

Instant Web Scraping With Java

Mastering Data Analysis with R Enterprise Application Integration with XML and Java Learning Scrapy Java and XSLT OpenJDK Cookbook Mastering Concurrency in Python Encyclopedia of Criminal Activities and the Deep Web Java for Data Science Automate the Boring Stuff with Python Professional JavaScript Web Scraping with Python Numerical Methods in Engineering with Python Webbots, Spiders, and Screen Scrapers, 2nd Edition Data Wrangling with Python Tiny House Living Advanced Python Programming Web Application Vulnerabilities Instant PHP Web Scraping Architecting Modern Java EE Applications Instant Web Scraping with Java PHP 7 Data Structures and Algorithms Python Web Scraping RESTful Web Services Data Visualization with Python and JavaScript Learning Robotic Process Automation Instant Nokogiri Getting Started with Beautiful Soup Python Natural Language Processing Instant Messaging in Java Python Requests Essentials Web Scraping with Python Learning Python Application Development Automated Data Collection with R Go Web Scraping Quick Start Guide Instant Jsoup How-To Jupyter for Data Science Python Web Scraping Cookbook Building a RESTful Web Service with Spring Programming Collective Intelligence How to Borrow Your Way Out of Debt

Mastering Data Analysis with R

Take Python beyond scripting to build robust, reusable, and efficient applications About This Book Get to grips with Python techniques that address

commonly encountered problems in general application development. Develop, package, and deploy efficient applications in a fun way. All-practical coverage of the major areas of application development, including best practices, exception handling, testing, refactoring, design patterns, performance, and GUI application development. Who This Book Is For Do you know the basics of Python and object oriented programming? Do you want to go an extra mile and learn techniques to make your Python application robust, extensible, and efficient? Then this book is for you. What You Will Learn Build a robust application by handling exceptions. Modularize, package, and release the source distribution. Document the code and implement coding standards. Create automated tests to catch bugs in the early development stage. Identify and re-factor badly written code to improve application life. Detect recurring problems in the code and apply design patterns. Improve code efficiency by identifying performance bottlenecks and fixing them. Develop simple GUI applications using Python. In Detail Python is one of the most widely used dynamic programming languages, supported by a rich set of libraries and frameworks that enable rapid development. But fast paced development often comes with its own baggage that could bring down the quality, performance, and extensibility of an application. This book will show you ways to handle such problems and write better Python applications. From the basics of simple command-line applications, develop your skills all the way to designing efficient and advanced Python apps. Guided by a light-hearted fantasy learning theme, overcome the real-world problems of

complex Python development with practical solutions. Beginning with a focus on robustness, packaging, and releasing application code, you'll move on to focus on improving application lifetime by making code extensible, reusable, and readable. Get to grips with Python refactoring, design patterns and best practices. Techniques to identify the bottlenecks and improve performance are covered in a series of chapters devoted to performance, before closing with a look at developing Python GUIs. Style and approach The book uses a fantasy game theme as a medium to explain various topics. Specific aspects of application development are explained in different chapters. In each chapter the reader is presented with an interesting problem which is then tackled using hands-on examples with easy-to-follow instructions.

Enterprise Application Integration with XML and Java

Create distributed applications with clever design patterns to solve complex problems Key Features Set up and run distributed algorithms on a cluster using Dask and PySpark Master skills to accurately implement concurrency in your code Gain practical experience of Python design patterns with real-world examples Book Description This Learning Path shows you how to leverage the power of both native and third-party Python libraries for building robust and responsive applications. You will learn about profilers and reactive programming, concurrency and parallelism, as well as tools for making your apps quick and efficient. You will discover how to write

code for parallel architectures using TensorFlow and Theano, and use a cluster of computers for large-scale computations using technologies such as Dask and PySpark. With the knowledge of how Python design patterns work, you will be able to clone objects, secure interfaces, dynamically choose algorithms, and accomplish much more in high performance computing. By the end of this Learning Path, you will have the skills and confidence to build engaging models that quickly offer efficient solutions to your problems. This Learning Path includes content from the following Packt products: Python High Performance - Second Edition by Gabriele Lanaro Mastering Concurrency in Python by Quan Nguyen Mastering Python Design Patterns by Sakis Kasampalis What you will learn Use NumPy and pandas to import and manipulate datasets Achieve native performance with Cython and Numba Write asynchronous code using asyncio and RxPy Design highly scalable programs with application scaffolding Explore abstract methods to maintain data consistency Clone objects using the prototype pattern Use the adapter pattern to make incompatible interfaces compatible Employ the strategy pattern to dynamically choose an algorithm Who this book is for This Learning Path is specially designed for Python developers who want to build high-performance applications and learn about single core and multi-core programming, distributed concurrency, and Python design patterns. Some experience with Python programming language will help you get the most out of this Learning Path.

Learning Scrapy

Learn how to turn raw data into rich, interactive web visualizations with the powerful combination of Python and JavaScript. With this hands-on guide, author Kyran Dale teaches you how build a basic dataviz toolchain with best-of-breed Python and JavaScript libraries—including Scrapy, Matplotlib, Pandas, Flask, and D3—for crafting engaging, browser-based visualizations. As a working example, throughout the book Dale walks you through transforming Wikipedia’s table-based list of Nobel Prize winners into an interactive visualization. You’ll examine steps along the entire toolchain, from scraping, cleaning, exploring, and delivering data to building the visualization with JavaScript’s D3 library. If you’re ready to create your own web-based data visualizations—and know either Python or JavaScript—this is the book for you. Learn how to manipulate data with Python Understand the commonalities between Python and JavaScript Extract information from websites by using Python’s web-scraping tools, BeautifulSoup and Scrapy Clean and explore data with Python’s Pandas, Matplotlib, and Numpy libraries Serve data and create RESTful web APIs with Python’s Flask framework Create engaging, interactive web visualizations with JavaScript’s D3 library

Java and XSLT

"This book provides a unified framework of web scraping and information extraction from text data with R for the social sciences"--

OpenJDK Cookbook

Design RPA solutions to perform a wide range of transactional tasks with minimal cost and maximum ROI Key Features A beginner's guide to learn Robotic Process Automation and its impact on the modern world Design, test, and perform enterprise automation task with UiPath Create Automation apps and deploy them to all the computers in your department. Book Description Robotic Process Automation (RPA) enables automating business processes using software robots. Software robots interpret, trigger responses, and communicate with other systems just like humans do. Robotic processes and intelligent automation tools can help businesses improve the effectiveness of services faster and at a lower cost than current methods. This book is the perfect start to your automation journey, with a special focus on one of the most popular RPA tools: UiPath. Learning Robotic Process Automation takes you on a journey from understanding the basics of RPA to advanced implementation techniques. You will become oriented in the UiPath interface and learn about its workflow. Once you are familiar with the environment, we will get hands-on with automating different applications such as Excel, SAP, Windows and web applications, screen and web scraping, working with user events, as well as understanding exceptions and debugging. By the end of the book, you'll not only be able to build your first software bot, but also you'll wire it to perform various automation tasks with the help of best practices for bot deployment. What you will learn Understand Robotic

Process Automation technology Learn UiPath programming techniques to deploy robot configurations Explore various data extraction techniques Learn about integrations with various popular applications such as SAP and MS Office Debug a programmed robot including logging and exception handling Maintain code version and source control Deploy and control Bots with UiPath Orchestrator Who this book is for If you would like to pursue a career in Robotic Process Automation or improve the efficiency of your businesses by automating common tasks, then this book is perfect for you. Prior programming knowledge of either Visual Basic or C# will be useful.

Mastering Concurrency in Python

In this book, we aim to describe how to make a computer bend to your will by finding and exploiting vulnerabilities specifically in Web applications. We will describe common security issues in Web applications, tell you how to find them, describe how to exploit them, and then tell you how to fix them. We will also cover how and why some hackers (the bad guys) will try to exploit these vulnerabilities to achieve their own end. We will also try to explain how to detect if hackers are actively trying to exploit vulnerabilities in your own Web applications. Learn to defend Web-based applications developed with AJAX, SOAP, XMLRPC, and more. See why Cross Site Scripting attacks can be so devastating.

Encyclopedia of Criminal Activities and the Deep Web

Three super-hot topics come together in this first complete guide to Enterprise Application Integration with XML and Java. The book teaches readers to identify data exchange requirements and meet them with Java and XML. It contains easy-to-read, well-documented code throughout. The CD-ROM contains extensive source code from the book, plus a library of leading-edge software and trialware.

Java for Data Science

Find out how to craft effective, business-oriented Java EE 8 applications that target customer's demands in the age of Cloud platforms and container technology. About This Book Understand the principles of modern Java EE and how to realize effective architectures Gain knowledge of how to design enterprise software in the age of automation, Continuous Delivery and Cloud platforms Learn about the reasoning and motivations behind state-of-the-art enterprise Java technology, that focuses on business Who This Book Is For This book is for experienced Java EE developers who are aspiring to become the architects of enterprise-grade applications, or software architects who would like to leverage Java EE to create effective blueprints of applications. What You Will Learn What enterprise software engineers should focus on Implement applications, packages, and components in a modern way Design and structure application architectures Discover how to realize technical and cross-cutting aspects Get to grips with containers and container orchestration technology Realize zero-dependency, 12-factor, and Cloud-native applications

Implement automated, fast, reliable, and maintainable software tests Discover distributed system architectures and their requirements In Detail Java EE 8 brings with it a load of features, mainly targeting newer architectures such as microservices, modernized security APIs, and cloud deployments. This book will teach you to design and develop modern, business-oriented applications using Java EE 8. It shows how to structure systems and applications, and how design patterns and Domain Driven Design aspects are realized in the age of Java EE 8. You will learn about the concepts and principles behind Java EE applications, and how to effect communication, persistence, technical and cross-cutting concerns, and asynchronous behavior. This book covers Continuous Delivery, DevOps, infrastructure-as-code, containers, container orchestration technologies, such as Docker and Kubernetes, and why and especially how Java EE fits into this world. It also covers the requirements behind containerized, zero-dependency applications and how modern Java EE application servers support these approaches. You will also learn about automated, fast, and reliable software tests, in different test levels, scopes, and test technologies. This book covers the prerequisites and challenges of distributed systems that lead to microservice, shared-nothing architectures. The challenges and solutions of consistency versus scalability will further lead us to event sourcing, event-driven architectures, and the CQRS principle. This book also includes the nuts and bolts of application performance as well as how to realize resilience, logging, monitoring and tracing in a modern enterprise world. Last but not least the demands of securing enterprise systems are covered.

By the end, you will understand the ins and outs of Java EE so that you can make critical design decisions that not only live up to, but also surpass your clients' expectations. Style and approach This book focuses on solving business problems and meeting customer demands in the enterprise world. It covers how to create enterprise applications with reasonable technology choices, free of cargo-cult and over-engineering. The aspects shown in this book not only demonstrate how to realize a certain solution, but also explain its motivations and reasoning.

Automate the Boring Stuff with Python

Inside: Learn how to manage debt (the good, the bad, and the truly ugly) How to invest with taxes in mind How to use leverage safely to ramp-up your returns (in other words, what the rich do that the average investor doesn't) Special, little-known investments, some previously available only to a select few (e.g. syndicated mortgages) explained How to find the right financial advisor for you, plus when you absolutely must seek the advice of other professionals BONUS: The Not So Wonderful World of Student Debt And much more

Professional JavaScript

Gain sharp insights into your data and solve real-world data science problems with R—from data munging to modeling and visualization About This Book Handle your data with precision and care for optimal business intelligence Restructure and

transform your data to inform decision-making Packed with practical advice and tips to help you get to grips with data mining Who This Book Is For If you are a data scientist or R developer who wants to explore and optimize your use of R's advanced features and tools, this is the book for you. A basic knowledge of R is required, along with an understanding of database logic. What You Will Learn Connect to and load data from R's range of powerful databases Successfully fetch and parse structured and unstructured data Transform and restructure your data with efficient R packages Define and build complex statistical models with glm Develop and train machine learning algorithms Visualize social networks and graph data Deploy supervised and unsupervised classification algorithms Discover how to visualize spatial data with R In Detail R is an essential language for sharp and successful data analysis. Its numerous features and ease of use make it a powerful way of mining, managing, and interpreting large sets of data. In a world where understanding big data has become key, by mastering R you will be able to deal with your data effectively and efficiently. This book will give you the guidance you need to build and develop your knowledge and expertise. Bridging the gap between theory and practice, this book will help you to understand and use data for a competitive advantage. Beginning with taking you through essential data mining and management tasks such as munging, fetching, cleaning, and restructuring, the book then explores different model designs and the core components of effective analysis. You will then discover how to optimize your use of machine learning algorithms for classification and

recommendation systems beside the traditional and more recent statistical methods. Style and approach
Covering the essential tasks and skills within data science, Mastering Data Analysis provides you with solutions to the challenges of data science. Each section gives you a theoretical overview before demonstrating how to put the theory to work with real-world use cases and hands-on examples.

Web Scraping with Python

Immerse yourself in the world of Python concurrency and tackle the most complex concurrent programming problems
Key Features Explore the core syntaxes, language features and modern patterns of concurrency in Python Understand how to use concurrency to keep data consistent and applications responsive Utilize application scaffolding to design highly-scalable programs
Book Description Python is one of the most popular programming languages, with numerous libraries and frameworks that facilitate high-performance computing. Concurrency and parallelism in Python are essential when it comes to multiprocessing and multithreading; they behave differently, but their common aim is to reduce the execution time. This book serves as a comprehensive introduction to various advanced concepts in concurrent engineering and programming. Mastering Concurrency in Python starts by introducing the concepts and principles in concurrency, right from Amdahl's Law to multithreading programming, followed by elucidating multiprocessing programming, web scraping, and asynchronous I/O, together with

common problems that engineers and programmers face in concurrent programming. Next, the book covers a number of advanced concepts in Python concurrency and how they interact with the Python ecosystem, including the Global Interpreter Lock (GIL). Finally, you'll learn how to solve real-world concurrency problems through examples. By the end of the book, you will have gained extensive theoretical knowledge of concurrency and the ways in which concurrency is supported by the Python language

What you will learn

- Explore the concepts of concurrency in programming
- Explore the core syntax and features that enable concurrency in Python
- Understand the correct way to implement concurrency
- Abstract methods to keep the data consistent in your program
- Analyze problems commonly faced in concurrent programming
- Use application scaffolding to design highly-scalable programs

Who this book is for

This book is for developers who wish to build high-performance applications and learn about single-core, multicore programming or distributed concurrency. Some experience with Python programming language is assumed.

Numerical Methods in Engineering with Python

Learn how some Go-specific language features help to simplify building web scrapers along with common pitfalls and best practices regarding web scraping.

Key Features

- Use Go libraries like Goquery and Colly to scrape the web
- Common pitfalls and best practices

to effectively scrape and crawl Learn how to scrape using the Go concurrency model Book Description Web scraping is the process of extracting information from the web using various tools that perform scraping and crawling. Go is emerging as the language of choice for scraping using a variety of libraries. This book will quickly explain to you, how to scrape data data from various websites using Go libraries such as Colly and Goquery. The book starts with an introduction to the use cases of building a web scraper and the main features of the Go programming language, along with setting up a Go environment. It then moves on to HTTP requests and responses and talks about how Go handles them. You will also learn about a number of basic web scraping etiquettes. You will be taught how to navigate through a website, using a breadth-first and then a depth-first search, as well as find and follow links. You will get to know about the ways to track history in order to avoid loops and to protect your web scraper using proxies. Finally the book will cover the Go concurrency model, and how to run scrapers in parallel, along with large-scale distributed web scraping. What you will learn Implement Cache-Control to avoid unnecessary network calls Coordinate concurrent scrapers Design a custom, larger-scale scraping system Scrape basic HTML pages with Colly and JavaScript pages with chromedp Discover how to search using the "strings" and "regexp" packages Set up a Go development environment Retrieve information from an HTML document Protect your web scraper from being blocked by using proxies Control web browsers to scrape JavaScript sites Who this book is for Data

scientists, and web developers with a basic knowledge of Golang wanting to collect web data and analyze them for effective reporting and visualization.

Webbots, Spiders, and Screen Scrapers, 2nd Edition

Successfully scrape data from any website with the power of Python About This Book A hands-on guide to web scraping with real-life problems and solutions Techniques to download and extract data from complex websites Create a number of different web scrapers to extract information Who This Book Is For This book is aimed at developers who want to use web scraping for legitimate purposes. Prior programming experience with Python would be useful but not essential. Anyone with general knowledge of programming languages should be able to pick up the book and understand the principals involved. What You Will Learn Extract data from web pages with simple Python programming Build a threaded crawler to process web pages in parallel Follow links to crawl a website Download cache to reduce bandwidth Use multiple threads and processes to scrape faster Learn how to parse JavaScript-dependent websites Interact with forms and sessions Solve CAPTCHAs on protected web pages Discover how to track the state of a crawl In Detail The Internet contains the most useful set of data ever assembled, largely publicly accessible for free. However, this data is not easily reusable. It is embedded within the structure and style of websites and needs to be carefully extracted to be useful. Web scraping is becoming increasingly

useful as a means to easily gather and make sense of the plethora of information available online. Using a simple language like Python, you can crawl the information out of complex websites using simple programming. This book is the ultimate guide to using Python to scrape data from websites. In the early chapters it covers how to extract data from static web pages and how to use caching to manage the load on servers. After the basics we'll get our hands dirty with building a more sophisticated crawler with threads and more advanced topics. Learn step-by-step how to use Ajax URLs, employ the Firebug extension for monitoring, and indirectly scrape data. Discover more scraping nitty-gritties such as using the browser renderer, managing cookies, how to submit forms to extract data from complex websites protected by CAPTCHA, and so on. The book wraps up with how to create high-level scrapers with Scrapy libraries and implement what has been learned to real websites.

Style and approach This book is a hands-on guide with real-life examples and solutions starting simple and then progressively becoming more complex. Each chapter in this book introduces a problem and then provides one or more possible solutions.

Data Wrangling with Python

Filled with practical, step-by-step instructions and clear explanations for the most important and useful tasks. This book will take a how-to approach, focusing on recipes that demonstrate Jsoup. If you are working in data scraping, data crawling, or within a similar area using Java, then this book is the one for you. This

book acts as a fast-paced and simple guide to enhance your HTML data manipulating skills using one of the most well-known libraries, Jsoup.

Tiny House Living

There's a wealth of data online, but sorting and gathering it by hand can be tedious and time consuming. Rather than click through page after endless page, why not let bots do the work for you? Webbots, Spiders, and Screen Scrapers will show you how to create simple programs with PHP/CURL to mine, parse, and archive online data to help you make informed decisions. Michael Schrenk, a highly regarded webbot developer, teaches you how to develop fault-tolerant designs, how best to launch and schedule the work of your bots, and how to create Internet agents that:

- Send email or SMS notifications to alert you to new information quickly
- Search different data sources and combine the results on one page, making the data easier to interpret and analyze
- Automate purchases, auction bids, and other online activities to save time

Sample projects for automating tasks like price monitoring and news aggregation will show you how to put the concepts you learn into practice. This second edition of Webbots, Spiders, and Screen Scrapers includes tricks for dealing with sites that are resistant to crawling and scraping, writing stealthy webbots that mimic human search behavior, and using regular expressions to harvest specific data. As you discover the possibilities of web scraping, you'll see how webbots can save you precious time and give you much greater control over

the data available on the Web.

Advanced Python Programming

Get to grips with a new technology, understand what it is and what it can do for you, and then get to work with the most important features and tasks. A concise, illustrated guide to extracting information available on the Internet using Nokogiri."Instant Nokogiri" is the perfect choice for the aspiring Ruby developer looking to incorporate screen scraping and parsing technology in their applications. Beginner level Ruby, basic HTML, and CSS experience is suggested.

Web Application Vulnerabilities

Successfully scrape data from any website with the power of Python 3.x About This Book A hands-on guide to web scraping using Python with solutions to real-world problems Create a number of different web scrapers in Python to extract information This book includes practical examples on using the popular and well-maintained libraries in Python for your web scraping needs Who This Book Is For This book is aimed at developers who want to use web scraping for legitimate purposes. Prior programming experience with Python would be useful but not essential. Anyone with general knowledge of programming languages should be able to pick up the book and understand the principals involved. What You Will Learn Extract data from web pages with simple Python programming Build a concurrent

crawler to process web pages in parallel Follow links to crawl a website Extract features from the HTML Cache downloaded HTML for reuse Compare concurrent models to determine the fastest crawler Find out how to parse JavaScript-dependent websites Interact with forms and sessions In Detail The Internet contains the most useful set of data ever assembled, most of which is publicly accessible for free. However, this data is not easily usable. It is embedded within the structure and style of websites and needs to be carefully extracted. Web scraping is becoming increasingly useful as a means to gather and make sense of the wealth of information available online. This book is the ultimate guide to using the latest features of Python 3.x to scrape data from websites. In the early chapters, you'll see how to extract data from static web pages. You'll learn to use caching with databases and files to save time and manage the load on servers. After covering the basics, you'll get hands-on practice building a more sophisticated crawler using browsers, crawlers, and concurrent scrapers. You'll determine when and how to scrape data from a JavaScript-dependent website using PyQt and Selenium. You'll get a better understanding of how to submit forms on complex websites protected by CAPTCHA. You'll find out how to automate these actions with Python packages such as mechanize. You'll also learn how to create class-based scrapers with Scrapy libraries and implement your learning on real websites. By the end of the book, you will have explored testing websites with scrapers, remote scraping, best practices, working with images, and many other relevant topics. Style and approach This hands-on guide is full of real-life examples and

solutions starting simple and then progressively becoming more complex. Each chapter in this book introduces a problem and then provides one or more possible solutions.

Instant PHP Web Scraping

Python is one of the most popular programming languages of our era; the Python Requests library is one of the world's best clients, with the highest number of downloads. It allows hassle-free interactions with web applications using simple procedures. You will be shown how to mock HTTP Requests using HTTPretty, and will learn to interact with social media using Requests. This book will help you to grasp the art of web scraping with the BeautifulSoup and Python Requests libraries, and will then paddle you through Requests impressive ability to interact with APIs. It will empower you with the best practices for seamlessly drawing data from web apps. Last but not least, you will get the chance to polish your skills by implementing a RESTful Web API with Python and Flask!

Architecting Modern Java EE Applications

Simplify your ETL processes with these hands-on data hygiene tips, tricks, and best practices. Key Features Focus on the basics of data wrangling Study various ways to extract the most out of your data in less time Boost your learning curve with bonus topics like random data generation and data integrity checks Book Description For data to be useful and

meaningful, it must be curated and refined. Data Wrangling with Python teaches you the core ideas behind these processes and equips you with knowledge of the most popular tools and techniques in the domain. The book starts with the absolute basics of Python, focusing mainly on data structures. It then delves into the fundamental tools of data wrangling like NumPy and Pandas libraries. You'll explore useful insights into why you should stay away from traditional ways of data cleaning, as done in other languages, and take advantage of the specialized pre-built routines in Python. This combination of Python tips and tricks will also demonstrate how to use the same Python backend and extract/transform data from an array of sources including the Internet, large database vaults, and Excel financial tables. To help you prepare for more challenging scenarios, you'll cover how to handle missing or wrong data, and reformat it based on the requirements from the downstream analytics tool. The book will further help you grasp concepts through real-world examples and datasets. By the end of this book, you will be confident in using a diverse array of sources to extract, clean, transform, and format your data efficiently. What you will learn

- Use and manipulate complex and simple data structures
- Harness the full potential of DataFrames and numpy.array at run time
- Perform web scraping with BeautifulSoup4 and html5lib
- Execute advanced string search and manipulation with RegEX
- Handle outliers and perform data imputation with Pandas
- Use descriptive statistics and plotting techniques
- Practice data wrangling and modeling using data generation techniques

Who this book is for Data Wrangling with

Python is designed for developers, data analysts, and business analysts who are keen to pursue a career as a full-fledged data scientist or analytics expert. Although, this book is for beginners, prior working knowledge of Python is necessary to easily grasp the concepts covered here. It will also help to have rudimentary knowledge of relational database and SQL.

Instant Web Scraping with Java

If you are an experienced Java developer using Java 7 platform and want to get your grips on OpenJDK for Java development, this is the book for you. JDK users who wish to migrate to OpenJDK will find this book very useful.

PHP 7 Data Structures and Algorithms

Leverage the power of machine learning and deep learning to extract information from text data About This Book Implement Machine Learning and Deep Learning techniques for efficient natural language processing Get started with NLTK and implement NLP in your applications with ease Understand and interpret human languages with the power of text analysis via Python Who This Book Is For This book is intended for Python developers who wish to start with natural language processing and want to make their applications smarter by implementing NLP in them. What You Will Learn Focus on Python programming paradigms, which are used to develop NLP applications Understand corpus analysis and different

types of data attribute. Learn NLP using Python libraries such as NLTK, Polyglot, SpaCy, Stanford CoreNLP and so on Learn about Features Extraction and Feature selection as part of Features Engineering. Explore the advantages of vectorization in Deep Learning. Get a better understanding of the architecture of a rule-based system. Optimize and fine-tune Supervised and Unsupervised Machine Learning algorithms for NLP problems. Identify Deep Learning techniques for Natural Language Processing and Natural Language Generation problems. In Detail This book starts off by laying the foundation for Natural Language Processing and why Python is one of the best options to build an NLP-based expert system with advantages such as Community support, availability of frameworks and so on. Later it gives you a better understanding of available free forms of corpus and different types of dataset. After this, you will know how to choose a dataset for natural language processing applications and find the right NLP techniques to process sentences in datasets and understand their structure. You will also learn how to tokenize different parts of sentences and ways to analyze them. During the course of the book, you will explore the semantic as well as syntactic analysis of text. You will understand how to solve various ambiguities in processing human language and will come across various scenarios while performing text analysis. You will learn the very basics of getting the environment ready for natural language processing, move on to the initial setup, and then quickly understand sentences and language parts. You will learn the power of Machine Learning and Deep Learning to extract information from text data. By the

end of the book, you will have a clear understanding of natural language processing and will have worked on multiple examples that implement NLP in the real world. Style and approach This book teaches the readers various aspects of natural language Processing using NLTK. It takes the reader from the basic to advance level in a smooth way.

Python Web Scraping

Want to tap the power behind search rankings, product recommendations, social bookmarking, and online matchmaking? This fascinating book demonstrates how you can build Web 2.0 applications to mine the enormous amount of data created by people on the Internet. With the sophisticated algorithms in this book, you can write smart programs to access interesting datasets from other web sites, collect data from users of your own applications, and analyze and understand the data once you've found it. Programming Collective Intelligence takes you into the world of machine learning and statistics, and explains how to draw conclusions about user experience, marketing, personal tastes, and human behavior in general -- all from information that you and others collect every day. Each algorithm is described clearly and concisely with code that can immediately be used on your web site, blog, Wiki, or specialized application. This book explains:

- Collaborative filtering techniques that enable online retailers to recommend products or media
- Methods of clustering to detect groups of similar items in a large dataset
- Search engine features -- crawlers, indexers,

query engines, and the PageRank algorithm
Optimization algorithms that search millions of possible solutions to a problem and choose the best one
Bayesian filtering, used in spam filters for classifying documents based on word types and other features
Using decision trees not only to make predictions, but to model the way decisions are made
Predicting numerical values rather than classifications to build price models
Support vector machines to match people in online dating sites
Non-negative matrix factorization to find the independent features in a dataset
Evolving intelligence for problem solving -- how a computer develops its skill by improving its own code the more it plays a game
Each chapter includes exercises for extending the algorithms to make them more powerful. Go beyond simple database-backed applications and put the wealth of Internet data to work for you. "Bravo! I cannot think of a better way for a developer to first learn these algorithms and methods, nor can I think of a better way for me (an old AI dog) to reinvigorate my knowledge of the details." -- Dan Russell, Google
"Toby's book does a great job of breaking down the complex subject matter of machine-learning algorithms into practical, easy-to-understand examples that can be directly applied to analysis of social interaction across the Web today. If I had this book two years ago, it would have saved precious time going down some fruitless paths." -- Tim Wolters, CTO, Collective Intellect

RESTful Web Services

Numerical Methods in Engineering with Python, a student text, and a reference for practicing engineers.

Data Visualization with Python and JavaScript

Your one-stop guide to building an efficient data science pipeline using Jupyter About This Book Get the most out of your Jupyter notebook to complete the trickiest of tasks in Data Science Learn all the tasks in the data science pipeline—from data acquisition to visualization—and implement them using Jupyter Get ahead of the curve by mastering all the applications of Jupyter for data science with this unique and intuitive guide Who This Book Is For This book targets students and professionals who wish to master the use of Jupyter to perform a variety of data science tasks. Some programming experience with R or Python, and some basic understanding of Jupyter, is all you need to get started with this book. What You Will Learn Understand why Jupyter notebooks are a perfect fit for your data science tasks Perform scientific computing and data analysis tasks with Jupyter Interpret and explore different kinds of data visually with charts, histograms, and more Extend SQL's capabilities with Jupyter notebooks Combine the power of R and Python 3 with Jupyter to create dynamic notebooks Create interactive dashboards and dynamic presentations Master the best coding practices and deploy your Jupyter notebooks efficiently In Detail Jupyter Notebook is a web-based environment that enables interactive computing in notebook documents. It allows you to create

documents that contain live code, equations, and visualizations. This book is a comprehensive guide to getting started with data science using the popular Jupyter notebook. If you are familiar with Jupyter notebook and want to learn how to use its capabilities to perform various data science tasks, this is the book for you! From data exploration to visualization, this book will take you through every step of the way in implementing an effective data science pipeline using Jupyter. You will also see how you can utilize Jupyter's features to share your documents and codes with your colleagues. The book also explains how Python 3, R, and Julia can be integrated with Jupyter for various data science tasks. By the end of this book, you will comfortably leverage the power of Jupyter to perform various tasks in data science successfully. Style and approach This book is a perfect blend of concepts and practical examples, written in a way that is very easy to understand and implement. It follows a logical flow where you will be able to build on your understanding of the different Jupyter features with every chapter.

Learning Robotic Process Automation

Learn web scraping and crawling techniques to access unlimited data from any web source in any format. With this practical guide, you'll learn how to use Python scripts and web APIs to gather and process data from thousands—or even millions—of web pages at once. Ideal for programmers, security professionals, and web administrators familiar with Python, this book not only teaches basic web scraping

mechanics, but also delves into more advanced topics, such as analyzing raw data or using scrapers for frontend website testing. Code samples are available to help you understand the concepts in practice. Learn how to parse complicated HTML pages Traverse multiple pages and sites Get a general overview of APIs and how they work Learn several methods for storing the data you scrape Download, read, and extract data from documents Use tools and techniques to clean badly formatted data Read and write natural languages Crawl through forms and logins Understand how to scrape JavaScript Learn image processing and text recognition

Instant Nokogiri

Increase your productivity by implementing data structures About This Book Gain a complete understanding of data structures using a simple approach Analyze algorithms and learn when you should apply each solution Explore the true potential of functional data structures Who This Book Is For This book is for those who want to learn data structures and algorithms with PHP for better control over application-solution, efficiency, and optimization. A basic understanding of PHP data types, control structures, and other basic features is required What You Will Learn Gain a better understanding of PHP arrays as a basic data structure and their hidden power Grasp how to analyze algorithms and the Big O Notation Implement linked lists, double linked lists, stack, queues, and priority queues using PHP Work with sorting, searching, and recursive algorithms

Make use of greedy, dynamic, and pattern matching algorithms Implement tree, heaps, and graph algorithms Apply PHP functional data structures and built-in data structures and algorithms In Detail PHP has always been the the go-to language for web based application development, but there are materials and resources you can refer to to see how it works. Data structures and algorithms help you to code and execute them effectively, cutting down on processing time significantly. If you want to explore data structures and algorithms in a practical way with real-life projects, then this book is for you. The book begins by introducing you to data structures and algorithms and how to solve a problem from beginning to end using them. Once you are well aware of the basics, it covers the core aspects like arrays, listed lists, stacks and queues. It will take you through several methods of finding efficient algorithms and show you which ones you should implement in each scenario. In addition to this, you will explore the possibilities of functional data structures using PHP and go through advanced algorithms and graphs as well as dynamic programming. By the end, you will be confident enough to tackle both basic and advanced data structures, understand how they work, and know when to use them in your day-to-day work Style and approach An easy-to-follow guide full of examples of implementation of data structures and real world examples to solve the problems faced. Each topic is first explained in general terms and then implemented using step by step explanation so that developers can understand each part of the discussion without any problem.

Getting Started with Beautiful Soup

A hands-on guide to building an enterprise-grade, scalable RESTful web service using the Spring Framework About This Book Follow best practices and explore techniques such as clustering and caching to achieve a scalable web service Leverage the Spring Framework to quickly implement RESTful endpoints Learn to implement a client library for a RESTful web service using the Spring Framework Who This Book Is For This book is intended for those who want to learn to build RESTful web services with the Spring Framework. To make best use of the code samples included in the book, you should have a basic knowledge of the Java language. Previous experience with the Spring Framework would also help you get up and running quickly. What You Will Learn Deep dive into the principles behind REST Expose CRUD operations through RESTful endpoints with the Spring Framework Devise response formats and error handling strategies, offering a consistent and flexible structure to simplify integration for service consumers Follow the best approaches for dealing with a service's evolution while maintaining backward compatibility Understand techniques to secure web services Comply with the best ways to test RESTful web services, including tips for load testing Optimise and scale web services using techniques such as caching and clustering In Detail REST is an architectural style that tackles the challenges of building scalable web services. In today's connected world, APIs have taken a central role on the web. APIs provide the fabric through which systems interact,

and REST has become synonymous with APIs. The depth, breadth, and ease of use of Spring makes it one of the most attractive frameworks in the Java ecosystem. Marrying the two technologies is therefore a very natural choice. This book takes you through the design of RESTful web services and leverages the Spring Framework to implement these services. Starting from the basics of the philosophy behind REST, you'll go through the steps of designing and implementing an enterprise-grade RESTful web service. Taking a practical approach, each chapter provides code samples that you can apply to your own circumstances. This book goes beyond the use of Spring and explores approaches to tackle resilience, security, and scalability concerns. You'll learn techniques to deal with security in Spring and discover how to implement unit and integration test strategies. Finally, the book ends by walking you through building a Java client for your RESTful web service, along with some scaling techniques for it.

Style and approach This book is a step-by-step, hands-on guide to designing and building RESTful web services. The book follows the natural cycle of developing these services and includes multiple code samples to help you.

Python Natural Language Processing

Learn the art of efficient web scraping and crawling with Python About This Book Extract data from any source to perform real time analytics. Full of techniques and examples to help you crawl websites and extract data within hours. A hands-on guide to

web scraping and crawling with real-life problems and solutions Who This Book Is For If you are a software developer, data scientist, NLP or machine-learning enthusiast or just need to migrate your company's wiki from a legacy platform, then this book is for you. It is perfect for someone , who needs instant access to large amounts of semi-structured data effortlessly. What You Will Learn Understand HTML pages and write XPath to extract the data you need Write Scrapy spiders with simple Python and do web crawls Push your data into any database, search engine or analytics system Configure your spider to download files, images and use proxies Create efficient pipelines that shape data in precisely the form you want Use Twisted Asynchronous API to process hundreds of items concurrently Make your crawler super-fast by learning how to tune Scrapy's performance Perform large scale distributed crawls with scrapyd and scrapinghub In Detail This book covers the long awaited Scrapy v 1.0 that empowers you to extract useful data from virtually any source with very little effort. It starts off by explaining the fundamentals of Scrapy framework, followed by a thorough description of how to extract data from any source, clean it up, shape it as per your requirement using Python and 3rd party APIs. Next you will be familiarised with the process of storing the scrapped data in databases as well as search engines and performing real time analytics on them with Spark Streaming. By the end of this book, you will perfect the art of scarping data for your applications with ease Style and approach It is a hands on guide, with first few chapters written as a tutorial, aiming to motivate you and get you started quickly. As the book

progresses, more advanced features are explained with real world examples that can be referred while developing your own web applications.

Instant Messaging in Java

Explores the philosophies behind the tiny house lifestyle, and provides guidance for the transition to a smaller space, including practical strategies for paring down possessions and design tricks to maximize function in available space.

Python Requests Essentials

Develop your JavaScript programming skills by learning strategies and techniques commonly used in modern full-stack application development Key Features Write and deploy full-stack applications efficiently with JavaScript Delve into JavaScript's multiple programming paradigms Get up to speed with core concepts such as modularity and functional programming to write efficient code Book Description In depth knowledge of JavaScript makes it easier to learn a variety of other frameworks, including React, Angular, and related tools and libraries. This book is designed to help you cover the core JavaScript concepts you need to build modern applications. You'll start by learning how to represent an HTML document in the Document Object Model (DOM). Then, you'll combine your knowledge of the DOM and Node.js to create a web scraper for practical situations. As you read through further lessons, you'll create a Node.js-based RESTful API using the Express

library for Node.js. You'll also understand how modular designs can be used for better reusability and collaboration with multiple developers on a single project. Later lessons will guide you through building unit tests, which ensure that the core functionality of your program is not affected over time. The book will also demonstrate how constructors, `async/await`, and events can load your applications quickly and efficiently. Finally, you'll gain useful insights into functional programming concepts such as immutability, pure functions, and higher-order functions. By the end of this book, you'll have the skills you need to tackle any real-world JavaScript development problem using a modern JavaScript approach, both for the client and server sides. What you will learn

- Apply the core concepts of functional programming
- Build a Node.js project that uses the Express.js library to host an API
- Create unit tests for a Node.js project to validate it
- Use the Cheerio library with Node.js to create a basic web scraper
- Develop a React interface to build processing flows
- Use callbacks as a basic way to bring control back

Who this book is for

If you want to advance from being a frontend developer to a full-stack developer and learn how Node.js can be used for hosting full-stack applications, this is an ideal book for you. After reading this book, you'll be able to write better JavaScript code and learn about the latest trends in the language. To easily grasp the concepts explained here, you should know the basic syntax of JavaScript and should've worked with popular frontend libraries such as jQuery. You should have also used JavaScript with HTML and CSS but not necessarily Node.js.

Web Scraping with Python

This book describes how to create Instant Messaging applications in Java and covers the Jabber IM protocols. If you want to create new IM systems, integrate them with your existing software, or wish to know more about the Jabber protocols, this is the book for you.

Learning Python Application Development

This book is a practical, hands-on guide that takes you through the techniques of web scraping using BeautifulSoup. Getting Started with BeautifulSoup is great for anybody who is interested in website scraping and extracting information. However, a basic knowledge of Python, HTML tags, and CSS is required for better understanding.

Automated Data Collection with R

Untangle your web scraping complexities and access web data with ease using Python scripts
Key Features
Hands-on recipes for advancing your web scraping skills to expert level
One-stop solution guide to address complex and challenging web scraping tasks using Python
Understand web page structures and collect data from a website with ease
Book Description
Python Web Scraping Cookbook is a solution-focused book that will teach you techniques to develop high-performance Scrapers, and deal with cookies, hidden form fields, Ajax-based sites and

proxies. You'll explore a number of real-world scenarios where every part of the development or product life cycle will be fully covered. You will not only develop the skills to design reliable, high-performing data flows, but also deploy your codebase to Amazon Web Services (AWS). If you are involved in software engineering, product development, or data mining or in building data-driven products, you will find this book useful as each recipe has a clear purpose and objective. Right from extracting data from websites to writing a sophisticated web crawler, the book's independent recipes will be extremely helpful while on the job. This book covers Python libraries, requests, and BeautifulSoup. You will learn about crawling, web spidering, working with AJAX websites, and paginated items. You will also understand to tackle problems such as 403 errors, working with proxy, scraping images, and LXML. By the end of this book, you will be able to scrape websites more efficiently and deploy and operate your scraper in the cloud. What you will learn Use a variety of tools to scrape any website and data, including Scrapy and Selenium Master expression languages, such as XPath and CSS, and regular expressions to extract web data Deal with scraping traps such as hidden form fields, throttling, pagination, and different status codes Build robust scraping pipelines with SQS and RabbitMQ Scrape assets like image media and learn what to do when Scraper fails to run Explore ETL techniques of building a customized crawler, parser, and convert structured and unstructured data from websites Deploy and run your scraper as a service in AWS Elastic Container Service Who this book is for This book is ideal for

Python programmers, web administrators, security professionals, and anyone who wants to perform web analytics. Familiarity with Python and basic understanding of web scraping will be useful to make the best of this book.

Go Web Scraping Quick Start Guide

If you've ever spent hours renaming files or updating hundreds of spreadsheet cells, you know how tedious tasks like these can be. But what if you could have your computer do them for you? In *Automate the Boring Stuff with Python*, you'll learn how to use Python to write programs that do in minutes what would take you hours to do by hand—no prior programming experience required. Once you've mastered the basics of programming, you'll create Python programs that effortlessly perform useful and impressive feats of automation to:

- Search for text in a file or across multiple files
- Create, update, move, and rename files and folders
- Search the Web and download online content
- Update and format data in Excel spreadsheets of any size
- Split, merge, watermark, and encrypt PDFs
- Send reminder emails and text notifications
- Fill out online forms

Step-by-step instructions walk you through each program, and practice projects at the end of each chapter challenge you to improve those programs and use your newfound skills to automate similar tasks. Don't spend your time doing work a well-trained monkey could do. Even if you've never written a line of code, you can make your computer do the grunt work. Learn how in *Automate the Boring Stuff with Python*.

Note: The programs in this book are written to run on Python 3.

Instant Jsoup How-To

Filled with practical, step-by-step instructions and clear explanations for the most important and useful tasks. Short, concise recipes to learn a variety of useful web scraping techniques using PHP. This book is aimed at those new to web scraping, with little or no previous programming experience. Basic knowledge of HTML and the Web is useful, but not necessary.

Jupyter for Data Science

As society continues to rely heavily on technological tools for facilitating business, e-commerce, banking, and communication, among other applications, there has been a significant rise in criminals seeking to exploit these tools for their nefarious gain. Countries all over the world are seeing substantial increases in identity theft and cyberattacks, as well as illicit transactions, including drug trafficking and human trafficking, being made through the dark web internet. Sex offenders and murderers explore unconventional methods of finding and contacting their victims through Facebook, Instagram, popular dating sites, etc., while pedophiles rely on these channels to obtain information and photographs of children, which are shared on hidden community sites. As criminals continue to harness technological advancements that are outpacing legal and ethical standards, law enforcement and government officials

are faced with the challenge of devising new and alternative strategies to identify and apprehend criminals to preserve the safety of society. The Encyclopedia of Criminal Activities and the Deep Web is a three-volume set that includes comprehensive articles covering multidisciplinary research and expert insights provided by hundreds of leading researchers from 30 countries including the United States, the United Kingdom, Australia, New Zealand, Germany, Finland, South Korea, Malaysia, and more. This comprehensive encyclopedia provides the most diverse findings and new methodologies for monitoring and regulating the use of online tools as well as hidden areas of the internet, including the deep and dark web. Highlighting a wide range of topics such as cyberbullying, online hate speech, and hacktivism, this book will offer strategies for the prediction and prevention of online criminal activity and examine methods for safeguarding internet users and their data from being tracked or stalked. Due to the techniques and extensive knowledge discussed in this publication it is an invaluable addition for academic and corporate libraries as well as a critical resource for policy makers, law enforcement officials, forensic scientists, criminologists, sociologists, victim advocates, cybersecurity analysts, lawmakers, government officials, industry professionals, academicians, researchers, and students within this field of study.

Python Web Scraping Cookbook

This book is full of short, concise recipes to learn a

variety of useful web scraping techniques using Java. You will start with a simple basic recipe of setting up your Java environment and gradually learn some more advanced recipes such as using complex Scrapers. Instant Web Scraping with Java is aimed at developers who, while not necessarily familiar with Java, are at least ready to dive into the complexities of this language with simple, step-by-step instructions leading the way. It is assumed that you have at least an intermediate knowledge of HTML, some knowledge of MySQL, and access to an Internet-connected computer while doing most of the exercises (after all, scraping the Web is difficult if your code can't get online!)

Building a RESTful Web Service with Spring

Examine the techniques and Java tools supporting the growing field of data science About This Book Your entry ticket to the world of data science with the stability and power of Java Explore, analyse, and visualize your data effectively using easy-to-follow examples Make your Java applications more capable using machine learning Who This Book Is For This book is for Java developers who are comfortable developing applications in Java. Those who now want to enter the world of data science or wish to build intelligent applications will find this book ideal. Aspiring data scientists will also find this book very helpful. What You Will Learn Understand the nature and key concepts used in the field of data science Grasp how data is collected, cleaned, and processed

Become comfortable with key data analysis techniques See specialized analysis techniques centered on machine learning Master the effective visualization of your data Work with the Java APIs and techniques used to perform data analysis In Detail Data science is concerned with extracting knowledge and insights from a wide variety of data sources to analyse patterns or predict future behaviour. It draws from a wide array of disciplines including statistics, computer science, mathematics, machine learning, and data mining. In this book, we cover the important data science concepts and how they are supported by Java, as well as the often statistically challenging techniques, to provide you with an understanding of their purpose and application. The book starts with an introduction of data science, followed by the basic data science tasks of data collection, data cleaning, data analysis, and data visualization. This is followed by a discussion of statistical techniques and more advanced topics including machine learning, neural networks, and deep learning. The next section examines the major categories of data analysis including text, visual, and audio data, followed by a discussion of resources that support parallel implementation. The final chapter illustrates an in-depth data science problem and provides a comprehensive, Java-based solution. Due to the nature of the topic, simple examples of techniques are presented early followed by a more detailed treatment later in the book. This permits a more natural introduction to the techniques and concepts presented in the book. Style and approach This book follows a tutorial approach, providing examples of each of the major concepts covered. With a step-by-

step instructional style, this book covers various facets of data science and will get you up and running quickly.

Programming Collective Intelligence

"Every developer working with the Web needs to read this book." -- David Heinemeier Hansson, creator of the Rails framework "RESTful Web Services finally provides a practical roadmap for constructing services that embrace the Web, instead of trying to route around it." -- Adam Trachtenberg, PHP author and EBay Web Services Evangelist You've built web sites that can be used by humans. But can you also build web sites that are usable by machines? That's where the future lies, and that's what RESTful Web Services shows you how to do. The World Wide Web is the most popular distributed application in history, and Web services and mashups have turned it into a powerful distributed computing platform. But today's web service technologies have lost sight of the simplicity that made the Web successful. They don't work like the Web, and they're missing out on its advantages. This book puts the "Web" back into web services. It shows how you can connect to the programmable web with the technologies you already use every day. The key is REST, the architectural style that drives the Web. This book: Emphasizes the power of basic Web technologies -- the HTTP application protocol, the URI naming standard, and the XML markup language Introduces the Resource-Oriented Architecture (ROA), a common-sense set of rules for designing RESTful web services Shows how a RESTful

design is simpler, more versatile, and more scalable than a design based on Remote Procedure Calls (RPC) Includes real-world examples of RESTful web services, like Amazon's Simple Storage Service and the Atom Publishing Protocol Discusses web service clients for popular programming languages Shows how to implement RESTful services in three popular frameworks -- Ruby on Rails, Restlet (for Java), and Django (for Python) Focuses on practical issues: how to design and implement RESTful web services and clients This is the first book that applies the REST design philosophy to real web services. It sets down the best practices you need to make your design a success, and the techniques you need to turn your design into working code. You can harness the power of the Web for programmable applications: you just have to work with the Web instead of against it. This book shows you how.

How to Borrow Your Way Out of Debt

A guide for Java programmers explains how to use XSLT's ability to provide platform-independent data to build Web-based applications incorporating transformations as well as interactive Web site and wireless services.

[ROMANCE](#) [ACTION & ADVENTURE](#) [MYSTERY & THRILLER](#) [BIOGRAPHIES & HISTORY](#) [CHILDREN'S](#) [YOUNG ADULT](#) [FANTASY](#) [HISTORICAL FICTION](#) [HORROR](#) [LITERARY FICTION](#) [NON-FICTION](#) [SCIENCE FICTION](#)