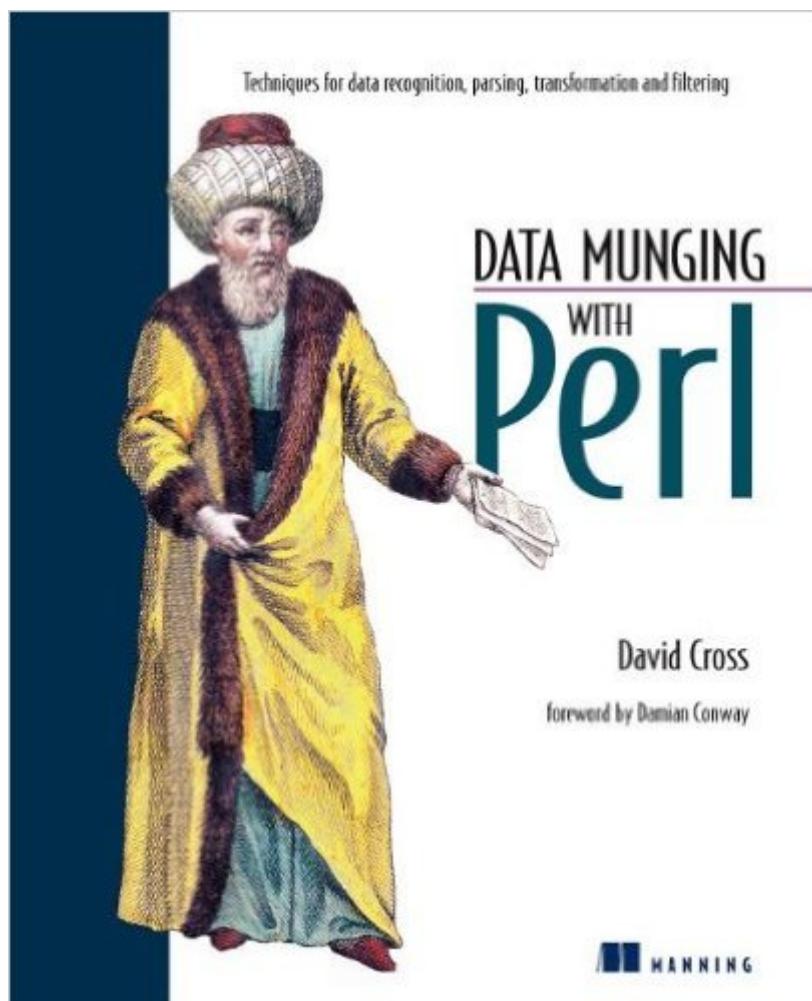


The book was found

# Data Munging With Perl



## **Synopsis**

Techniques for using Perl to recognize, parse, transform, and filter data.

## **Book Information**

Paperback: 300 pages

Publisher: Manning Publications (January 15, 2001)

Language: English

ISBN-10: 1930110006

ISBN-13: 978-1930110007

Product Dimensions: 7.4 x 0.6 x 9.2 inches

Shipping Weight: 1 pounds

Average Customer Review: 4.3 out of 5 stars [See all reviews](#) (13 customer reviews)

Best Sellers Rank: #1,254,808 in Books (See Top 100 in Books) #79 in Books > Computers & Technology > Databases & Big Data > Relational Databases #89 in Books > Computers & Technology > Programming > Languages & Tools > Perl #937 in Books > Computers & Technology > Databases & Big Data > Data Processing

## **Customer Reviews**

I was perusing the shelves at my local bookstore, when this title jumped out and grabbed me. Not only is it a unique and interesting title (Something uncommon for computer books), but it is also the most succinct description of my job, and of Perl for that matter, that I have ever seen. I am an avid reader, and in the interest of furthering my career most of my reading is work related. Normally, it takes me more than a month to read a book, though I am in the process of reading a half dozen or so at a time. I finished this book in less than a week. I couldn't put it down. The thing that is so great about this book, is that it delves into the heart of what Perl does best (And was designed to do).

Nine out of ten (more like ninety-nine out of one-hundred) jobs in Perl involve taking some sort of raw data, munging it, and spitting it out to some other process. This book is about doing that, many of the different forms that that can take, and some of the many techniques that perl (and a pragmatic approach) make available to do that. While I hesitate to say that this is the best Perl book I've ever read (It's got some good competition), I can say that there is no application of Perl that I am aware of where this book and the principles it explores would not be of value. In my opinion this book belongs on every serious Perl programmer's bookshelf right next to the "Camel" and the "Cookbook".

Dave Cross's new book, published by Manning, which means it has a figure from an old guide to native dress of the peoples of the world on the cover instead of some kind of animal, tells everything you need to know about using Perl for what it is most suited for: manipulating data. Starting with the source/filter/sink theory of data manipulation and demonstrating every tip and technique with clear and efficient examples, without severe digressions into mythological whimsy, this book would make an excellent second text on the Perl language, or a suitable first for someone who is good with programming languages. Many of the techniques contained in it are of "trade secret" quality; they are the sort of write-the-number-of-gallons-of-paint-it-took-to-paint-the-room-on-the-back-of-the-light-switch-cover practices that until now had to be learned or happened upon by every programmer, alone, or by example, rather than in the context of a coherent theory. The theoretical side, in which "munging" is defined and most software activity is described in terms of it, is clear enough that the book might be an interesting read for management, to answer the question "Just what is it about Perl that makes those who use it regularly so confoundedly fanatical?" If you've ever been mystified by a Perl wizard who found it easier to export the records from the fancy GUI database into a comma delimited text file and then sort and display the data with mysterious little programs rather than use the GUI's native report generator, and want to find out why, or if you would like to become such a person yourself, or if you already are such a person but would like to get better at it, this book is for you.

It's a guide. David takes you through the different "data munging" tasks ( record oriented data ? binary data ? fixed-width data ? XML ? ) and shows you his proper ways of dealing with them ( or, at least, thinking about them ). It's not an encyclopedia of "data munging", the book is 300 pages and many of them ( too many, may be ) are detailed descriptions of useful CPAN modules ( which I wasn't reading as careful as the rest of the book, since POD was always enough ), so it covers only a usual data processing tasks letting you to go deeper by yourself for more advanced topics. After you'll finish it much less "data sources" will scare you - the solutions and references are inside. As I said, it may be good for data-processing beginners, but Perl experts will hardly find lot's of new information in it. P.S. I trust him and therefore follow his advices in every script I start to think of ( especially the one about "UNIX filter model" ).

This book isn't about arcane corners of Perl theory. It's about how to write Perl programs that perform the "simple" task of converting data from one format to another. Need to get every headline from an RSS feed? Or report the three users with the most processes running, as listed by `ps`? Or

extract the first paragraph from each of a thousand HTML files? Or make a .tsv file based on all the "From:" and "Subject:" lines in your mailbox file? If those sorts of tasks sound familiar to you, then this is the book you've been looking for. It has working code for doing these sorts of things, involving lots of different common kinds of formats. By tech book standards, this book is short (300 pages), but it's clear and direct and to the point -- no bloat here. Every page tells you something you need to know, with useful examples for every idea that it explains.

This book, written by Perlmonk [...] David Cross, is an excellent, easy to read, and easy to follow guide into what Perl does best: Data Munging. For those who don't know, Munging Data means taking data from one format and putting it into another. Perl excels at this, and the author shows you the how and the why. The author gives you enough information, and background to start working with the more advanced Perl functions like map, grep, pack, unpack, etc. It is possible to write Perl without ever having to use these modules, but David Cross shows you how they are more effective, more powerful. This book will expand your Perl vocabulary by leaps and bounds. I know that some people would say that the book is too thin, and it is thinner than many computer books today, but the thickness of a book does not determine its merit. Effective Perl Programming by Joseph Hall and Randal Schwartz is often cited as one of the best Perl books ever and it's thinner than this one. If you are a junior to intermediate level programmer, and you want to improve your Perl skills, pick up this book. You won't be disappointed.

[Download to continue reading...](#)

Data Munging with Perl Data Analytics: What Every Business Must Know About Big Data And Data Science (Data Analytics for Business, Predictive Analysis, Big Data) Data Analytics: Practical Data Analysis and Statistical Guide to Transform and Evolve Any Business. Leveraging the Power of Data Analytics, Data ... (Hacking Freedom and Data Driven) (Volume 2) Analytics: Data Science, Data Analysis and Predictive Analytics for Business (Algorithms, Business Intelligence, Statistical Analysis, Decision Analysis, Business Analytics, Data Mining, Big Data) Ruby: Programming, Master's Handbook: A TRUE Beginner's Guide! Problem Solving, Code, Data Science, Data Structures & Algorithms (Code like a PRO in ... web design, tech, perl, ajax, swift, python,) Java Programming: Master's Handbook: A TRUE Beginner's Guide! Problem Solving, Code, Data Science, Data Structures & Algorithms (Code like a PRO in ... web design, tech, perl, ajax, swift, python) Perl: Crash Course - The Ultimate Beginner's Course to Learning Perl Programming in Under 12 Hours Effective Perl Programming: Ways to Write Better, More Idiomatic Perl (2nd Edition) (Effective Software Development Series) Mastering Perl/Tk: Graphical User Interfaces in Perl Perl

Programming Success in a Day: Beginners Guide to Fast, Easy, and Efficient Learning of Perl  
Programming Programming the Perl DBI: Database programming with Perl Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data (Data-Centric Systems and Applications) Rsmeans Assemblies Cost Data: Assemblies Cost Data Data and Goliath: The Hidden Battles to Capture Your Data and Control Your World Efficient SAP R/3-Data Archiving: How to Handle Large Data Volumes Big Data Appliances for In-Memory Computing: A Real-World Research Guide for Corporations to Tame and Wrangle Their Data Professional Java Data: RDBMS, JDBC, SQLJ, OODBMS, JNDI, LDAP, Servlets, JSP, WAP, XML, EJBs, CMP2.0, JDO, Transactions, Performance, Scalability, Object and Data Modeling Data Structure and Algorithmic Thinking with Python: Data Structure and Algorithmic Puzzles

[Dmca](#)