and attention PART 3 - GETTING REAL (REAL-WORLD NLP CHALLENGES) Information extraction (named entity extraction and question answering) Getting chatty (dialog engines) Scaling up (optimization, parallelization, and batch processing)

Computational Analysis and Understanding of Natural Languages: Principles, Methods and Applications, Volume 38, the latest release in this monograph that provides a cohesive and integrated exposition of these advances and associated applications, includes new chapters on Linguistics: Core Concepts and Principles, Grammars, Open-Source Libraries, Application Frameworks, Workflow Systems, Mathematical Essentials, Probability, Inference and Prediction Methods, Random Processes, Bayesian Methods, Machine Learning, Artificial Neural Networks for Natural Language Processing, Information Retrieval, Language Core Tasks, Language Understanding Applications, and more. The synergistic confluence of linguistics, statistics, big data, and high-performance computing is the underlying force for the recent and dramatic advances in analyzing and understanding natural languages, hence making this series all the more important. Provides a thorough treatment of open-source libraries, application frameworks and workflow systems for natural language analysis and understanding Presents new chapters on Linguistics: Core Concepts and Principles, Grammars, Open-Source Libraries, Application Frameworks, Workflow Systems, Mathematical Essentials, Probability, and more

The digital age has presented an exponential growth in the amount of data available to individuals looking to draw conclusions based on given or collected information across industries. Challenges associated with the analysis, security, sharing, storage, and visualization of large and complex data sets continue to plague data scientists and analysts alike as traditional data processing applications struggle to adequately manage big data. The Handbook of Research on Big Data Storage and Visualization Techniques is a critical scholarly resource that explores big data analytics and technologies and their role in developing a broad understanding of issues pertaining to the use of big data in multidisciplinary fields. Featuring coverage on a broad range of topics, such as architecture patterns, programing systems, and computational energy, this publication is geared towards professionals, researchers, and students seeking current research and application topics on the subject.

Helen Macdonald, een ervaren valkenier, had nooit de behoefte een van de allerfelste roofvogels af te richten: de havik. Maar dan sterft plotseling haar vader, de bekende fotograaf en valkenier Alisdair Macdonald, in Londen, midden op straat. Overweldigd door rouw en verdriet besluit Helen een havik te nemen om haar verlies te verwerken. Door het vurige dier te temmen verkent ze haar grenzen en verandert haar leven voorgoed. In De H is van havik beschrijft Helen Macdonald hoe ze haar havik, Mabel, stapje voor stapje tam maakt. Het is een roerend en humoristisch relaas over rouwverwerking en de aantrekkingskracht van een buitengewoon beest. Macdonald betreedt onontgonnen gebied met dit hartveroverende boek dat tegelijkertijd een memoir, een boek over de natuur en een spirituele queeste is, en waarin ze onderzoekt of de dood zich laat verzoenen met het leven en de liefde.

[Personal Analytics and Privacy. An Individual and Collective Perspective](#)

[Geslacht](#)

[Big Crisis Data](#)

[Taming The Duke \(Mills & Boon Historical\)](#)

[Taming Text](#)

[Learning Technology for Education Challenges](#)

[Case Studies Using Open-Source Tools](#)

[Practical Text Analytics](#)

[Als je terug bent](#)

[Als wij schurken waren](#)

[De lange weg naar de vrijheid](#)

Taming TextHow to Find, Organize, and Manipulate ItSimon and Schuster

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 PART 1 GETTING STARTED The case for the digital Babel fish Getting started with Tika The information landscape PART 2 TIKA IN DETAIL Document type detection Content extraction Understanding metadata Language detection What's in a file? PART 3 INTEGRATION AND ADVANCED USE The big picture Tika and the Lucene search stack Extending Tika PART 4 CASE STUDIES Powering NASA science data systems Content management with Apache Jackrabbit Curating cancer research data with Tika The classic search engine example

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, Syracuse University Table of Contents Getting started taming text Foundations of taming text Searching Fuzzy string matching Identifying people, places, and things Clustering text Classification, categorization, and tagging Building an example question answering system Untamed text: exploring the next frontier

In an age where customer opinion and feedback can have an immediate, major effect upon the success of a business or organization, marketers must have the ability to analyze unstructured data in everything from social media and internet reviews to customer surveys and phone logs. Practical Text Analytics is an essential daily reference resource, providing real-world guidance on the effective application of text analytics. The book presents the analysis process so that it is immediately understood by the marketing professionals who must use it, so they can apply proven concepts and methods correctly and with confidence. By decoding industry terminology and demonstrating practical application of data models once reserved for experts, Practical Text Analytics shows marketers how to frame the right questions, identify key themes and find hidden meaning from unstructured data. Readers will learn to develop powerful new marketing strategies to elevate customer experience, solidify brand value and elevate reputation. Online resources include self-test questions, chapter review Q&A and an Instructor's Manual with text sources and instructions.

Grokking Simplicity is a friendly, practical guide that will change the way you approach software design and development. Summary Distributed across servers, difficult to test, and resistant to modification—modern software is complex. Grokking Simplicity is a friendly, practical guide that will change the way you approach software design and development. It introduces a unique approach to functional programming that explains why certain features of software are prone to complexity, and teaches you the functional techniques you can use to simplify these systems so that they 're easier to test and debug. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the technology Developers rightly fear the unintended complexity that infects most code. This book shows you how to write software that keeps complexity close to its inherent minimum. As you write software you should distinguish between code that alters your system 's state, and code that does not. Once you learn to make that distinction, you can refactor much of your state-altering "actions" into stateless "calculations." Your software will be simpler. About the book The book also teaches you to solve the complex timing bugs that inevitably creep into asynchronous and multithreaded code. In advanced sections of the book you learn how composable abstractions help avoid repeating code and open up new levels of expressivity. What's inside Patterns for simpler code Powerful time modeling approaches to simplify asynchronous code How higher-order functions can make code reusable and composable About the reader For intermediate and advanced developers building complex software. Exercises, illustrations, self-assessments, and hands-on examples lock in each new idea. About the author Eric Normand is an expert software developer who has been an influential teacher of functional programming since 2007. Table of Contents 1 Welcome to Grokking Simplicity 2 Functional thinking in action PART 1 - ACTIONS, CALCULATIONS, AND DATA 3 Distinguishing actions, calculations, and data 4 Extracting calculations from actions 5 Improving the design of actions 6 Staying immutable in a mutable language 7 Staying immutable with untrusted code 8 Stratified design, part 1 9 Stratified design, part 2 PART 2 - FIRST-CLASS ABSTRACTIONS 10 First-class functions, part 1 11 First-class functions, part 2 12 Functional iteration 13 Chaining functional tools 14 Functional tools for nested data 15 Isolating timelines 16 Sharing resources between timelines 17 Coordinating timelines 18 Reactive and onion architectures 19 The functional journey ahead

IEDEREEN KENT EEN NICE GUY. Misschien ben je zelf zo'n man... Een man die fantastisch kan luisteren, wiens vriendinnen vaak tegen hem zeggen dat hij zo'n aanwinst is, maar die nooit een echte date heeft op een zaterdagavond. De echtgenoot die wanhopig probeert zijn vrouw te behagen, maar wiens huwelijk emotioneel en seksueel verre van bevredigend is. De betrouwbare vriend bij wie iedereen aanklopt voor klusjes en adviezen, maar wiens eigen leven een rommeltje is. Volgens dr. Glover geloven deze mannen in de mythe dat ze geliefd zullen worden en nooit ergens problemen mee zullen ondervinden, zolang ze maar 'aardig' zijn en altijd alles 'goed' doen - terwijl ze ondertussen hun eigen behoeften en mannelijkheid opofferen. Dit is wat dr. Robert A. Glover, psychotherapeut en codirecteur van het Center for Healing en Recovery (Washington), het 'Nice Guy-syndroom' noemt. Nooit meer Mister Nice Guy is gebaseerd op jarenlange klinische groepstherapie met mannen en stellen. Het boek vernietigt de schadelijke mythe waarin Nice Guys geloven. Ben je een Nice Guy? Dan leert dit boek je dat je de destructieve patronen kunt doorbreken en de controle over je eigen leven weer terug kunt krijgen door zelf de

verantwoordelijkheid te nemen voor het vervullen van je eigen behoeften. Je leert om: - te stoppen met zoeken naar goedkeuring van anderen; - grenzen te stellen en integer om te gaan met conflicten; - te genieten van intieme relaties die gebaseerd zijn op wederzijds respect; - een spannender seksleven te hebben; - je carrière naar een hoger niveau te brengen en de professionele aandacht te krijgen die je verdient. Nooit meer Mister Nice Guy staat vol bewezen strategieën en uit het leven gegrepen succesverhalen. Het boek biedt handvatten die voor jou (en/of je partner) van onschatbare waarde kunnen zijn om gezonde en meer bevredigende relaties op te bouwen en met vertrouwen van het leven te genieten.

While the term Big Data is open to varying interpretation, it is quite clear that the Volume, Velocity, and Variety (3Vs) of data have impacted every aspect of computational science and its applications. The volume of data is increasing at a phenomenal rate and a majority of it is unstructured. With big data, the volume is so large that processing it using traditional database and software techniques is difficult, if not impossible. The drivers are the ubiquitous sensors, devices, social networks and the all-pervasive web. Scientists are increasingly looking to derive insights from the massive quantity of data to create new knowledge. In common usage, Big Data has come to refer simply to the use of predictive analytics or other certain advanced methods to extract value from data, without any required magnitude thereon. Challenges include analysis, capture, curation, search, sharing, storage, transfer, visualization, and information privacy. While there are challenges, there are huge opportunities emerging in the fields of Machine Learning, Data Mining, Statistics, Human-Computer Interfaces and Distributed Systems to address ways to analyze and reason with this data. The edited volume focuses on the challenges and opportunities posed by "Big Data" in a variety of domains and how statistical techniques and innovative algorithms can help glean insights and accelerate discovery. Big data has the potential to help companies improve operations and make faster, more intelligent decisions. Review of big data research challenges from diverse areas of scientific endeavor Rich perspective on a range of data science issues from leading researchers Insight into the mathematical and statistical theory underlying the computational methods used to address big data analytics problems in a variety of domains

Text Mining and Visualization: Case Studies Using Open-Source Tools provides an introduction to text mining using some of the most popular and powerful open-source tools: KNIME, RapidMiner, Weka, R, and Python. The contributors—all highly experienced with text mining and open-source software—explain how text data are gathered and processed from a w

[How to Find, Organize, and Manipulate It](#)

[A Guide to Data Science](#)

[een bewezen plan om te krijgen wat je wilt in de liefde, seks en het leven.](#)

[Computational Analysis and Understanding of Natural Languages: Principles, Methods and Applications](#)

[Tika in Action](#)

[Perspectives in Business Informatics Research](#)

[Relevant Search](#)

[Social Media in Disasters and Time-Critical Situations](#)

[Natural Language Processing in Action](#)

[Keeping Customers Informed](#)

[8th International Workshop, LTEC 2019, Zamora, Spain, July 15 – 18, 2019, Proceedings](#)

[Handbook of Linguistic Annotation](#)

Users expect search to be simple: They enter a few terms and expect perfectly-organized, relevant results instantly. But behind this simple user experience, complex machinery is at work. Whether using Elasticsearch, Solr, or another search technology, the solution is never one size fits all. Returning the right search results requires conveying domain knowledge and business rules in the search engine's data structures, text analytics, and results ranking capabilities. Relevant Search demystifies relevance work. Using Elasticsearch, it tells how to return engaging search results to users, helping readers understand and leverage the internals of Lucene-based search engines. The book walks through several real-world problems using a cohesive philosophy that combines text analysis, query building, and score shaping to express business ranking rules to the search engine. It outlines how to guide the engineering process by monitoring search user behavior and shifting the enterprise to a search-first culture focused on humans, not computers. It also shows how the search engine provides a deeply pluggable platform for integrating search ranking with machine learning, ontologies, personalization, domain-specific expertise, and other enriching sources. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications.

Create data mining algorithms About This Book Develop a strong strategy to solve predictive modeling problems using the most popular data mining algorithms Real-world case studies will take you from novice to intermediate to apply data mining techniques Deploy cutting-edge sentiment analysis techniques to real-world social media data using R Who This Book Is For This Learning Path is for R developers who are looking to making a career in data analysis or data mining. Those who come across data mining problems of different complexities from web, text, numerical, political, and social media domains will find all information in this single learning path. What You Will Learn Discover how to manipulate data in R Get to know top classification algorithms written in R Explore solutions written in R based on R Hadoop projects Apply data management skills in handling large data sets Acquire knowledge about neural network concepts and their applications in data mining Create predictive models for classification, prediction, and recommendation Use various libraries on R CRAN for data mining Discover more about data potential, the pitfalls, and inferential gotchas Gain an insight into the concepts of supervised and

unsupervised learning Delve into exploratory data analysis Understand the minute details of sentiment analysis In Detail Data mining is the first step to understanding data and making sense of heaps of data. Properly mined data forms the basis of all data analysis and computing performed on it. This learning path will take you from the very basics of data mining to advanced data mining techniques, and will end up with a specialized branch of data mining—social media mining. You will learn how to manipulate data with R using code snippets and how to mine frequent patterns, association, and correlation while working with R programs. You will discover how to write code for various predication models, stream data, and time-series data. You will also be introduced to solutions written in R based on R Hadoop projects. Now that you are comfortable with data mining with R, you will move on to implementing your knowledge with the help of end-to-end data mining projects. You will learn how to apply different mining concepts to various statistical and data applications in a wide range of fields. At this stage, you will be able to complete complex data mining cases and handle any issues you might encounter during projects. After this, you will gain hands-on experience of generating insights from social media data. You will get detailed instructions on how to obtain, process, and analyze a variety of socially-generated data while providing a theoretical background to accurately interpret your findings. You will be shown R code and examples of data that can be used as a springboard as you get the chance to undertake your own analyses of business, social, or political data. This Learning Path combines some of the best that Packt has to offer in one complete, curated package. It includes content from the following Packt products: Learning Data Mining with R by Biter Makhabel R Data Mining Blueprints by Pradeepta Mishra Social Media Mining with R by Nathan Danneman and Richard Heimann Style and approach A complete package with which will take you from the basics of data mining to advanced data mining techniques, and will end up with a specialized branch of data mining—social media mining.

Derde deel uit de populaire Sea Breeze-serie Low heeft het hart van Cage gebroken nadat ze een relatie met Marcus in Alleen voor haar. Het leven van Cage begon langzaam van de rails te raken, eindigend in een arrestatie voor beschonken rijden. Om zijn baseballbeurs te kunnen behouden - het enige waar hij naast Low om geeft - moet Cage een zomerbaantje zien te vinden. Weg van Sea Breeze. Op een boerderij, met een hoop koeien en zonder hete meisjes. Misschien is dat wat Cage nodig heeft om alles weer op orde te krijgen. Maar de eigenaar van de boerderij blijkt een knappe maar chagrijnige dochter te hebben. Cage ziet een hoop geheimen en verdriet in haar ogen. Hij wil niets liever dan haar zowel mentaal als fysiek uitkleden in de schuur, maar is hij wel klaar voor de consequenties?

There was comfort in her touch... A kind of healing that brought peace and stirrings of love where Dalton Warfield never expected to find them – within himself! But could he truly surrender to the graciousness that was Alicia Spencer – when he was the Duke of Wexton and she outcast by the Ton?

This research monograph brings AI to the field of Customer Relationship Management (CRM) to make a customer experience with a product or service smart and enjoyable. AI is here to help customers to get a refund for a canceled flight, unfreeze a banking account or get a health test result. Today, CRM has evolved from storing and analyzing customers' data to predicting and understanding their behavior by putting a CRM system in a customers' shoes. Hence advanced reasoning with learning from small data, about customers' attitudes, introspection, reading between the lines of customer communication and explainability need to come into play. Artificial Intelligence for Customer Relationship Management leverages a number of Natural Language Processing (NLP), Machine Learning (ML), simulation and reasoning techniques to enable CRM with intelligence. An effective and robust CRM needs to be able to chat with customers, providing desired information, completing their transactions and resolving their problems. It introduces a systematic means of ascertaining a customers' frame of mind, their intents and attitudes to determine when to provide a thorough answer, a recommendation, an explanation, a proper argument, timely advice and promotion or compensation. The author employs a spectrum of ML methods, from deterministic to statistical to deep, to predict customer behavior and anticipate possible complaints, assuring customer retention efficiently. Providing a forum for the exchange of ideas in AI, this book provides a concise yet comprehensive coverage of methodologies, tools, issues, applications, and future trends for professionals, managers, and researchers in the CRM field together with AI and IT professionals.

De lange weg naar de vrijheid is de beroemde autobiografie van een van de grootste mannen van de twintigste eeuw. Nelson Mandela beschrijft de lange weg die hij heeft moeten afleggen van onwetende jongen tot charismatisch staatsman. Dit is het verhaal van misschien wel de wonderbaarlijkste omwenteling in de geschiedenis, verteld door de man die het allemaal heeft meegemaakt en in gang gezet. Het verhaal van Mandela, door Mandela.

Waar zij was, zou altijd mijn thuis zijn. Het werd tijd dat ik de volgende stap zette en haar bewees dat ik er helemaal voor ging. Voor altijd. Een toevallige ontmoeting in Rosemary Beach verandert in een sprookjesromance als Reese op het punt staat om samen te gaan wonen met Mase op de familieranch in Texas. Met de liefdevolle ondersteuning van de Colts, en de recente ontdekking dat ze zelf ook familie heeft, kan Reese eindelijk de vreselijke gebeurtenissen uit haar verleden achter zich laten en de eerste stappen zetten naar een rooskleurige toekomst. Maar dit dreigt te worden gedwardsboomd. Hoewel Aida denkt dat ze een nicht van Mase is, is het voor Reese wel duidelijk dat de aantrekkelijke blondine Mase niet als familie heeft liefheeft. En dan is er nog Captain - mysterieus en knap - die Reese vertelt wat hij van Mase vindt, en wat hij voor haar voelt. Als de strijd om de harten van Mase en Reese aanwakkert, begint het verleden van Reese zicht te ontrafelen. Schokkende waarheden worden onthuld die het leven van het gelukkige stel voor altijd zullen veranderen.

This book constitutes the proceedings of the 16th International Conference on Perspectives in Business Informatics Research, BIR 2017, held in Copenhagen, Denmark, in August 2017. This year the BIR conference attracted 59 submissions from 23 countries. They were reviewed by 45 members of the Program Committee, and as a result, 17 full papers and 3 short papers were selected for presentation at the conference and publication in this volume. They are organized in sections on enterprise architecture, business process management, business analytics, information systems applications, and information systems development. In addition, the summaries of the two conference keynotes are also included. This year, the conference theme was the digital transformation, which will impact most businesses, organizations and societies and call for new and radical approaches to how we adopt, use and manage IT.

[Grokking Simplicity](#)

[Hadoop in Practice](#)

[Includes 104 Techniques](#)

[Bevrijdend licht](#)

[De havik](#)

[Methodologies and Applications](#)

[16th International Conference, BIR 2017, Copenhagen, Denmark, August 28–30, 2017, Proceedings](#)

[ze is adembenemend en levensgevaarlijk](#)

[Zolang het duurt](#)

[de autobiografie](#)

[De H is van Havik](#)