

# 1

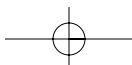
## You: Programmer and Search Engine Marketer

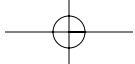
*Googling* for information on the World Wide Web is such a common activity these days that it is hard to imagine that just a few years ago this verb did not even exist. Search engines are now an integral part of our lifestyle, but this was not always the case. Historically, systems for finding information were driven by data organization and classification performed by humans. Such systems are not entirely obsolete — libraries still keep their books ordered by categories, author names, and so forth. Yahoo! itself started as a manually maintained directory of web sites, organized into categories. Those were the good old days.

Today, the data of the World Wide Web is enormous and rapidly changing; it cannot be confined in the rigid structure of the library. The format of the information is extremely varied, and the individual bits of data — coming from blogs, articles, web services of all kinds, picture galleries, and so on — form an almost infinitely complex virtual organism. In this environment, making information *findable* necessitates something more than the traditional structures of data organization or classification.

Introducing the ad-hoc query and the modern search engine. This functionality reduces the aforementioned need for organization and classification; and since its inception, it has become quite pervasive. Google's popular email service, Gmail, features its searching capability that permits a user to find emails that contain a particular set of keywords. Microsoft Windows Vista now integrates an instant search feature as part of the operating system, helping you quickly find information within any email, Word document, or database on your hard drive from the Start menu regardless of the underlying file format. But, by far, the most popular use of this functionality is in the World Wide Web search engine.

These search engines are the exponents of the explosive growth of the Internet, and an entire industry has grown around their huge popularity. Each visit to a search engine potentially generates business for a particular vendor. Looking at Figure 1-1 it is easy to figure out where people in Manhattan are likely to order pizza online. Furthermore, the traffic resulting from non-sponsored, or organic, search results costs nothing to the vendor. These are highlighted in Figure 1-1.





## Chapter 1: You: Programmer and Search Engine Marketer

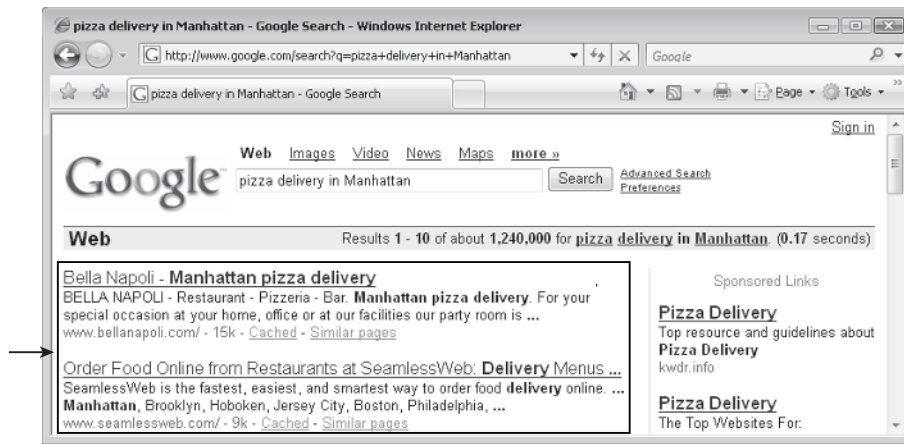


Figure 1-1

The less obvious effect of the search engine explosion phenomenon is that web developers are now directly involved in the search engine marketing process. To rank well in these organic results, it may not be enough to “write relevant content,” as your typical search engine marketing tutorial drones. Rather, the web application developer must work together with the marketing team, and he or she must build a web site fully aware that certain features or technologies may interfere with a search engine marketing campaign. An improperly designed web site can interfere with a search engine’s need to periodically navigate and index the information contained therein. In the worst case, the search engine may not be able to index the content at all.

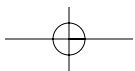
So, ironically, while users are becoming less interested in understanding the structure of data on the Internet, the structure of a web site is becoming an increasingly important facet in search engine marketing! This structure — the architecture of a web site — is the primary focus of this book.

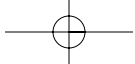
We hope that this brief introduction whets your appetite! The remainder of this chapter tells you what to expect from this book. You will also configure your development machine to ensure you won’t have any problems following the technical exercises in the later chapters.

## Who Are You?

Maybe you’re a great programmer or IT professional, but marketing isn’t your thing. Or perhaps you’re a tech-savvy search engine marketer who wants a peek under the hood of a search engine optimized web site. Search engine marketing is a field where technology and marketing are both critical and interdependent, because small changes in the implementation of a web site can make you or break you in search engine rankings. Furthermore, the fusion of technology and marketing know-how can create web site features that attract more visitors.

The *raison d’être* of this book is to help web developers create web sites that rank well with the major search engines, and to teach search engine marketers how to use technology to their advantage. We assert that neither marketing nor IT can exist in a vacuum, and it is essential that they not see themselves as opposing forces in an organization. They *must* work together. This book aims to educate both sides in that regard.





## Chapter 1: You: Programmer and Search Engine Marketer

### The Story

So how do a search engine marketer from the USA (Jaimie) and a programmer from Romania (Cristian) meet? To answer, we need to tell you a funny little story. A while ago, Jaimie happened to purchase a book (that shall remain nameless) written by Cristian, and was not pleased with one particular aspect of its contents. Jaimie proceeded to grill him with some critical comments on a public web site. Ouch!

Cristian contacted Jaimie courteously, and explained most of it away. No, we're not going to tell you the name of the book, what the contents were, or whether it is still in print. But things did eventually get more amicable, and we started to correspond about what we do for a living. Jaimie is a web site developer and search engine marketer, and Cristian is a software engineer who has published quite a few books in the technology sector. As a result of those discussions, the idea of a technology-focused search engine optimization book came about. The rest is more or less history.

## What Do You Need to Learn?

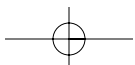
As with anything in the technology-related industry, one must constantly learn and research to keep apprised of the latest news and trends. *How exhausting!* Fortunately, there are fundamental truths with regard to search engine optimization that are both easy to understand and probably won't change in time significantly — so a solid foundation that you build now will likely stand the test of time.

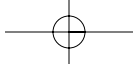
We remember the days when search engine optimization was a black art of analyzing and improving on-page factors. Search engine marketers were obsessed over keyword density and which HTML tags to use. Many went so far as to recommend optimizing content for different search engines individually, thusly creating different pages with similar content optimized with different densities and tags. Today, that would create a problem called *duplicate content*.

The current struggle is creating a site with interactive content and navigation with a minimal amount of duplicate content, with URLs that do not confuse web spiders, and a tidy internal linking structure. There is a thread on SearchEngineWatch (<http://www.searchenginewatch.com>) where someone asked which skill everyone reading would like to hone. Almost all of them enumerated programming as one of the skills (<http://forums.searchenginewatch.com/showthread.php?t=11945>). This does not surprise us. Having an understanding of both programming and search engine marketing will serve one well in the pursuit of success on the Internet.

When people ask us where we'd suggest spending money in an SEO plan, we always recommend making sure that one is starting with a sound basis. If your web site has architectural problems, it's tantamount to trumpeting your marketing message atop a house of cards. *Professional Search Engine Optimization with PHP: A Developer's Guide to SEO* aims to illustrate how to build a solid foundation.

To get the most out of this journey, you should be familiar with a bit of programming (PHP, preferably). You can also get quite a bit out this book by only reading the explanations. And another strategy to reading this book is to do just that — then hand this book to the web developer with a list of concerns and directives in order to ensure the resulting product is search engine optimized. In that case, don't get bogged down in the exercises — just skim them.





## Chapter 1: You: Programmer and Search Engine Marketer

---

We cover a quick introduction to SEO in Chapter 2, which should nail down the foundations of that subject. However, PHP and MySQL are vast subjects; and this book cannot afford to also be a PHP and MySQL tutorial. The code samples are explained step by step, but if you have never written a line of PHP or SQL before, and want to follow the examples in depth, you should also consider reading a PHP and MySQL tutorial book, such as the following:

- ❑ *PHP and MySQL for Dynamic Web Sites: Visual QuickPro Guide, 2nd edition* (Larry Ulman, Peachpit Press, 2005)
- ❑ *Build Your Own Database Driven Website Using PHP & MySQL, 3<sup>rd</sup> Edition* (Kevin Yank, Sitepoint, 2005)
- ❑ *Teach Yourself PHP in 10 Minutes* (Chris Newman, Sams, 2005)

### SEO and the Site Architecture

A web site's architecture is what grounds all future search engine marketing efforts. The content rests on top of it, as shown in Figure 1-2. An optimal web site architecture facilitates a search engine in traversing and understanding the site. Therefore, creating a web site with a search engine optimized architecture is a major contributing factor in achieving and maintaining high search engine rankings.

Architecture should also be considered throughout a web site's lifetime by the web site developer, alongside other factors such as aesthetics and usability. If a new feature does not permit a search engine to access the content, hinders it, or confuses it, the effects of good content may be reduced substantially. For example, a web site that uses Flash or AJAX technologies inappropriately may obscure the majority of its content from a search engine.

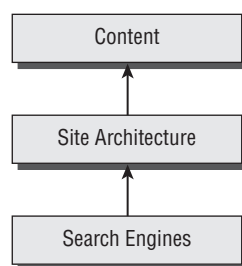
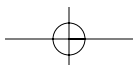
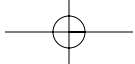


Figure 1-2

We do not cover copywriting concepts in detail, or provide much coaching as to how to create persuasive page titles. These are also very important topics, which are masterfully covered by Bryan and Jeffrey Eisenberg in *Persuasive Online Copywriting: How to Take Your Words to the Bank* (Wizard Academy Press, 2002), and by John Caples and Fred E. Hahn in *Tested Advertising Methods, 5<sup>th</sup> edition* (Prentice Hall, 1998). Shari Thurow also has an excellent section on creating effective titles in her book, *Search Engine Visibility* (New Riders Press, 2002). Writing copy and titles that rank well are obviously not successful if they do not convert or result in click-throughs, respectively. We do give some pointers, though, to get you started.

We also do not discuss concepts related to search engine optimization such as usability and user psychology in depth, though they are strong themes throughout the book.





## Chapter 1: You: Programmer and Search Engine Marketer

---

Optimizing a site's architecture frequently involves tinkering with variables that also affect usability and the overall user perception of your site. When we encounter such situations, we alert you to why these certain choices were made. Chapter 5, "Duplicate Content," highlights a typical problem with breadcrumbs and presents some potential solutions. Sometimes we find that SEO enhancements run counter to usability. Likewise, not all designs that are user friendly are search engine friendly. Either way, a compromise must be struck to satisfy both kinds of visitors — users and search engines.

### ***SEO Cannot Be an Afterthought***

One common misconception is that search engine optimization efforts can be made after a web site is launched. This is frequently incorrect. Whenever possible, a web site can and should be designed to be search engine friendly as a fundamental concern.

Unfortunately, when a preexisting web site is designed in a way that poses problems for search engines, search engine optimization can become a much larger task. If a web site has to be redesigned, or partially redesigned, the migration process frequently necessitates special technical considerations. For example, old URLs must be *properly* redirected to new ones with similar relevant content.

The majority of this book documents best practices for design from scratch as well as how to mitigate redesign problems and concerns. The rest is dedicated to discretionary enhancements.

### ***Communicating Architectural Decisions***

The aforementioned scenario regarding URL migration is a perfect example of how the technical team and marketing team must communicate. The programmer must be instructed to add the proper redirects to the web application. Otherwise existing search rankings may be hopelessly lost forever. Marketers must know that such measures must be taken in the first place.

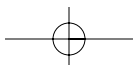
In a world where organic rankings contribute to the bottom line, a one-line redirect command in a web server configuration file may be much more important than one may think. This particular topic, URL migration, is discussed in Chapter 4.

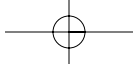
### ***Architectural Minutiae Can Make or Break You***

So you now understand that small mistakes in implementation can be quite insidious. Another common example would be the use of JavaScript-based navigation, and failing to provide an HTML-based alternative. Spiders would be lost, because they, for the most part, do not interpret JavaScript.

The search engine spider is "the third browser." Many organizations will painstakingly test the efficacy and usability of a design in Internet Explorer and Firefox with dedicated QA teams. *Unfortunately, many fall short by neglecting to design and test for the spider.* Perhaps this is because you have to design in the abstract for the spider; we don't have a Google spider at our disposal after all; and we can't interview it afterward with regard to what it thought of our "usability." However, that does not make its assessment any less important.

The Spider Simulator tool located at <http://www.seochat.com/seo-tools/spider-simulator/> shows you the contents of a web page from the perspective of a hypothetical search engine. The tool is very simplistic, but if you're new to SEO, using it can be an enlightening experience.





## Chapter 1: You: Programmer and Search Engine Marketer

---

### Preparing Your Playground

This book contains many exercises, and all of them assume that you've prepared your environment as explained in the next few pages. If you're a PHP and MySQL veteran, here's the quick list of software requirements. If you have these, you can skip to the end of the chapter, where you're instructed to create a MySQL database for the few exercises in this book that use it.

- Apache 2 or newer, with the `mod_rewrite` module
- PHP 4.1 or newer
- MySQL

Your PHP installation should have these modules:

- `php_mysql` (necessary for the chapters that work with MySQL)
- `php_gd2` (necessary for exercises in Chapter 5 and Chapter 10)
- `php_curl` (necessary for exercises in Chapter 11)

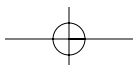
**The programming exercises in this book assume prior experience with PHP and MySQL. However, if you follow the exercises with discipline, exactly as described, everything should work as planned.**

If you already have PHP but you aren't sure which modules you have installed, view your `php.ini` configuration file. On a default Windows installation, this file is located in the `Windows` folder; if you install PHP through XAMPP as shown in the exercise that follows, the path is `\Program Files\xampp\apache\bin`. To enable a module, remove the leading `;` from the `extension=module_name.dll` line, and restart Apache.

After installing the necessary software, you'll create a virtual host named `sephp.example.com`, which will point to a folder on your machine, which will be your working folder for this book. All exercises you build in this book will be accessible on your machine through `http://sephp.example.com`.

Lastly, you'll prepare a MySQL database named `sephp`, which will be required for a few of the exercises in this book. Creating the database isn't a priority for now, so you can leave this task for when you'll actually need it for an exercise.

**The next few pages cover the exact installation procedure assuming that you're running Microsoft Windows. If you're running Linux or using a web hosting account, we assume you already have Apache, PHP, and MySQL installed with necessary modules.**



## Chapter 1: You: Programmer and Search Engine Marketer

### Installing XAMPP

XAMPP is a package created by Apache Friends (<http://www.apachefriends.org>), which includes Apache, PHP, MySQL, and many other goodies. If you don't have these already installed on your machine, the easiest way to have them running is to install XAMPP.

Here are the steps you should follow:

1. Visit <http://www.apachefriends.org/en/xampp.html>, and go to the XAMPP page specific for your operating system.
2. Download the XAMPP installer package, which should be an executable file named like `xampp-win32-version-installer.exe`.
3. Execute the installer executable. When asked, choose to install Apache and MySQL as services, as shown in Figure 1-3. Then click Install.
4. You'll be asked to confirm the installation of each of these as services. Don't install the FileZilla FTP Server service unless you need it for particular purposes (you don't need it for this book), but do install Apache and MySQL as services.

**Note that you can't have more web servers working on port 80 (the default port used for HTTP communication). If you already have a web server on your machine, such as IIS, you should either make it use another port, uninstall it, or deactivate it. Otherwise, Apache won't work. The exercises in this book assume that your Apache server works on port 80; they may not work otherwise.**

5. In the end, confirm the execution of the XAMPP Control Panel, which can be used for administering the installed services. Figure 1-4 shows the XAMPP Control Panel.

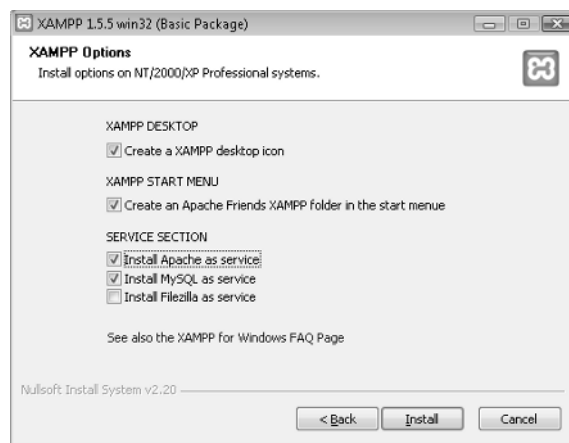
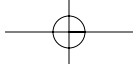


Figure 1-3



## Chapter 1: You: Programmer and Search Engine Marketer

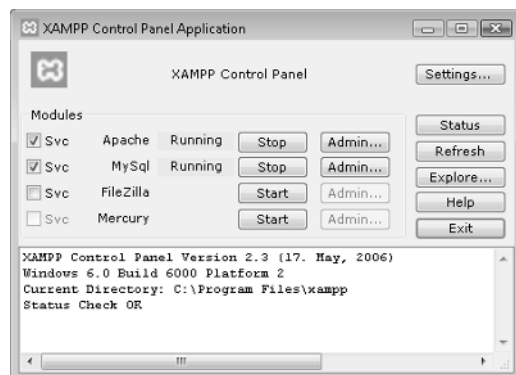


Figure 1-4

The XAMPP Control Panel is particularly useful when you need to stop or start the Apache server. Every time you make a change to the Apache configuration files, you'll need to restart Apache.

6. To test that Apache installed correctly, load `http://localhost/` using your web browser. An XAMPP welcome screen like the one in Figure 1-5 should load.
7. Finally, after you've tested that both Apache and PHP work, it's recommended to turn on PHP error reporting. Apache logs errors in a file named `error.log`, located in the `xampp\apache\logs` folder; looking at the latest entries of this file when something goes wrong with your application can be very helpful at times. To enable PHP error reporting, open for editing the `php.ini` configuration file, located by default in the `xampp\apache\bin\` folder. There, locate this entry:

```
display_errors = Off
```

and change it to:

```
display_errors = On
```

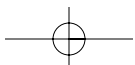
8. To configure what kind of errors you want reported, you can alter the value of the PHP `error_reporting` value. We recommend the following setting to report all errors, except for PHP notices:

```
error_reporting = E_ALL & ~E_NOTICE
```

### Preparing the Working Folder

Now you'll create a virtual host named `seophp.example.com` on your local machine, which will point to a local folder named `seophp`. The `seophp` folder will be your working folder for all the exercises in this book, and you'll load the sample pages through `http://seophp.example.com`.

*The `seophp.example.com` as virtual host won't interfere with any existing online applications, because `example.com` is a special domain name reserved by IANA to be used for documentation and demonstration purposes. See `http://example.com` for the official information.*



## Chapter 1: You: Programmer and Search Engine Marketer

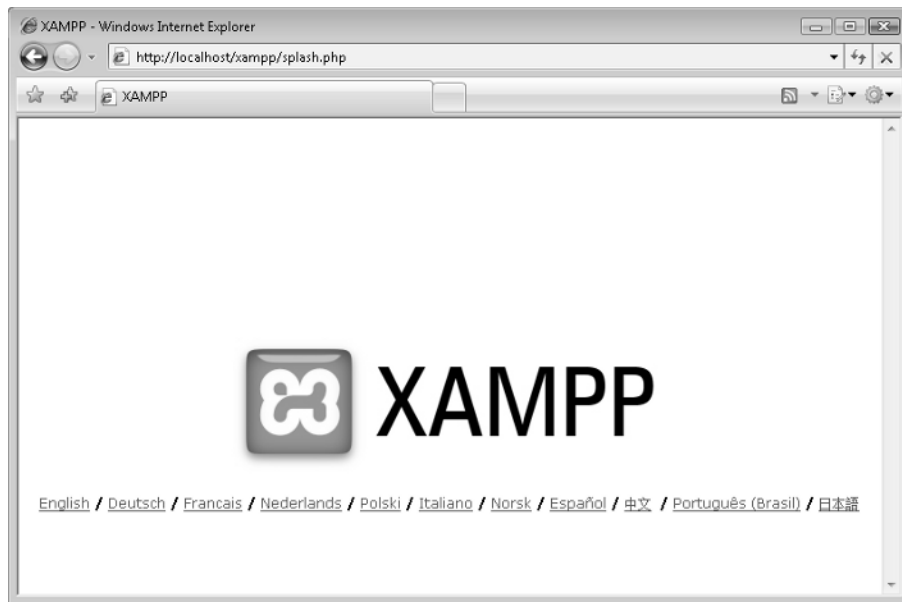


Figure 1-5

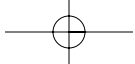
Follow these steps to create and test the virtual host on your machine:

1. First, you need to add `seophp.example.com` to the Windows `hosts` file. The following line will tell Windows that all domain name resolution requests for `seophp.example.com` should be handled by the local machine instead of your configured DNS. Open the `hosts` file, which is located by default in `C:\Windows\System32\drivers\etc\hosts`, and add this line to it:

```
127.0.0.1    localhost
127.0.0.1    seophp.example.com
```

2. Now create a new folder named `seophp`, which will be used for all the work you do in this book. You might find it easiest to create it in the root folder (`C:\`), but you can create it anywhere else if you like.
3. Finally, you need to configure a virtual host for `seophp.example.com` in Apache. Right now, all requests to `http://localhost/` and `http://seophp.example.com/` are handled by Apache, and both yield the same result. You want requests to `http://seophp.example.com/` to be served from your newly created folder, `seophp`. This way, you can work with this book without interfering with the existing applications on your web server.

To create the virtual host, you need to edit the Apache configuration file. In typical Apache installations there is a single configuration file named `httpd.conf`. XAMPP ships with more configuration files, which handle different configuration areas. To add a virtual host, add the following lines to `xampp\apache\conf\extra\httpd-vhosts.conf`. (If you installed XAMPP with the default options, the `xampp` folder should be under `\Program Files`.)



## Chapter 1: You: Programmer and Search Engine Marketer

---

```
NameVirtualHost 127.0.0.1:80

<VirtualHost 127.0.0.1:80>
  DocumentRoot "C:/Program Files/xampp/htdocs"
  ServerName localhost
</VirtualHost>

<VirtualHost 127.0.0.1:80>
  DocumentRoot C:/seophp/
  ServerName seophp.example.com
  <Directory C:/seophp/>
    Options Indexes FollowSymLinks
    AllowOverride All
    Order allow,deny
    Allow from all
  </Directory>
</VirtualHost>
```

**In order for `http://localhost/` to continue working after you create a virtual host, you need to define and configure it as a virtual host as well — this explains why we've included it in the `vhosts` file. If you have any important applications working under `http://localhost/`, make sure they continue to work after you restart Apache at the end of this exercise.**

4. To make sure `httpd-vhosts.conf` gets processed when Apache starts, open `xampp\apache\conf\httpd.conf` and make sure this line, located somewhere near the end of the file, isn't commented:

```
# Virtual hosts
include conf/extra/httpd-vhosts.conf
```

5. Restart Apache for the new configuration to take effect. The easiest way to restart Apache is to open the XAMPP Control Panel, and use it to stop and then start the Apache service.

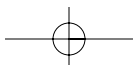
*In case you run into trouble, the first place to check is the Apache error log file. In the default XAMPP installation, this is `xampp\apache\logs\error.log`.*

6. To test your new virtual host, create a new file named `test.php` in your `seophp` folder, and type this code in it:

```
<?php
phpinfo();
?>
```

7. Then load `http://seophp.example.com/test.php` and expect to see a page like the one in Figure 1-6.

This way you've also tested that your PHP installation is working correctly.



## Chapter 1: You: Programmer and Search Engine Marketer

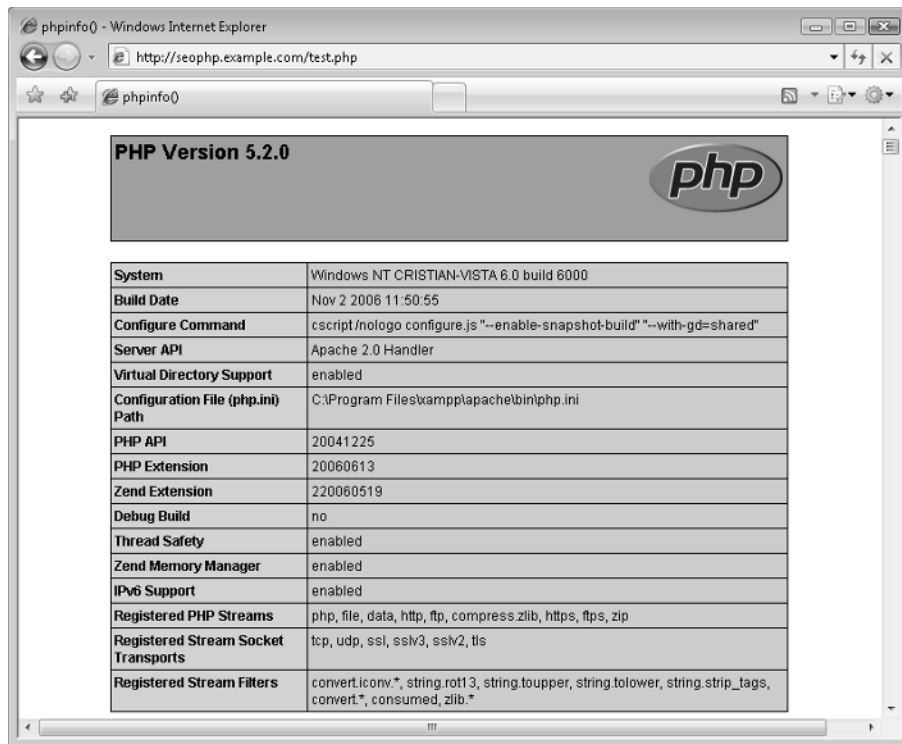


Figure 1-6

### Preparing the Database

The final step is to create a new MySQL database. You're creating a database named `seophp` that you will use for the exercises contained in this book. You'll also create a user named `seouser`, with the password `seomaster`, which will have full privileges to the `seophp` database.

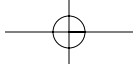
*You will be using this database only for the exercises in Chapter 11 and Chapter 14, so you can skip this database installation for now if desired.*

To prepare your database environment, follow these steps. Note that this exercise uses the MySQL console application to send commands to the database server.

Follow these steps:

1. Load a Windows Command Prompt window by going to Start ⇨ Run and executing `cmd.exe`. In Windows Vista, you can type `cmd` or `Command Prompt` in the search box of the Start menu.
2. Change your current directory to the `bin` folder of your MySQL installation. With the default XAMPP installation, that folder is `\Program Files\xampp\mysql\bin`. Change the directory using the following command:

```
cd \Program Files\xampp\mysql\bin
```



## Chapter 1: You: Programmer and Search Engine Marketer

---

3. Start the MySQL console application using the following command (this loads an executable file named `mysql.exe` located in the directory you have just browsed to):

```
mysql -u root
```

*If you have a password set for the root account, you should also add the `-p` option, which will have the tool ask you for the password. By default, after installing XAMPP, the `root` user doesn't have a password. Needless to say, you may want to change this for security reasons.*

4. Create the `seophp` database by typing this at the MySQL console:

```
CREATE DATABASE seophp;
```

*MySQL commands, such as `CREATE DATABASE`, are not case sensitive. If you like, you can type `create database` instead of `CREATE DATABASE`. However, database objects, such as the `seophp` database, may or may not be case sensitive, depending on the server settings and operating system. For this reason, it's important to always use consistent casing. (This book uses uppercase for MySQL commands, and lowercase for object names.)*

5. Switch context to the `seophp` database.

```
USE seophp;
```

6. Create a database user with full access to the new `seophp` database:

```
GRANT ALL PRIVILEGES ON seophp.*
TO seouser@localhost IDENTIFIED BY "seomaster";
```

7. Make sure all commands executed successfully, as shown in Figure 1-7.

8. Exit the console by typing:

```
exit;
```

```

C:\Program Files\xampp\mysql\bin\mysql.exe
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 2 to server version: 5.0.24a-community-nt

Type 'help;' or '\h' for help. Type '\c' to clear the buffer.

mysql> CREATE DATABASE seophp;
Query OK, 1 row affected (0.00 sec)

mysql> USE seophp;
Database changed
mysql> GRANT ALL PRIVILEGES ON seophp.*
-> TO seouser@localhost IDENTIFIED BY "seomaster";
Query OK, 0 rows affected (0.00 sec)

mysql> _

```

Figure 1-7

## Summary

Congratulations! You will soon be ready to write some code and delve into more advanced SEO concepts! The next chapter takes you through a quick SEO tutorial, and builds the foundation for the chapters to come.

