Book Review: Lucene In Action

Posted At : 15 August 2008 16:08 | Posted By : Gareth
Related Categories: lucene, book review, java

I first heard about Lucene a few years ago, as it's used to power the search functionality in *BlueDragon*. Although that makes it sound like an open-source version of the *Verity Search Engine* software, it is so much more. I've been using Lucene on my current project and have turned into a big fan. I bought the book *Lucene In Action* to help, and I've been equally impressed with the book.

If you're new to Lucene, it is technically known as *Information Retrieval (IR)* software. If that sounds scary, it's used to build search engines. You'll find implmentations of it in many application servers - BlueDragon, Ruby on Rails, PHP, Java and more. It started as a Java project, but there are now implementations in other languages such as .NET.

Oops!

Message

Detail

Extended Info

Tag Context

You're most likely to use it to build a site search facility in your website, but there are so many other applications. This book has case studies on some of the varied uses Lucene has been put to. The most impressive study was of a website which sold art prints, how they used Lucene to index the colours in the prints, and how this was used to suggest similar products.

Lucene is very simple to get up and running with, and there are loads of tutorials on the web. As with all projects though, clients always want something special, and this book show you how to get the most out.

Lucene In Action explains how Lucene works under the hood, how it all fits together and how to extend it to add your own functionality. You don't need a PhD in Information Retrieval, as this book explains the basics of the science and how it works in Lucene.

It's not all theory either. The book is littered with practical examples, all in the form of *JUnit* test cases. Many of the examples, such as *Stemming* and *Metaphone replacement*, I've used directly.

This book was printed over 3 years ago, and unfortunately, Lucene has moved on a lot since then. I can't blame the book for that, but it did mean that I struggled a bit trying to implement some of the new Lucene features, such as payloads. The good news however, is that the 2nd edition is due out in early 2009, and covers all of this.

It's rare that I've been so impressed with a book. My current project has required some in-depth customisation with Lucene. Given that I was a new-comer to the subject, I don't know how I would have managed without this book. The only reason it doesn't get 10/10 is that it's aged a little - that should change with the 2nd edition.

Author	Erik Hatcher and Otis Gospodnetic
Publisher	Manning Publications (January 2005)
ISBN	978-1932394283
Rating	9/10