



Learn by doing: less theory, more results

PHP and MongoDB Web Development

Combine the power of PHP and MongoDB to build
dynamic web 2.0 applications

Beginner's Guide

Rubayeet Islam

[PACKT] open source*
PUBLISHING community experience distilled

PHP and MongoDB Web Development

Beginner's Guide

Combine the power of PHP and MongoDB to build dynamic web 2.0 applications

Rubayeet Islam

[PACKT] open source 
PUBLISHING community experience distilled

BIRMINGHAM - MUMBAI

PHP and MongoDB Web Development

Beginner's Guide

Copyright © 2011 Packt Publishing

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, without the prior written permission of the publisher, except in the case of brief quotations embedded in critical articles or reviews.

Every effort has been made in the preparation of this book to ensure the accuracy of the information presented. However, the information contained in this book is sold without warranty, either express or implied. Neither the author, nor Packt Publishing, and its dealers and distributors will be held liable for any damages caused or alleged to be caused directly or indirectly by this book.

Packt Publishing has endeavored to provide trademark information about all of the companies and products mentioned in this book by the appropriate use of capitals. However, Packt Publishing cannot guarantee the accuracy of this information.

First published: November 2011

Production Reference: 1181111

Published by Packt Publishing Ltd.
Livery Place
35 Livery Street
Birmingham B3 2PB, UK.

ISBN 978-1-84951-362-3

www.packtpub.com

Cover Image by Charwak A (charwak86@gmail.com)

Credits

Author

Rubayeet Islam

Reviewers

Sam Millman

Sigert de Vries

Nurul Ferdous

Vidyasagar N V

Acquisition Editor

Usha Iyer

Development Editor

Susmita Panda

Technical Editors

Joylita D'Souza

Veronica Fernandes

Lubna Shaikh

Copy Editor

Laxmi Subramanian

Project Coordinator

Kushal Bhardwaj

Proofreader

Matthew Humphries

Indexer

Tejal Daruwale

Graphics

Valentina D'silva

Production Coordinator

Prachali Bhiwandkar

Cover Work

Prachali Bhiwandkar

About the Author

Rubayeeet Islam is a Software Developer with over 4 years of experience in large-scale web application development on open source technology stacks (LAMP, Python/Django, Ruby on Rails). He is currently involved in developing cloud-based distributed software that use MongoDB as their analytics and metadata backend. He has also spoken in seminars to promote the use of MongoDB and NoSQL databases in general. He graduated from the University of Dhaka with a B.S. in Computer Science and Engineering.

I thank the Almighty for giving me such a blessed life and my parents for letting me follow my passion. My friend and colleague, Nurul Ferdous, for inspiring me to be an author in the first place. Finally, the amazing people at Packt – Usha Iyer, Kushal Bhardwaj, Priya Mukherji, and Susmita Panda, without your help and guidance this book would not have been possible to write.

About the Reviewers

Sam Millman, after achieving a B.Sc. in Computing from the University of Plymouth, immediately moved to advance his knowledge within Web development, specifically PHP. He is a fully self-taught professional Web Developer and IT Administrator working for a company in the south of England.

He first started to show an interest in MongoDB when he went in search of something new to learn. Now he is an active user of the MongoDB Google User Group and is about to release a new site written in PHP with MongoDB as the primary data store.

Sigert de Vries (1983) is a professional Web Developer working in The Netherlands. He has worked in several companies as a System Administrator and Web Developer. He is a specialist in high performance websites and is an open source enthusiast. With his communicative skills, he translates advanced technical issues to "normal" human language.

Sigert is currently working at worldticketshop.com, helping them to be one of the largest ticket marketplaces in Europe. Within the company, there's plenty of room to use NoSQL solutions such as MongoDB.

I would like to thank Packt publishing for asking me to review this book, it has been a pleasure!

Vidyasagar N V was interested in Computer Science since an early age. Some of his serious work in computers and computer networks began during his high school days. Later, he went to the prestigious Institute Of Technology, Banaras Hindu University for his B.Tech. He has been working as a Software Developer and Data Expert, developing and building scalable systems. He has worked with a variety of 2nd, 3rd, and 4th generation languages. He has worked with flat files, indexed files, hierarchical databases, network databases, relational databases, NoSQL databases, Hadoop, and related technologies. Currently, he is working as a Senior Developer at Ziva Software Pvt. Ltd., developing big database-structured data-extraction techniques for the Web and local information. He enjoys producing high-quality software, web-based solutions, and designing secure and scalable data systems.

I would like to thank my parents, Mr. N Srinivasa Rao and Mrs. Latha Rao, and my family who supported and backed me throughout my life. My friends for being friends, and all those people willing to donate their time, effort, and expertise by participating in open source software projects. Thank you Packt Publishing for selecting me as one of the technical reviewers on this wonderful book. It is my honor to be a part of this book. You can contact me at vidyasagar1729@gmail.com.

www.PacktPub.com

Support files, eBooks, discount offers and more

You might want to visit www.PacktPub.com for support files and downloads related to your book.

Did you know that Packt offers eBook versions of every book published, with PDF and ePub files available? You can upgrade to the eBook version at www.PacktPub.com and as a print book customer, you are entitled to a discount on the eBook copy. Get in touch with us at service@packtpub.com for more details.

At www.PacktPub.com, you can also read a collection of free technical articles, sign up for a range of free newsletters and receive exclusive discounts and offers on Packt books and eBooks.



<http://PacktLib.PacktPub.com>

Do you need instant solutions to your IT questions? PacktLib is Packt's online digital book library. Here, you can access, read and search across Packt's entire library of books.

Why Subscribe?

- Fully searchable across every book published by Packt
- Copy and paste, print and bookmark content
- On demand and accessible via web browser

Free Access for Packt account holders

If you have an account with Packt at www.PacktPub.com, you can use this to access PacktLib today and view nine entirely free books. Simply use your login credentials for immediate access.

Table of Contents

Preface	1
Chapter 1: Getting Started with MongoDB	7
The NoSQL movement	8
Types of NoSQL databases	8
MongoDB – A document-based NoSQL database	9
Why MongoDB?	9
Who is using MongoDB?	9
MongoDB concepts—Databases, collections, and documents	10
Anatomy of document	10
BSON—The data exchange format for MongoDB	11
Similarity with relational databases	11
Downloading, installing, and running MongoDB	12
System requirements	12
Time for action – downloading and running MongoDB on Windows	13
Installing the 64-bit version	14
Time for action – downloading and running MongoDB on Linux	15
Installing MongoDB on OS X	17
Configuring MongoDB	17
Command-line parameters	17
File-based configuration	18
Stopping MongoDB	18
Hitting Control + C	18
From the mongo shell	19
Sending INT or TERM signal in UNIX	19
Creating databases, collections, and documents	19
Time for Action – creating databases, collections, and documents	19
Installing the PHP driver for MongoDB	21
Time for Action – installing PHP driver for MongoDB on Windows	21
Installing the PHP-MongoDB driver on Unix	23

Connecting to the MongoDB server from PHP	23
Creating a PHP-Mongo connection	23
Time for action – creating a connection to the MongoDB server from PHP	24
Configuring the PHP-MongoDB connection	26
Specifying timeout for the connection attempt	26
Summary	27
Chapter 2: Building your First MongoDB Powered Web App	29
A MongoDB powered blog	30
Have the MongoDB server running	30
Inserting documents in MongoDB	30
Time for action – building the Blog Post Creator	30
Creating databases and collections implicitly	35
Performing 'safe' inserts	35
Benefits of safe inserts	36
Specifying a timeout on insert	36
Setting the user generated _id	37
The MongoDate object	37
Querying documents in a collection	38
Time for action – retrieving articles from a database	38
The Mongo Query Language	42
The MongoClient object	42
Conditional Queries	44
Doing advanced queries in MongoDB	45
Time for action – building the Blog Dashboard	45
Returning a subset of fields	49
Sorting the query results	49
Using count, skip, and limit	49
Performing range queries on dates	50
Updating documents in MongoDB	51
Time for action – building the Blog Editor	51
Optional arguments to the update method	55
Performing 'upsert'	55
Using update versus using save	56
Using modifier operations	56
Setting with \$set	56
Incrementing with \$inc	57
Deleting fields with \$unset	57
Renaming fields with \$rename	57
Deleting documents in MongoDB	58
Time for action – deleting blog posts	58
Optional arguments to remove	63

Managing relationships between documents	63
Embedded documents	64
Referenced documents	64
Time for action – posting comments to blog posts	64
Embedded versus referenced – Which one to use?	69
Querying embedded objects	69
Summary	71
Chapter 3: Building a Session Manager	73
Understanding HTTP sessions	74
Understanding PHP native session handling	74
Time for action – testing native PHP session handling	74
Limitations of native PHP session handling	78
Implementing session handling with MongoDB	78
Extending session handling with session_set_save_handler	78
The SessionManager class	79
Time for action – building the SessionManager class	79
How the SessionManager works	83
The constructor	83
The open and close methods	84
The read method	84
The write method	84
The destroy method	84
The gc method	85
Putting the SessionManager in action	85
Time for action – putting SessionManager into action	86
Building the user authentication module	88
Time for action – building the User class	89
Creating the login, logout, and user profile page	92
Time for action – creating the login, logout, and profile page	93
Using good session practices	99
Setting low expiry times of session cookies	99
Using session timeouts	100
Setting proper domains for session cookies	100
Checking for browser consistency	100
Summary	101
Chapter 4: Aggregation Queries	103
Generating sample data	104
Time for action – generating sample data	104
Understanding MapReduce	107
Visualizing MapReduce	108
Performing MapReduce in MongoDB	109

Time for action – counting the number of articles for each author	110
Defining the Map function	111
Defining the Reduce function	112
Applying the Map and Reduce	112
Viewing the results	113
Performing MapReduce on a subset of the collection	114
Concurrency	114
Performing MongoDB MapReduce within PHP	114
Time for action – creating a tag cloud	115
Performing aggregation using group()	120
Time for action – calculating the average rating per author	121
Grouping by custom keys	124
MapReduce versus group()	124
Listing distinct values for a field	125
Time for action – listing distinct categories of articles	125
Using distinct() in mongo shell	127
Counting documents with count()	127
Summary	128
Chapter 5: Web Analytics using MongoDB	129
Why MongoDB is a good choice as a web analytics backend	130
Logging with MongoDB	131
Time for action – logging page visits with MongoDB	131
Capped collections	134
Sorting in natural order	135
Updating and deleting documents in a capped collection	135
Specifying the size of a regular collection	135
Convert a regular collection to a capped one	136
Extracting analytics data with MapReduce	136
Time for action – finding total views and average response time per blog post	137
The map, reduce, and finalize functions	140
Displaying the result	140
Running MapReduce in real time versus running it in the background	141
Real-time analytics using MongoDB	141
Time for action – building a real-time page visit counter	141
Summary	146
Chapter 6: Using MongoDB with Relational Databases	147
The motivation behind using MongoDB and an RDBMS together	148
Potential use cases	148
Defining the relational model	149
Time for action – creating the database in MySQL	150

Caching aggregation results in MongoDB	153
Time for action – storing the daily sales history of products in MongoDB	153
Benefits of caching queries in MongoDB	160
Storing results of expensive JOINS	160
Using MongoDB for data archiving	160
Time for action – archiving old sales records in MongoDB	161
Challenges in archiving and migration	163
Dealing with foreign key constraints	163
Preserving data types	163
Storing metadata in MongoDB	164
Time for action – using MongoDB to store customer metadata	164
Problems with using MongoDB and RDBMS together	173
Summary	173
Chapter 7: Handling Large Files with GridFS	175
What is GridFS?	175
The rationale of GridFS	176
The specification	176
Advantages over the filesystem	177
Storing files in GridFS	178
Time for action – uploading images to GridFS	178
Looking under the hood	181
Serving files from GridFS	182
Time for action – serving images from GridFS	183
Updating metadata of a file	186
Deleting files	186
Reading files in chunks	187
Time for action – reading images in chunks	187
When should you not use GridFS	189
Summary	190
Chapter 8: Building Location-aware Web Applications with MongoDB and PHP	191
A geolocation primer	192
Methods to determine location	192
Detecting the location of a web page visitor	193
The W3C Geolocation API	193
Browsers that support geolocation	194
Time for action – detecting location with W3C API	194
The Geolocation object	198
The <code>getCurrentPosition()</code> method	198
Drawing the map using the Google Maps API	199

Geospatial indexing	200
Time for action – creating geospatial indexes	201
Geospatial indexing – Important things to know	202
Performing location queries	203
Time for action – finding restaurants near your location	203
The geoNear() command	208
Bounded queries	210
Geospatial haystack indexing	210
Time for action – finding nearby restaurants that serve burgers	211
Summary	215
Chapter 9: Improving Security and Performance	217
Enhancing query performance using indexes	217
Time for action – creating an index on a MongoDB collection	218
The _id index	221
Unique indexes	221
Compound keys indexes	222
Indexing embedded document fields	223
Indexing array fields	224
Deleting indexes	224
When indexing cannot be used	224
Indexing guidelines	225
Choose the keys wisely	225
Keep an eye on the index size	225
Avoid using low-selectivity single key indexes	225
Be aware of indexing costs	226
On a live database, run indexing in the background	226
Optimizing queries	227
Explaining queries using explain()	227
Optimization rules	228
Using hint()	228
Profiling queries	229
Understanding the output	229
Optimization rules	230
Securing MongoDB	230
Time for action – adding user authentication in MongoDB	230
Creating an admin user	232
Creating regular user	233
Viewing, changing, and deleting user accounts	233
User authentication through PHP driver	234
Filtering user input	235
Running MongoDB server in a secure environment	235

Ensuring data durability	236
Journaling	236
Performance	237
Using fsync	237
Replication	238
Summary	239
Chapter 10: Easy MongoDB Administration with RockMongo and phpMoAdmin	241
<hr/>	
Administering MongoDB with RockMongo	242
Time for action – installing RockMongo on your computer	242
Exploring data with RockMongo	244
Querying	245
Updating, deleting, and creating documents	245
Importing and exporting data	247
Viewing stats	248
Miscellaneous	248
Using phpMoAdmin to administer MongoDB	249
Time for action – installing phpMoAdmin on your computer	249
Viewing databases and collections	250
Querying documents	251
Saving and deleting objects	252
Importing and exporting data	252
Viewing stats	253
Other features	253
RockMongo versus phpMoAdmin	254
The verdict	255
Summary	256
Pop Quiz Answers	257
<hr/>	
Chapter 1, Getting Started with MongoDB	257
Chapter 2, Building your First MongoDB Powered Web App	257
Chapter 3, Building a Session Manager	258
Chapter 4, Aggregation Queries	258
Chapter 5, Web Analytics using MongoDB	258
Chapter 7, Handling Large Files with GridFS	259
Chapter 8, Building Location-aware Web Applications with MongoDB and PHP	259
Chapter 9, Improving Security and Performance	259
Index	237

Preface

MongoDB is an open source, non-relational database system designed to meet the needs of modern Web 2.0 applications. It is currently being used by some of the most popular websites in the world. This book introduces MongoDB to the web developer who has some background building web applications using PHP. This book teaches what MongoDB is, how it is different from relational database management systems, and when and why developers should use it instead of a relational database for storing data.

You will learn how to build PHP applications that use MongoDB as the data backend; solve common problems, such as HTTP session handling, user authentication, and so on.

You will also learn to solve interesting problems with MongoDB, such as web analytics with MapReduce, storing large files in GridFS, and building location-aware applications using Geospatial indexing.

Finally, you will learn how to optimize MongoDB to boost performance, improve security, and ensure data durability. The book will demonstrate the use of some handy GUI tools that makes database management easier.

What this book covers

Chapter 1, Getting Started with MongoDB introduces the underlying concepts of MongoDB, provides a step-by-step guide on how to install and run a MongoDB server on a computer, and make PHP and MongoDB talk to each other.

Chapter 2, Building your First MongoDB Powered Web App shows you how to build a simple blogging web application using PHP and MongoDB. Through the examples in this chapter, you will learn how to create/read/update/delete data in MongoDB using PHP.

Chapter 3, Building a Session Manager shows you how PHP and MongoDB can be used to handle HTTP sessions. You will build a stand-alone session manager module and learn how to perform user authentication/authorization using the module.

Chapter 4, Aggregation Queries introduces MapReduce, a powerful functional programming paradigm and shows you how it can be used to perform aggregation queries in MongoDB.

Chapter 5, Web Analytics using MongoDB shows you how you can store website traffic data in MongoDB in real time and use MapReduce to extract important analytics.

Chapter 6, Using MongoDB with Relational Databases explores use cases where MongoDB can be used alongside a relational database. You will learn how to archive data in MongoDB, use it for caching expensive query results, and store non-structured metadata about different objects in the domain.

Chapter 7, Handling Large Files with GridFS introduces GridFS, a specification in MongoDB that allows us to store large files in the database.

Chapter 8, Building Location-aware Web Applications with MongoDB and PHP, uses PHP, HTML5, JavaScript, and the Geospatial Indexing feature of MongoDB to build a web application that helps you find restaurants close to your current location.

Chapter 9, Improving Security and Performance shows you how to boost query performance using indexes, use built-in tools for analyzing and fine-tuning queries, improve database security, and ensure data durability.

Chapter 10, Easy MongoDB Administration with RockMongo and phpMoAdmin demonstrates the use of a couple of PHP-based GUI tools for managing MongoDB server—RockMongo and phpMoAdmin.

What you need for this book

Apache web server (or IIS if you are on Windows) running PHP 5.2.6 or higher.

A web browser that supports the W3C Geolocation API (Internet Explorer 9.0+, Google Chrome 5.0+, Firefox 3.5+ or Safari 5.0+).

Chapter 6, Using MongoDB with Relational Databases requires that you have MySQL installed on your machine.

Who this book is for

This book assumes that you have some background in web application development using PHP, HTML, and CSS. Some of the chapters require that you know JavaScript and are familiar with AJAX. Having a working knowledge of using a relational database system, such as MySQL will help you grasp some of the concepts quicker, but it is not strictly mandatory. No prior knowledge of MongoDB is required.

Conventions

In this book, you will find several headings appearing frequently.

To give clear instructions of how to complete a procedure or task, we use:

Time for action – heading

1. Action 1
2. Action 2
3. Action 3

Instructions often need some extra explanation so that they make sense, so they are followed with:

What just happened?

This heading explains the working of tasks or instructions that you have just completed.

You will also find some other learning aids in the book, including:

Pop quiz – heading

These are short multiple choice questions intended to help you test your own understanding.

Have a go hero – heading

These set practical challenges and give you ideas for experimenting with what you have learned.

You will also find a number of styles of text that distinguish between different kinds of information. Here are some examples of these styles, and an explanation of their meaning.

Code words in text are shown as follows: "The value for the first field, `_id`, is autogenerated."

A block of code is set as follows:

```
try {
    $mongo = new Mongo($options=array('timeout'=> 100))
} catch(MongoConnectionException $e) {
    die("Failed to connect to database ".$e->getMessage());
}
```

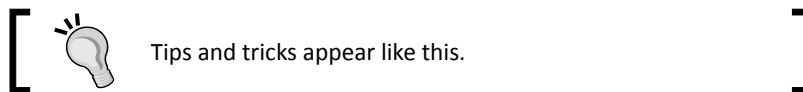
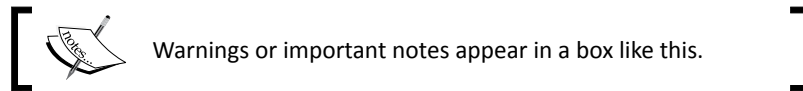
When we wish to draw your attention to a particular part of a code block, the relevant lines or items are set in bold:

```
{
  _id      : ObjectId("4dcd2abe5981aec801010000"),
  title    : "The only perfect site is hind-site",
  content  : "Loren ipsum dolor sit amet...",
  saved_at : ISODate('2011-05-16T18:42:57.949Z'),
  author_id : ObjectId("4dd491695072aefc456c9aca")
}
```

Any command-line input or output is written as follows:

```
>db.movies.find({"genre":"sci-fi"})
{ "_id" : ObjectId("4db439153ec7b6fd1c9093ec"), "name" : "Source Code",
  "genre" : "sci-fi", "year" : 2011 }
```

New terms and **important words** are shown in bold. Words that you see on the screen, in menus or dialog boxes for example, appear in the text like this: "Click on the **Delete** link on any one article."



Reader feedback

Feedback from our readers is always welcome. Let us know what you think about this book—what you liked or may have disliked. Reader feedback is important for us to develop titles that you really get the most out of.

To send us general feedback, simply send an e-mail to feedback@packtpub.com, and mention the book title via the subject of your message.

If there is a book that you need and would like to see us publish, please send us a note in the **SUGGEST A TITLE** form on www.packtpub.com or e-mail suggest@packtpub.com.

If there is a topic that you have expertise in and you are interested in either writing or contributing to a book, see our author guide on www.packtpub.com/authors.

Customer support

Now that you are the proud owner of a Packt book, we have a number of things to help you to get the most from your purchase.

Downloading the example code

You can download the example code files for all Packt books you have purchased from your account at <http://www.PacktPub.com>. If you purchased this book elsewhere, you can visit <http://www.PacktPub.com/support> and register to have the files e-mailed directly to you.

Errata

Although we have taken every care to ensure the accuracy of our content, mistakes do happen. If you find a mistake in one of our books—maybe a mistake in the text or the code—we would be grateful if you would report this to us. By doing so, you can save other readers from frustration and help us improve subsequent versions of this book. If you find any errata, please report them by visiting <http://www.packtpub.com/support>, selecting your book, clicking on the **errata submission form** link, and entering the details of your errata. Once your errata are verified, your submission will be accepted and the errata will be uploaded on our website, or added to any list of existing errata, under the Errata section of that title. Any existing errata can be viewed by selecting your title from <http://www.packtpub.com/support>.

Piracy

Piracy of copyright material on the Internet is an ongoing problem across all media. At Packt, we take the protection of our copyright and licenses very seriously. If you come across any illegal copies of our works, in any form, on the Internet, please provide us with the location address or website name immediately so that we can pursue a remedy.

Please contact us at copyright@packtpub.com with a link to the suspected pirated material.

We appreciate your help in protecting our authors, and our ability to bring you valuable content.

Questions

You can contact us at questions@packtpub.com if you are having a problem with any aspect of the book, and we will do our best to address it.

1

Getting Started with MongoDB

We are about to begin our journey in PHP and MongoDB web development. Since you picked up this book, I assume you have some background building web apps using PHP, and you are interested in learning to develop PHP applications with MongoDB as data backend. In case you have never heard of MongoDB before, it is an open source, document-oriented database that supports the concept of flexible schema. In this chapter, we will learn what MongoDB is, and what do we gain from using MongoDB instead of trusted old SQL databases. We will start by learning briefly about the NoSQL databases (a set of database technologies that are considered alternative to RDBM systems), the basics of MongoDB, and what distinguishes it from relational databases. Then we will move on to installing and running MongoDB and hooking it up with PHP.

To sum it up, in this chapter we will:

- Learn about the NoSQL movement
- Learn the basic concepts behind MongoDB
- Learn how to download, install, and run MongoDB on a computer
- Learn to use the `mongo` Interactive Shell
- Learn how to make PHP and MongoDB talk to each other

So let's get on with it...

- [click Dead Weight \(Lizzy Gardner, Book 2\) for free](#)
- [The Texas Food Bible: From Legendary Dishes to New Classics pdf, azw \(kindle\), epub](#)
- [download online The Man Who Walked Through Time: The Story of the First Trip Afoot Through the Grand Canyon pdf, azw \(kindle\)](#)
- [Rage Across Egypt \(Werewolf: The Apocalypse\) pdf, azw \(kindle\), epub, doc, mobi](#)
- [read Around the World in Seventy-Two Days and Other Writings \(Penguin Classics\)](#)
- [read Peach Cobbler Murder \(Hannah Swensen Mystery, Book 7\) for free](#)

- <http://unpluggedtv.com/lib/Dead-Weight--Lizzy-Gardner--Book-2-.pdf>
- <http://unpluggedtv.com/lib/The-Texas-Food-Bible--From-Legendary-Dishes-to-New-Classics.pdf>
- <http://metromekanik.com/ebooks/The-Man-Who-Walked-Through-Time--The-Story-of-the-First-Trip-Afoot-Through-the-Grand-Canyon.pdf>
- <http://honareavalmusic.com/?books/Der-Celista--Perry-Rhodan-Neo--Band-38--Das-Gro--e-Imperium--Band-2-.pdf>
- <http://www.mmastyles.com/books/l-m-a-Vegetarian.pdf>
- <http://twilightblogs.com/library/Peach-Cobbler-Murder--Hannah-Swensen-Mystery--Book-7-.pdf>