Webbots, Spiders, and Screen Scrapers
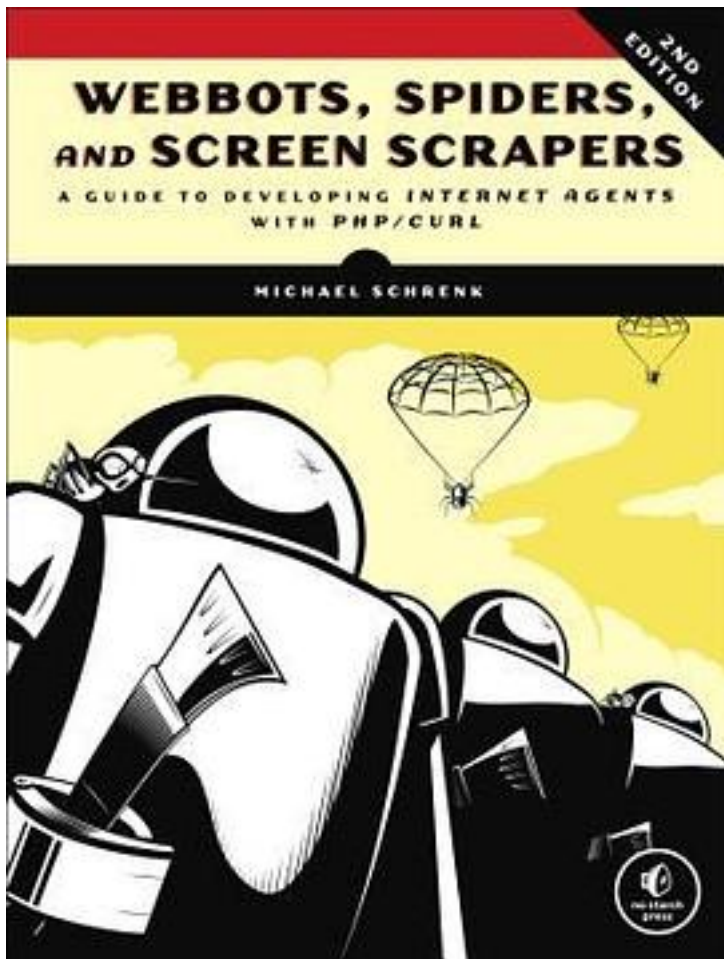

[Webbots, Spiders, and Screen Scrapers_下载链接1_](#)

著者:Michael Schrenk

出版者:No Starch Press

出版时间:2007-03-30

装帧:Paperback

isbn:9781593271206

The Internet is bigger and better than what a mere browser allows. Webbots, Spiders, and Screen Scrapers is for programmers and businesspeople who want to take full advantage of the vast resources available on the Web. There's no reason to let

browsers limit your online experience-especially when you can easily automate online tasks to suit your individual needs.

Learn how to write webbots and spiders that do all this and more:

Programmatically download entire websites

Effectively parse data from web pages

Manage cookies

Decode encrypted files

Automate form submissions

Send and receive email

Send SMS alerts to your cell phone

Unlock password-protected websites

Automatically bid in online auctions

Exchange data with FTP and NNTP servers

Sample projects using standard code libraries reinforce these new skills. You'll learn how to create your own webbots and spiders that track online prices, aggregate different data sources into a single web page, and archive the online data you just can't live without. You'll learn inside information from an experienced webbot developer on how and when to write stealthy webbots that mimic human behavior, tips for developing fault-tolerant designs, and various methods for launching and scheduling webbots. You'll also get advice on how to write webbots and spiders that respect website owner property rights, plus techniques for shielding websites from unwanted robots.

As a bonus, visit the author's website to test your webbots on sample target pages, and to download the scripts and code libraries used in the book.

Some tasks are just too tedious-or too important!- to leave to humans. Once you've automated your online life, you'll never let a browser limit the way you use the Internet again.

作者介绍:

目录:

Webbots, Spiders, and Screen Scrapers_下载链接1_

标签

网络爬虫

互联网

编程

programming

Spider

计算机

爬虫

数据挖掘

评论

读了大概三分之一，没看到亮点，把简单的东西故弄玄虚倒有点，浪费了这么好的封面和标题啊

故弄玄虚，技术过时

----
php

Webbots, Spiders, and Screen Scrapers_下载链接1

书评

[Webbots, Spiders, and Screen Scrapers_下载链接1_](#)