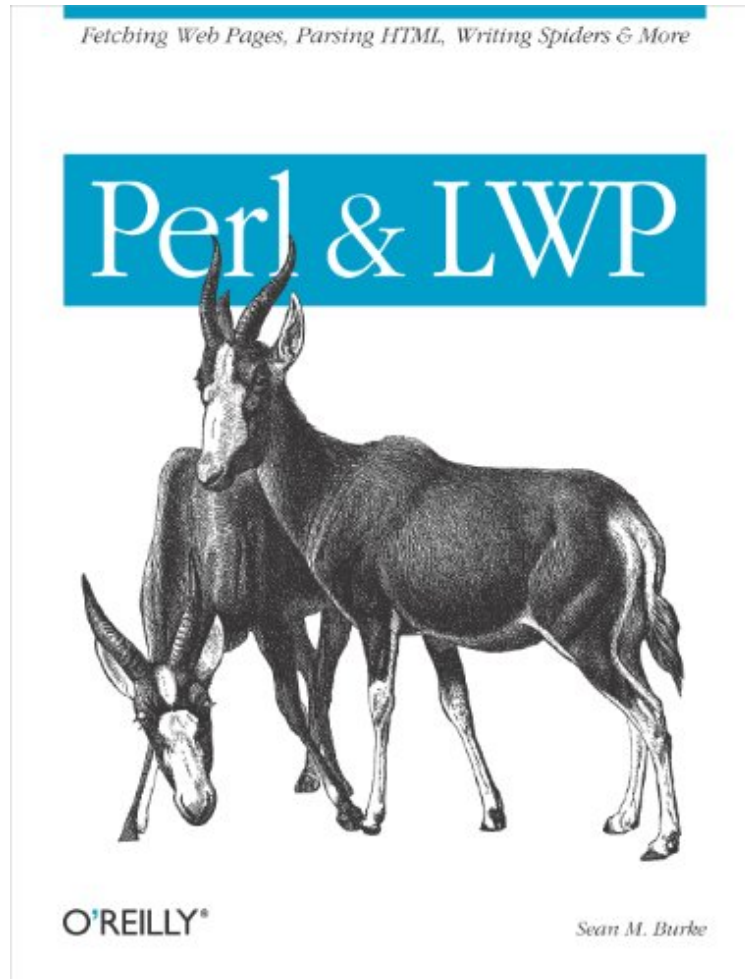


[Library ebook] Perl LWP: Fetching Web Pages, Parsing HTML, Writing Spiders More

Perl LWP: Fetching Web Pages, Parsing HTML, Writing Spiders More

Von Sean M. Burke

ePub | *DOC | audiobook | ebooks | Download PDF



DOWNLOAD



READ ONLINE

Produktinformation -Verkaufs-rang: #1318838 in eBooksVerffentlicht am: 2002-06-20Erscheinungsdatum: 2009-02-09File Name: B0026REAB0 | File size: 18.Mb

Von Sean M. Burke : Perl LWP: Fetching Web Pages, Parsing HTML, Writing Spiders More before purchasing it in order to gage whether or not it would be worth my time, and all praised Perl LWP: Fetching Web Pages, Parsing HTML, Writing Spiders More:

KundenrezensionenHilfreichste Kundenrezensionen9 von 9 Kunden fanden die folgende Rezension hilfreich. Perl zum Webseiten absaugen und analysierenVon Gregor TheisDas Buch Perl LWP" von Sean S. Burke beschreibt die Verwendung von Perl und der Perl LWP (libwww-perl) Bibliothek und einiger Zusatzmodule zum Herunterladen von Web Seiten mit Perl und zur anschließenden Analyse dieser Daten. Mit diesen Modulen kann man Programme schreiben, die Daten aus dem Web sammeln und diese nach eigenen Wnschen aufbereiten.HTML Seiten aus dem Internet zu laden ist verhlnismig einfach mit LWP. Schwieriger ist diese Seiten zu bearbeiten und die gewünschten

Daten zu extrahieren. Sean Burke geht sehr intensiv auf die Daten Extraktion ein und bemht sich uns ein robustes Vorgehen bei der Programmierung beizubringen (bei der Daten Analyse vor allem defensiv programmieren).Man wird nach diesem Buch nicht vollstndig um das Studium der Dokumentation der einzelnen Module herumkommen (z.B. um weitere Informationen zur Verwendung von Proxies mit Passwort zu sammeln). Dieses Buch ist ein Einfhrungsbuch in das Thema, nachdem man schon viele Dinge mit LWP machen kann und die Online Dokumentation verstehen kann.Sehr schn ist der Anhang G mit einem Abdruck eines Artikels von Sean S Burke zum Thema objektorientierte Module verwenden", was einen guten Einblick in die Verwendung von OO Modulen gibt.Ich wrde diesem Buch gerne 4,5 Sterne geben, da ich es fr ein sehr gutes Perl Buch halte aber es meiner Meinung nach nicht so gut ist wie Einfhrung in Perl" (bzw. Learning Perl"). Der Stil von "Perl LWP" ist auch sehr locker und gut zu Lesen.Leider gibt es in dieser Auflage eine Reihe von Fehlern im Druck. Man sollte sich die Errata auf dem Seite von O'Reilly zu diesem Buch ansehen, bevor man zu frustriert ist. Die meisten Beispiele in diesem Buch funktionieren nicht mehr, da sich diese Beispiele auf reale Seiten im World Wide Web zugreifen und sich diese schon strukturell verndert haben, so dass die Analyse nicht mehr funktioniert. Wenn man diese Beispiele debugged, wird man eine Menge zu LWP und den Randmodulen lernen.

KurzbeschreibungPerl soared to popularity as a language for creating and managing web content, but with LWP (Library for WWW in Perl), Perl is equally adept at consuming information on the Web. LWP is a suite of modules for fetching and processing web pages.The Web is a vast data source that contains everything from stock prices to movie credits, and with LWP all that data is just a few lines of code away. Anything you do on the Web, whether it's buying or selling, reading or writing, uploading or downloading, news to e-commerce, can be controlled with Perl and LWP. You can automate Web-based purchase orders as easily as you can set up a program to download MP3 files from a web site.Perl LWP covers:Understanding LWP and its designFetching and analyzing URLsExtracting information from HTML using regular expressions and tokensWorking with the structure of HTML documents using treesSetting and inspecting HTTP headers and response codesManaging cookiesAccessing information that requires authenticationExtracting linksCooperating with proxy cachesWriting web spiders (also known as robots) in a safe fashionPerl LWP includes many step-by-step examples that show how to apply the various techniques. Programs to extract information from the web sites of BBC News, Altavista, ABEBooks.com, and the Weather Underground, to name just a few, are explained in detail, so that you understand how and why they work.Perl programmers who want to automate and mine the web can pick up this book and be immediately productive. Written by a contributor to LWP, and with a foreword by one of LWP's creators, Perl LWP is the authoritative guide to this powerful and popular toolkit.