

Fundamentals of Web Development

Third Edition by Randy Connolly and Ricardo Hoar



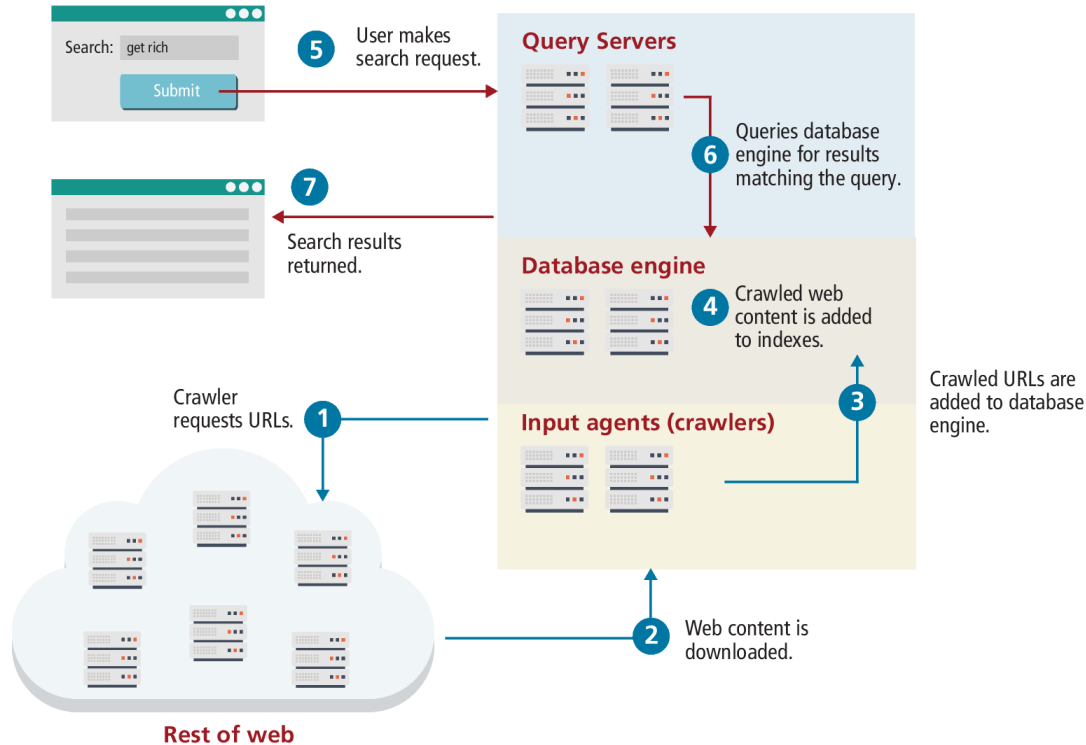
Chapter 18

Tools and Traffic

In this chapter you will learn . . .

- How search engines work
- Common search practices to improve your rank
- About social media integration
- How Content Management Systems make it easier to manage websites
- The basics of web advertising

Search Engine Overview



Web Crawlers

Web crawlers refer to a class of software that downloads pages, identifies the hyperlinks, and adds them to a database for future crawling. Crawlers are sometimes called web spiders, robots, worms, or wanderers and can be thought of as an automated text browser.

```
class Crawler {
    private $URLList;
    private $nextIndex;
    function __construct(){
        $this->nextIndex=0;
        $this->URLList =
            array("http://SEEDWEBSITE/");
    }
    private function
        getNextURLToCrawl(){
        return $this->URLList[$this-
            >nextIndex++];
        }
    private function printSummary(){
        echo count($this->URLList)." links.
        Index:".
        $this->nextIndex."<br>";
        foreach($this->URLList as $link){
            echo $link."<br>";
        }
    }
    public function doIteration(){
        $url = $self->getNextURLToCrawl();
        // Do not crawl if not allowed
        if (robotsDisallow($url))
            return;
        echo "Crawling ".$url."<br>";
        //this function finds the <a> links
        scrapeHyperlinks($url);
        $self->printSummary();
    }
}
```

LISTING 18.1 Simple crawler class in PHP

Web Scrapers

Scrapers are programs that identify certain pieces of information from the web to be stored in databases.

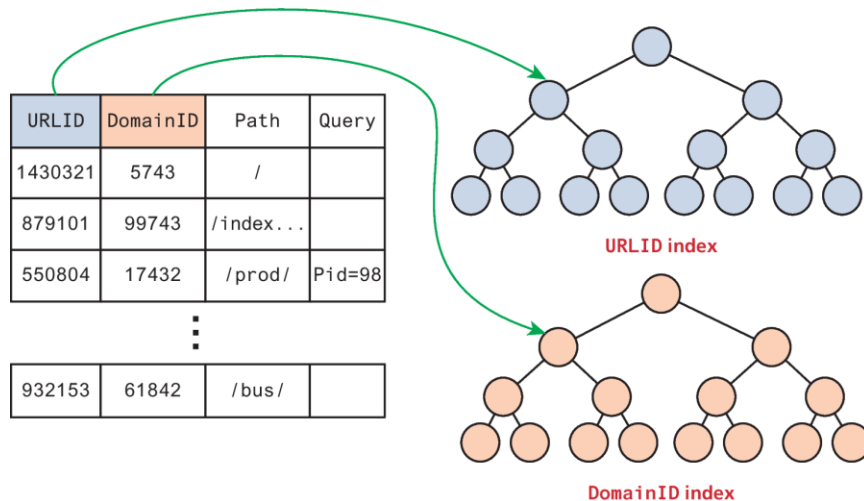
- **URL Scrapers** identify URLs inside of a page by seeking out all the <a> tags and extracting the value of the href attribute.
- **Email scrapers** harvest email accounts by seeking the words mailto: in the href attribute of a link.
- **Vulnerability scrapers** scan a website for information about the underlying software.
- **Word Scrapers** parse out is all of the text within a web page.

Indexing and Reverse Indexing

To understand indexing, consider what a crawler and a scraper might identify from a web page and how they might store it.

The URL is stored, as are rows for each link found to other URLs.

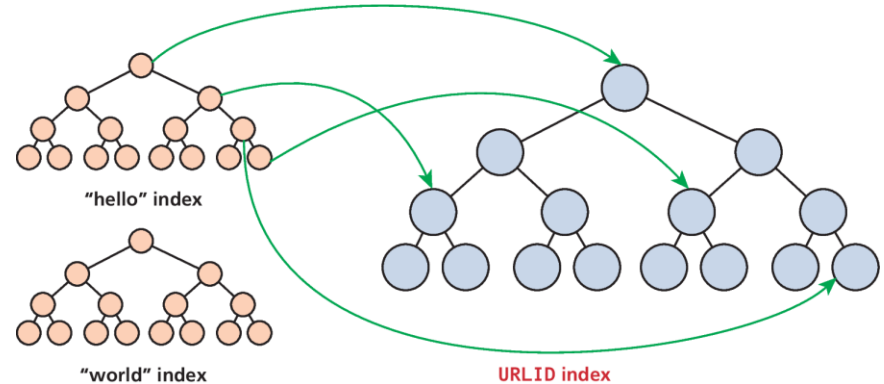
We can quickly get all the words associated with an index, but we want the opposite!



Reverse Index

We need to know which URLs contain a given word. For that, a **reverse index** is built, which indexes the words, rather than the URLs

Since engines are indexing words anyhow, there is an opportunity to improve the efficiency of the index by **stemming** the words first



PageRank and Result Order

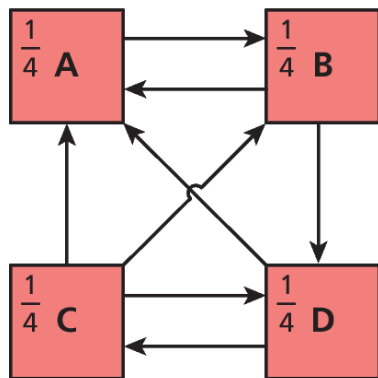
The simplified definition of a site n 's PageRank is:
$$PR(n) = \sum_{v \in B_n} \frac{PR(v)}{N_v}$$

The PageRank of a page, $PR(n)$, is determined by collecting every page v that links to n ($v \in B_n$), and summing their PageRanks $PR(v)$ divided by the number of links out (N_v).

In order to apply this algorithm, we begin by assigning each page the same rank: $1/(\text{number of pages})$.

PageRank Example

$$PR(n) = \sum_{v \in B_n} \frac{PR(v)}{N_v}$$



Iteration 0

Calculate the updated PageRank 1 at-a-time

$$PR(A) = \frac{PR(B)}{N_B} + \frac{PR(C)}{N_C} + \frac{PR(D)}{N_D}$$

We know the page ranks of B, C, D and can count the links out of each N_B , N_C , and N_D .

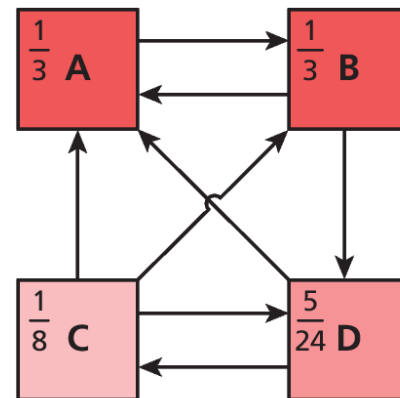
$$PR(A) = \frac{1/4}{2} + \frac{1/4}{3} + \frac{1/4}{2} = \frac{1}{3}$$

Since B has A and C backlinking to it:

$$PR(B) = \frac{PR(A)}{N_A} + \frac{PR(C)}{N_C} \Rightarrow \frac{1}{4} + \frac{1/4}{3} \Rightarrow \frac{1}{3}$$

C has only D backlinking to it and D has B and C backlinks, so

$$PR(C) = \frac{PR(D)}{N_D} \Rightarrow \frac{1/4}{2} \Rightarrow \frac{1}{8} \quad PR(D) = \frac{PR(B)}{N_B} + \frac{PR(C)}{N_C} \Rightarrow \frac{1/4}{2} + \frac{1/4}{3} \Rightarrow \frac{5}{24}$$



Iteration 1

Search Engine Optimization

Search engine optimization (SEO) is the process a webmaster undertakes to make a website more appealing to search engines, and by doing so, increases its ranking in search results for terms the webmaster is interested in targeting.

For many businesses, the optimization of their website is more important than the site itself.

SEO techniques can be broken down into two major categories:

- **white-hat SEO** that tries to honestly and ethically improve your site for search engines, and
- **black-hat SEO** that tries to game the results in your favor.

White Hat SEO

- **Title**
- **Meta Tags**
- **URLs**
- **Site Design**
- **Sitemaps**
- **Anchor Text**
- **Images**
- **Content**

Title Tag

The content of the <title> tag is how your site is identified in search engine results

- make it unique on each page of your site
- Include enough keywords to make it relevant in search engine results.
- Consider delimiting characters such as | or – to separate components of a title, allowing uniqueness and keywords.
- Although one should not overemphasize keywords, one should definitely include them when reasonable.

Meta Tags

Meta tags can set a page's charset, meta information, robots directives, HTTP redirects, and more

- **http-equiv** can perform HTTP-like operations like redirects and set headers.
- Meta *tags* where the **name** attribute is **description** have a corresponding **content** attribute, which contains a human-readable summary of your site

```
<meta name="description" content="The companion site for the textbook " />
```

- We can control some behavior of search engines through meta tags with the name attribute set to robots.

```
<meta name="robots" content="INDEX,NOFOLLOW" />
```

URLs

Some URLs work just fine but cannot be read by humans easily. A URL that identifies a product in a car parts website, for example, might look like the following :

`/products/index.php?productID=71829`

Descriptive Path Components provide search engines additional relevant material to index your site with.

`/products/AirFilters/index.php?productID=71829`

Descriptive File Names or Folders provide search engines additional information

`/products/AirFilters/BudgetBrandX100/`

Site Design

The design and layout of your site has a huge impact on your visibility to search engines. Sites that rely heavily on JavaScript or Flash for their content and navigation will suffer from poor indexing

Navigation links, are the primary means of navigating a site. If your site includes a menu, you should nest it inside of `<nav>` tags to demonstrate semantically that these links exist to navigate your site.

Recurring links are those that appear in a number of places but are not primary navigation.

Ad hoc links are links found in articles and content in general.

Sitemaps

A formal framework that captures website structure is known as a **sitemap**.

Thankfully there are tools to generate sitemaps based on the structure of your site. Google's sitemap generator bases your initial map on your server logs, while other commercial tools parse your entire site.

```
<?xml version="1.0" encoding="utf-8"?>
<urlset
xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>http://funwebdev.com/</loc>
    <lastmod>2013-09-29</lastmod>
    <changefreq>weekly</changefreq>
    <priority>1.0</priority>
  </url>
</urlset>
```

LISTING 18.6 Single page sitemap

Anchor Text

One of the things that is definitely indexed along with backlinks is the **anchor text** of the link.

Your hyperlinks should contain, as often as possible, anchor text that describes the link.

Links to a page of services and rates shouldn't say "*Click here to read more,*" it should read "*Services and Rates,*"

If a backlink to your site does not use some meaningful keywords, the link will not help your ranking for those keywords.

Images

Many search engines now have a separate site to search for images.

- The filename is the first element we can optimize, since it can be parsed for words. Rather than name an image of a rose **1.png**, we should call it **rose.png**.
- The judicious use of the alt attribute in the tag is another place where some textual description of the image can help your ranking.
- If you have a link to the image somewhere on our site, you should use descriptive anchor text such as “full size image of a red rose,” rather than generic text “full size.”

Content

When we refer to content in the SEO context, we are talking about the freshness of content on the whole. To increase the visibility of your pages in search results, you should definitely refresh your content as often as possible.

If your website can offer tools that allow users to comment or otherwise write content on your site, search engines will see the new content on subsequent passes, making the content as a whole look “fresh.”

Facebook, Twitter, MySpace, Slashdot, Reddit, Pinterest, and others all build on the user-submitted content model that ensures their sites are always *fresh*.

Black Hat SEO

We advise you not to use black-hat optimization techniques as you can get banned.

Keyword stuffing is a technique whereby you purposely add keywords into the site in a most unnatural way with the intention of increasing the affiliation between certain key terms and your URL.

Hidden Content builds on keyword stuffing, rather than remove the unwieldy content, tries to hide it using some simple CSS tricks.

Paid links is frowned upon by many search engines, since their intent is to discover good content by relying on referrals

Black Hat SEO (ii)

Doorway pages are pages written to be indexed by search engines and included in search results. Doorway pages are automatically generated pages crammed full of keywords, and effectively useless to real users of your site.

Hidden links are a link spam technique similar to hidden content. With hidden links, websites hide the color of the link to match the background, hoping that real users will not see them.

Comment spam is when automatically generated comments intended to increase links associate your site with spam.

Black Hat SEO (iii)

Google bombing is the technique of using anchor text in links throughout the web to encourage the search engine to associate the anchor text with the destination website.

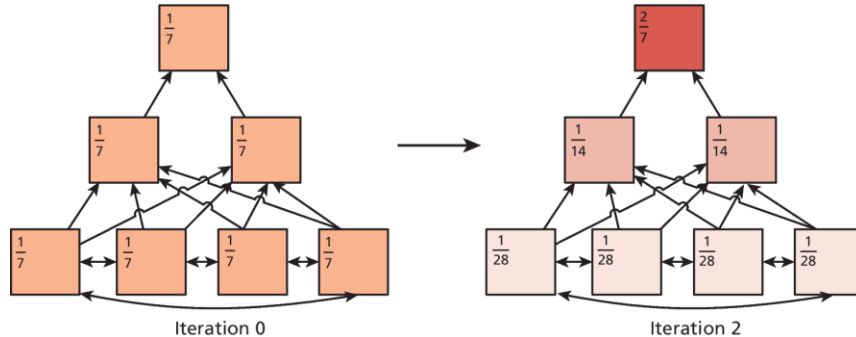
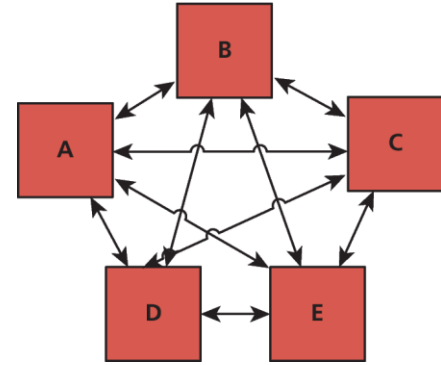
Cloaking refers to the process of identifying crawler requests and serving them content different from regular users. The **user-agent** header is the primary means of identifying crawler agents, which means a simple script can redirect users if **googlebot** is the **user-agent** to a page, normally stuffed with keywords.

Duplicate Content is stealing content to build a fake site

Black Hat: Link Farm/Pyramid

A **link farm** is a set of websites that all interlink each other

Link pyramids are similar to link farms but a pyramid has the intention of promoting one or two sites.



Social Networks

Social networks are web-based systems designed to bring people together by facilitating the exchange of text snippets, photos, links, and other content with other users.

Famous networks include Facebook, Twitter, MySpace, and LinkedIn, among a sea of others.

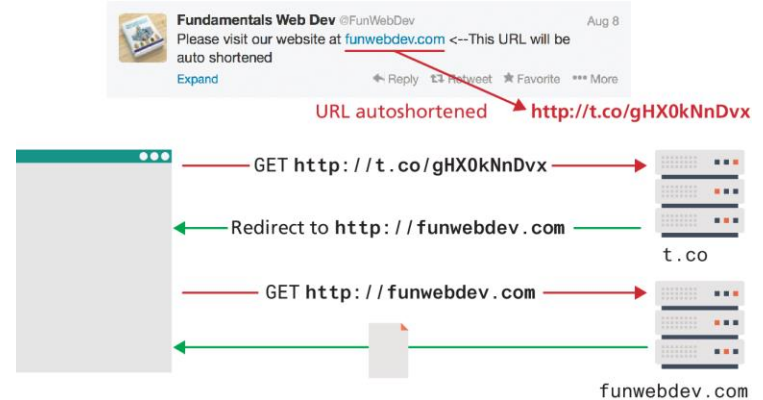
The modern study of social networks draws from psychology, sociology, graph theory, and computer science.

Billions of users are now logged on.

Social Network Integration

Social media lowers the barriers to entry for people who would never want to maintain an HTML page. It provides easy to create basic presence including:

- **Home Pages**
- **Links and Logos**
- **URL Shortening**

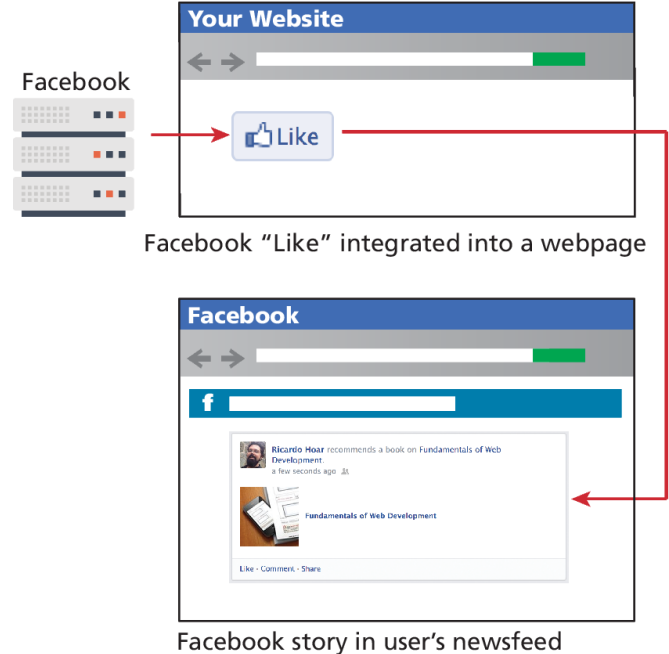


Facebook's Social Plugins

Facebook's social plugins include things you've probably seen before including the **Like** button, an **activity feed**, and **comments**.

For any of the plugins, you will have to choose between HTML5, the Facebook Markup Language (XFBML), or an `<iframe>` implementation.

You will also have to learn a little about the Open Graph API



Register and Plugin

To include the Facebook libraries in your website in the long term, you will have to first register as a developer and get an application ID.

The details of getting an application ID are straightforward. Log in to Facebook and check out <https://developers.facebook.com/>

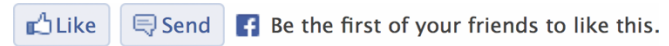
```
<script>
  window.fbAsyncInit = function() {
    FB.init({
      appId :'your-app-id',
      autoLogAppEvents: true,
      xfbml :true,
      version : 'v7.0'
    });
  };
</script>
<script async defer crossorigin="anonymous"
src="https://connect.facebook.net/en_US/sdk.js"></script>
```

LISTING 18.7 Including Facebook JS API

Like Button

With the Facebook classes loaded, you can use it to automatically parse your HTML page for certain tags, and replace them with common plugins.

The **Like button**, being the most widely used, can be included simply by defining a `<div>` element with the class `fb-like`, and some other custom attributes



```
<div class="fb-like"
  data-href="http://funwebdev.com"
  data-width="450"
  data-layout="standard"
  data-action="like"
  data-size="small"
  data-share="true">
</div>
```

LISTING 18.8 HTML5 markup to insert a Like button

More Facebook Plugins

The Facebook Markup Language (XFBML) adds some features.

Follow buttons, and comment sections are very easily added, once the library is loaded.

Note the use of FBML.

```
<fb:follow  
    href=  
    "https://facebook.com/fundamentalsOfWebDevelopment"  
    width="450"  
    show_faces="true">  
</fb:follow>
```

LISTING 18.9 Facebook Follow Me button social plugin

```
<fb:comments  
    href="http://funwebdev.com"  
    width="470">  
</fb:comments>
```

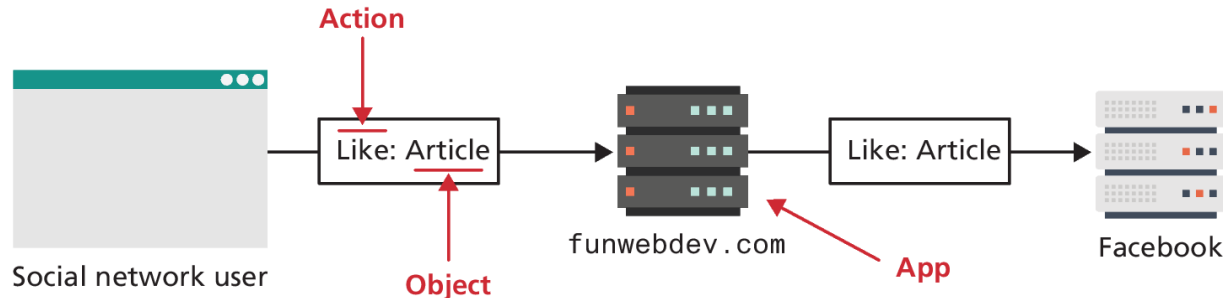
LISTING 18.10 Comment social widget

Open Graph

To control what Facebook uses when displaying items in your newsfeed, you must use Open Graph semantic tags to create **Open Graph Objects**

Open Graph makes use of actors, apps, actions, and objects

You can test your URL by visiting the Facebook Sharing debugger:
<https://developers.facebook.com/tools/debug/>



Open Graph Markup

```
<head prefix="og: http://ogp.me/ns#">
  <meta property="og:locale" content="en_US">
  <meta property="og:url" content="http://funwebdev.com/">
  <meta property="og:title" content="Fundamentals of Web
Development">
  <meta property="og:site_name" content="Fun Web Dev">
  <meta property="og:description" content="Randy Connolly
and Ricardo Hoar are working on a book">
  <meta property="og:image"
content="http://funwebdev.com/wpcontent/
uploads/2013/01/logo.png">
  <meta property="og:image:type" content="image/png">
  <meta property="og:image:width" content="424">
  <meta property="og:image:height" content="130">
  <meta property="og:type" content="book">
</head>
```



LISTING 18.11 Open Graph Markup to add semantic information to your page

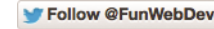
Twitter's Widgets

Like Facebook, Twitter follows the same pattern of including a JavaScript library and then using tags to embed simple social widgets.



```
<a href="https://twitter.com/share"
  class="twitter-share-button"
  data-hashtags="web">
  Tweet</a>
```

LISTING 18.13 Tweet This button markup to create a tweet with hashtag web

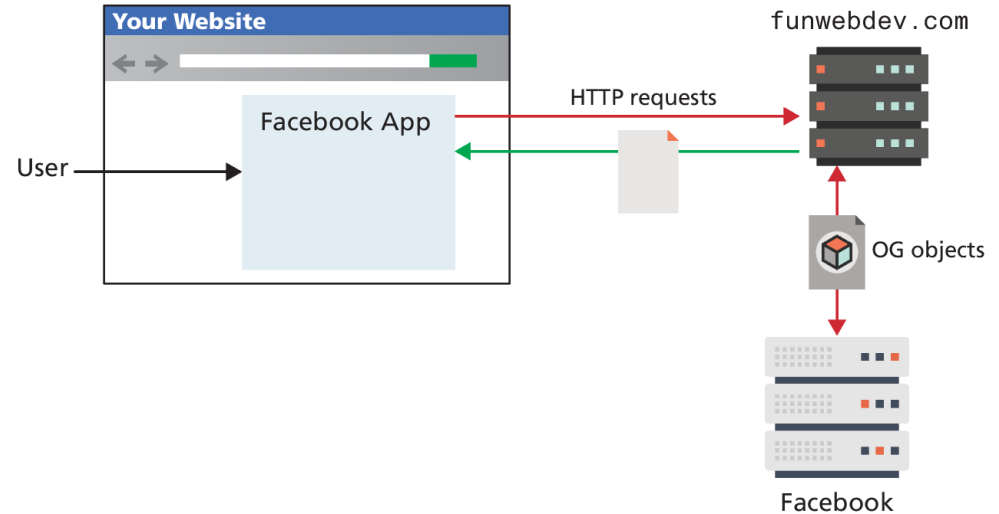


```
<a href="https://twitter.com/FunWebDev"
  class="twitter-follow-button"
  data-show-count="false">
  Follow @FunWebDev</a>
```

LISTING 18.14 Markup to define a Follow button for Twitter

Advanced Social Network Integration

- If your web application offers some sort of service aside from blog posts and static pages, you might want to consider integrating more completely by making use of server-side APIs which allow your server to act as an agent on behalf of users logged in through your site



Content Management Systems

- **Content management system (CMS)** is the name given to the category of software that easily manages websites with support for multiple users. Even for a sophisticated web developer, the challenge of implementing all this functionality would be daunting



Picking a CMS

When selecting a CMS there are several factors to consider including:

- **Technical requirements:** Each CMS has particular requirements in terms of the functionality it offers as well as the server software needed and the database it is compatible with.
- **System support:** Some systems have larger and more supportive communities/ companies than others. Since you are going to rely on the CMS to patch bugs and add new features, it's important that the CMS community be active
- **Ease of use:** Probably the most important consideration is that the system itself must be easy to use by nontechnical staff.

Some Popular CMSs

- **Drupal** Written in PHP, Drupal is a popular CMS with enterprise-level workflow functionality. It is a popular CMS used in many large organizations including **whitehouse.gov** and **data.gov.uk**.
- **Joomla!** Written in PHP, Joomla! Is one of the older free and open-source CMS (started in 2005). With many plugins and extensions available, it continues to be a popular CMS.
- **Contentful** A headless CMS; that is, it provides only the back-end CMS functionality and makes it available via a REST API.
- **SharePoint** SharePoint is an enterprise-focused, proprietary CMS from Microsoft that is especially popular in corporate intranet sites.

WordPress Overview

This chapter uses WordPress as its sample CMS. Originally a blogging engine, WordPress is by far the most popular CMS.

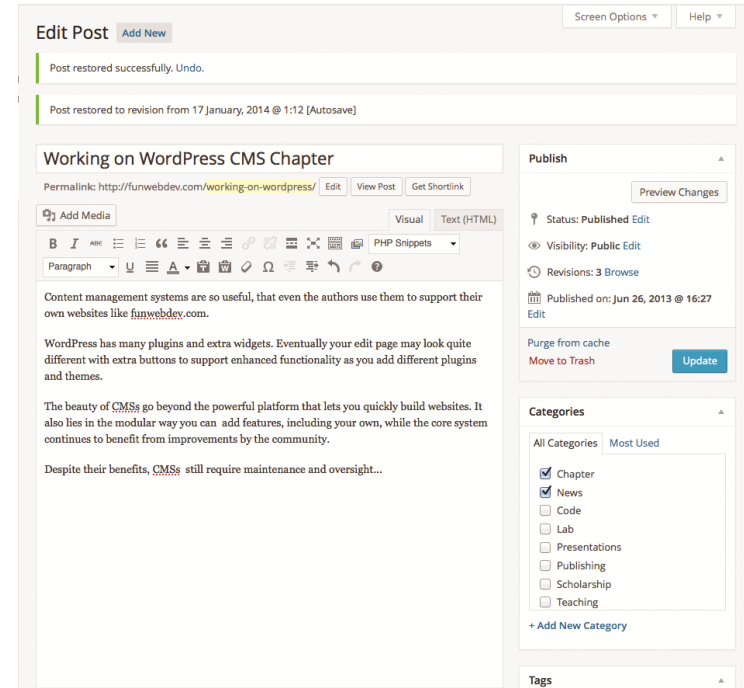
The ability to customize and adapt WordPress has become an important skill for many web developers.

As you will see throughout this chapter, it implements all the key pieces of a complete web management system, and goes beyond that, allowing you to leverage the work of thousands of developers and designers in the form of *plugins* and *themes* (written in PHP).

Post and Page Management

Posts are one important way of adding content to the site. Posts are usually displayed in reverse chronological order and are typically assigned to categories or tagged with keywords as a way of organizing them.

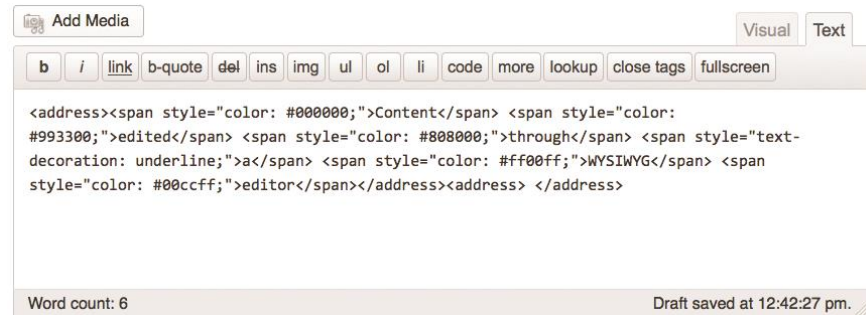
Pages contains content and typically do not display the date, categories, and tags that posts use. The main menu hierarchy of a CMS site will typically be constructed from pages.



WYSIWYG Editors

What You See Is What You Get (WYSIWYG) design is a user interface design pattern where you present the users with an exact (or close to it) view of what the final product will look like.

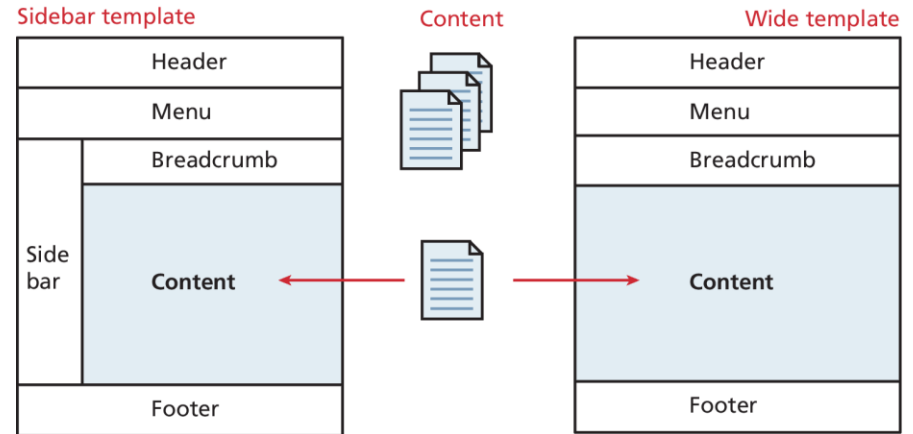
Users are not required to know HTML and CSS, allowing them to edit and create pages with a focus on the content



Template Management

Template management refers to the systems that manage the structure of a website, independently of the content of each particular page.

Several pages can use the same wireframe, but with distinct content as shown in Figure 18.30.



Menu Control

Some key pieces of functionality that should be supported in the **menu control** capability of a CMS include:

- Rearranging menu items and their hierarchy
- Changing the destination page or URL for any menu item
- Adding, editing, or removing menu items
- Changing the style and look/feel of the menu in one place
- Managing short URLs associated with each menu item

User Management and Roles

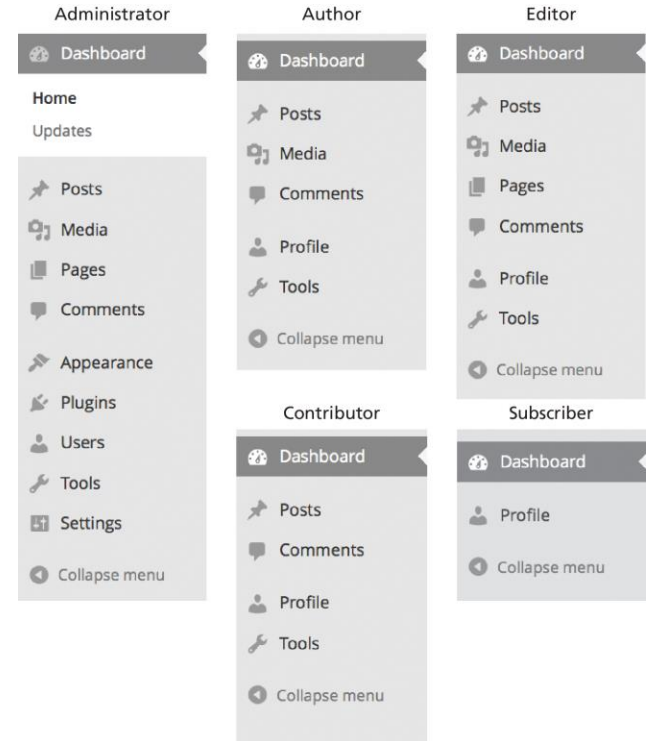
Users in a CMS are given a **user role**, which specifies which rights and privileges that user has. Roles in WordPress are analogous to roles in the publishing industry where the jobs of a journalist, editor, and photographer are distinct.

Content Creator	Content Publisher	Site Manager	Super Administrator
<ul style="list-style-type: none">• Create new web page• Edit existing web page• Save their edits as drafts• Upload media assets	<ul style="list-style-type: none">• Publish content	<ul style="list-style-type: none">• Manage the menu(s)• Manage installed widgets• Manage categories• Manage templates• Manage CMS user accounts• Manage assets	<ul style="list-style-type: none">• Install/Update CMS• Install/Manage plugins• Manage backups• Manage Site Manager• Interface with server

WordPress Roles

In WordPress the default roles are **Administrator**, **Author**, **Editor**, **Contributor**, and **Subscriber**, which are very similar to our generic roles with the Administrator being our super administrator and the Subscriber being a new type of role that is readonly.

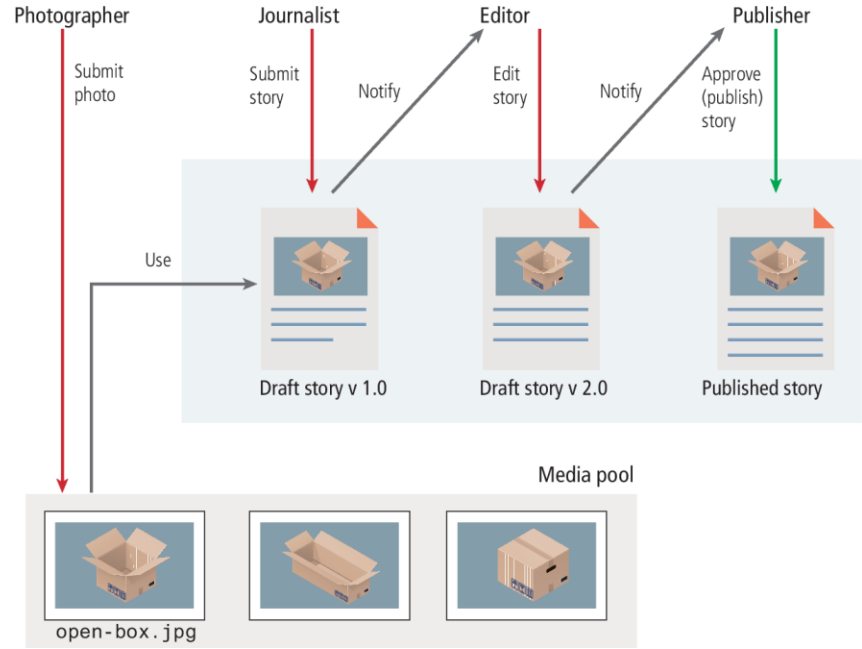
The diagram gives an overall sense of the capabilities.



Workflow and Version Control

Workflow refers to the process of approval for publishing content.

Using roles, you can see that the content created by content creators must eventually be approved or published by a higher-ranking user.



Asset Management

The basic functionality of digital **asset management** software enables the user to:

- Import new assets
- Edit the metadata associated with assets
- Delete assets
- Browse assets for inclusion in content
- Perform searches or apply filters to find assets

Search and Updates

Searching has become a core way that users navigate the web. Unfortunately, creating a fast and correct search of all your content is not straightforward. Thankfully plugins exist, such as WPSearch, which implement search indexes so that you can easily build an index to get faster user searches.

The security of your site is only as good as the weakest link, and an outdated version of WordPress (or any other CMS) may have publicly disclosed vulnerabilities that can be easily exploited.

You can configure automatic updates to improve the security of your system without manual intervention;

WordPress Technical Overview

By now it's obvious that WordPress meets the standards of a decent CMS from an end user's perspective.

This section delves deeper into the installation, configuration, and use of WordPress, including themes and plugins customizations.

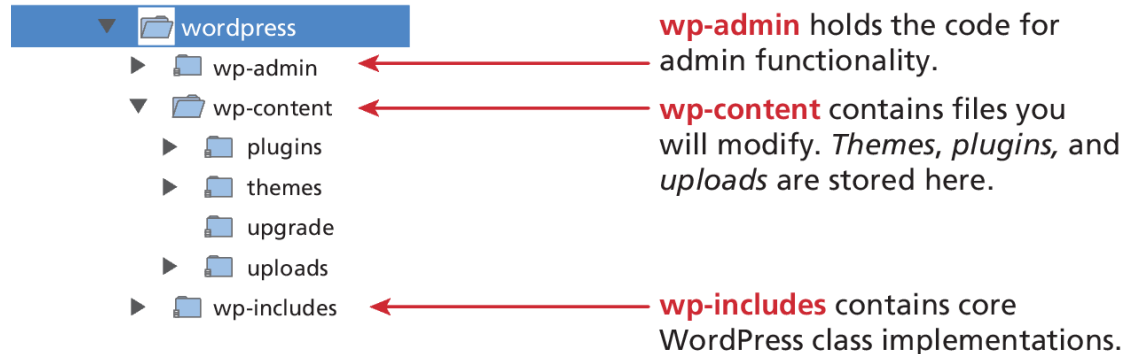
WordPress is written in PHP and relies on a database engine to function. You therefore require a server configured in much the same way as the systems you have used thus far.

The WordPress PHP code is distributed in a zipped folder and WordPress proudly boasts that it can be installed in five minutes

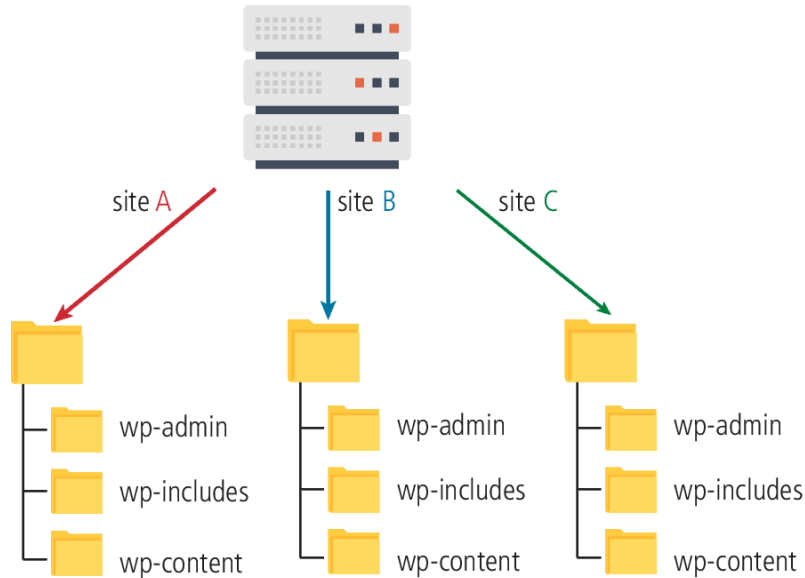
File Structure

A WordPress install comes with many PHP files, as well as images, style sheets, and simple plugins.

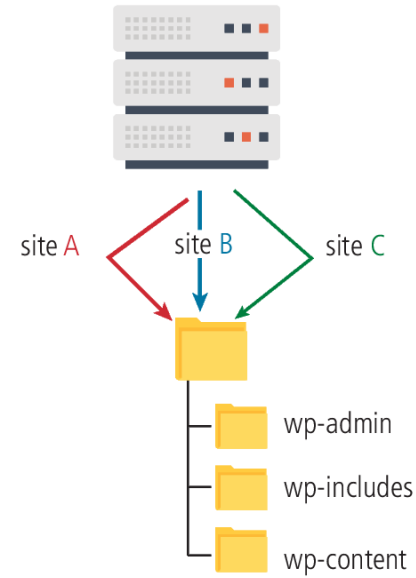
wp-content will contain files specific to your site including folders for user uploads, themes, templates, and plugins.



Multiple Sites with One WordPress



Server with multiple WordPress installations



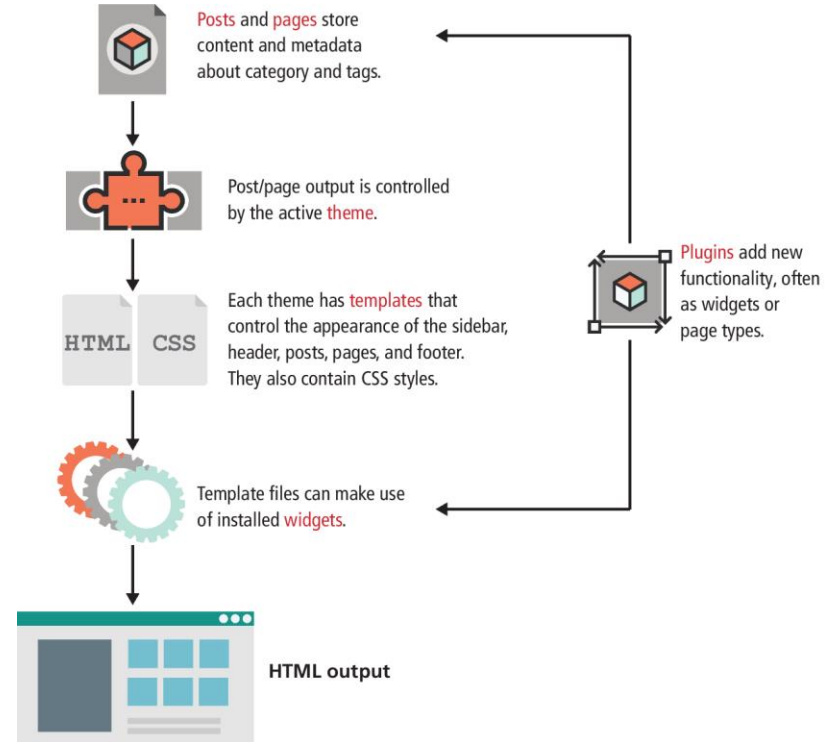
Multisite WordPress installation

WordPress Nomenclature

WordPress templates are the PHP files that control how content is pulled from the database and presented to the user.

WordPress themes are a collection of templates, images, styles, and other code snippets that together define the look and feel of your entire site.

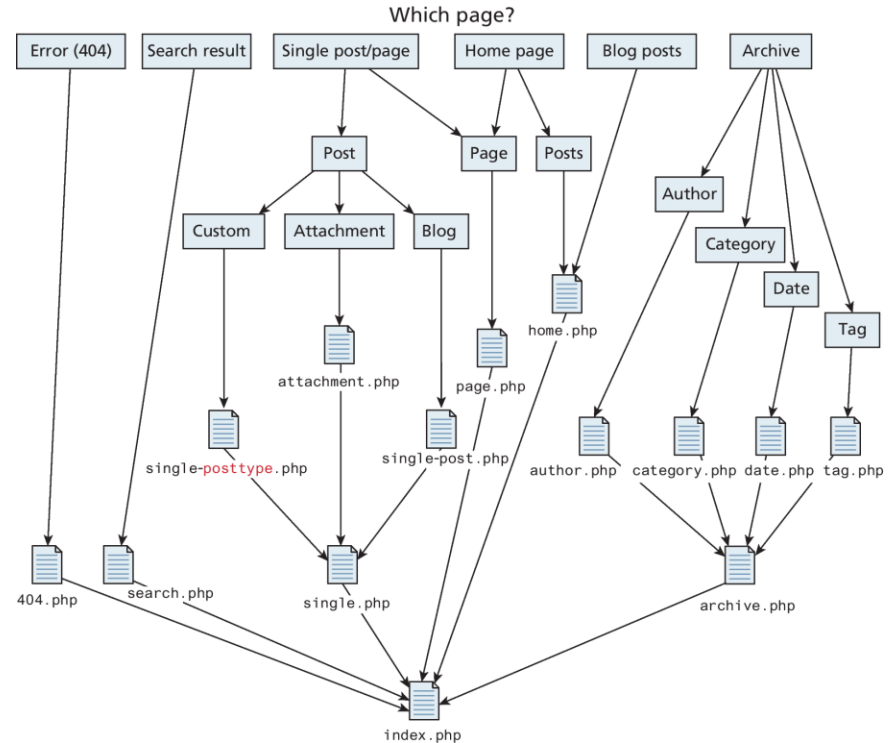
Plugins refer to the third-party add-ons that extend the functionality of WordPress



WordPress Template Hierarchy

The default WordPress installation comes with a default theme containing many templates

There are templates to display a single page or post, the home page, categories of posts including archive and categories



Modifying Themes

All the files you need to edit themes are found in the folder **/wp-content/themes/** with a subfolder containing every theme you have installed.

To start a child theme from an existing one where the only difference is a different **style.css** file, create a new folder on the server in the theme folder. In that folder create a **style.css** file with the comment from Listing 18.16

```
/*  
Theme Name: Twenty Sixteen Example Child  
Theme URI: http://funwebdev.com/  
Description: Theme to demonstrate child themes  
Author: Randy Connolly and Ricardo Hoar  
Author URI: http://funwebdev.com  
Template: twentysixteen  
Version: 1.0.0  
*/  
@import url("../twentysixteen/style.css");
```

LISTING 18.16 Comment to define a child theme and import its style sheet

Changing Theme Files

Best practice is to add the newly defined theme files to a child theme like the one we just started, leaving existing page templates alone. To tinker with the footer, we would make a copy of the existing footer.php in our new theme folder.

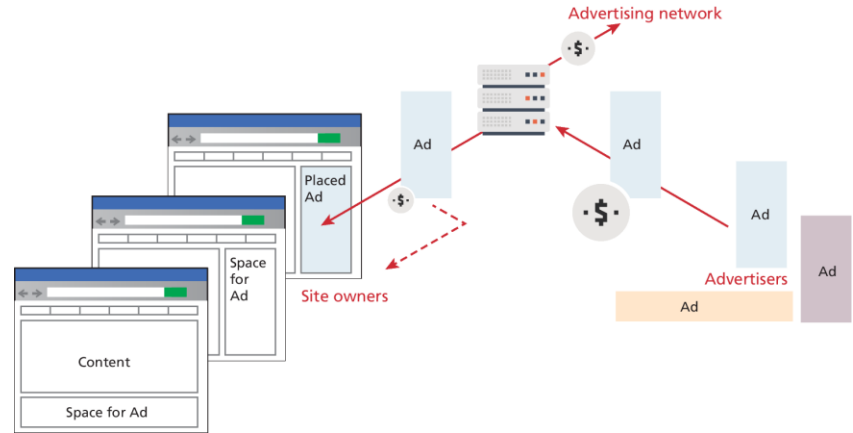
```
</div><!-- #main .wrapper -->
  <footer id="colophon" role="contentinfo">
    <div class="site-info">
      <a
href='http://funwebdev.com'>
        Supported by Fun Web Dev</a>
    </div><!-- .site-info -->
  </footer><!-- #colophon -->
</div><!-- #page -->
<?php wp_footer(); ?>
</body>
</html>
```

LISTING 18.17 A sample footer.php template file with the change from the original in red

Web Advertising Fundamentals

A website can become worth millions of dollars, and many of those millions of valuation are derived from projected advertising revenues.

- The vast majority of advertising is done through **advertising networks**.
- The actual advertisements are normally a little piece of JavaScript to embed on your page.



Ad Types

- There are many types of web advertisement that go beyond the basics such as the dreaded pop-up and the popular **interstitial ad** (where you must see the ad before proceeding to content).
- **Graphic ads** are the ones that serve a static image to the web browser. The image might contain text and graphics, enticing the user to click the ad, which will direct them to a URL.
- **Text ads** are low bandwidth, since they are entirely text-based. Like graphic ads, they too encourage the user to click and be directed to a destination URL.
- **Dynamic ads** are graphic ads with additional moving parts.

Web Advertising Commodities

- An **Ad View** (or *impression*) is a single viewing of an advertisement by a surfer. It is based on one loading of the page and although there may be multiple ads in the page, an impression is counted for each one.
- An **Ad Click** is an actual action by a surfer to go and check out the URL associated with an ad.
- An **Ad Action** is when the click on the ad results in a desired action on the advertiser's page. Advertisers may pay out, based on a successful account registration, survey completion, or product purchase, to name but a few.

Web Commodity Markets

- The **Cost per Click (CPC)** strategy is to decide how much money a click is worth, regardless of how many times it must be displayed.
- **Cost per Mille (CPM)** means cost per thousand impressions/views of the ad. Obviously this rate is lower than a CPC rate, since not every impression results in a click.
- The **Click-through Rate (CTR)** is the percentage of views that translate into clicks.
- **Cost per Action (CPA)** relates the cost of advertisement to some in-house action like buying a product, or filling out a registration form.

Support Tools and Analytics

Search Engine Webmaster Tools are tools from the search engines that show ways you can improve your site for the search engine's purposes

To sign up for these types of tools, go to **www.google.com/webmasters/tools/** and **<http://www.bing.com/webmaster>**.

Analytics refers to the class of useful software tools that provide website owners with data-driven information about their websites to help them make and assess change to their sites.

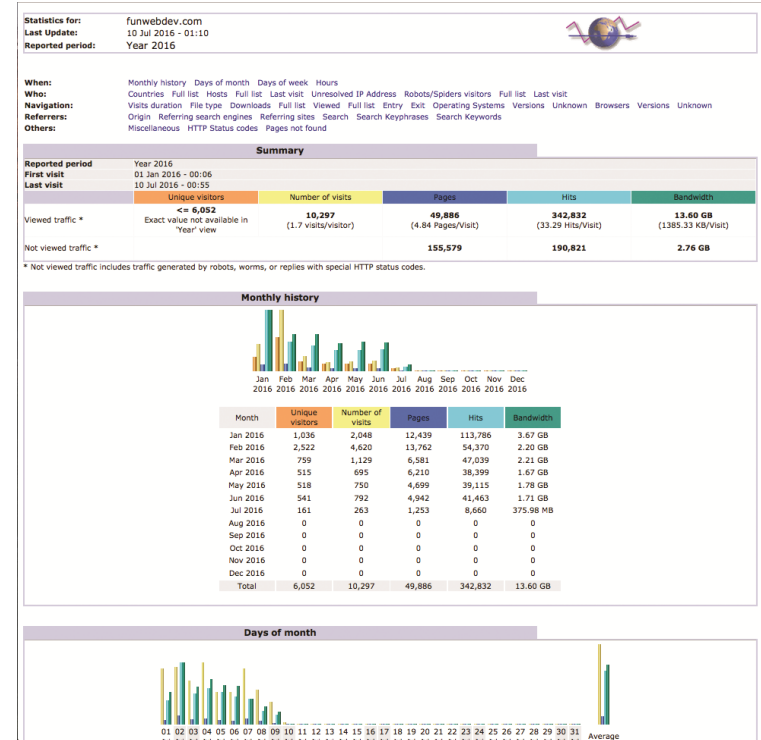
Metrics

Most metrics are included in analytics packages

- **Page Views** is a count of all the times a page was requested, even if requested multiple times by the same user/IP address.
- **Unique Page Views** counts page views but limits it to one request per page, per visit.
- **Average Visit Duration** tells you how long people are spending on your site. Longer visits indicate more engagement than shorter ones.
- **Bounce Rate** is the term given to the percentage of visitors who leave your site after visiting only one page.

Internal Analytics

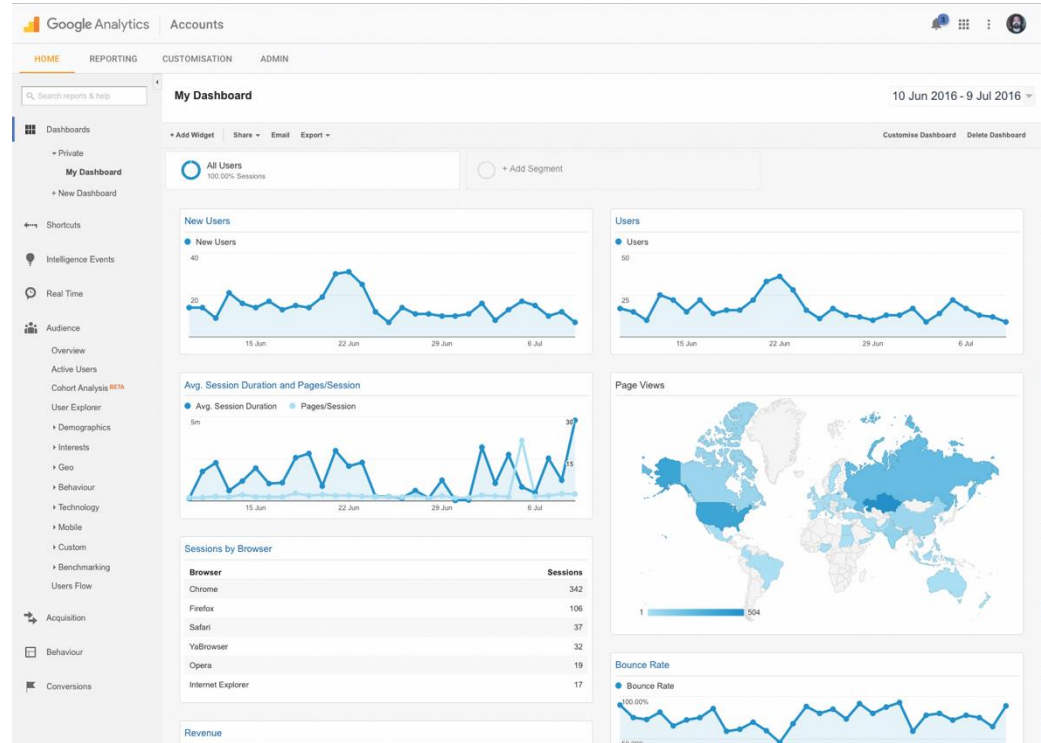
Open source analysis packages such as **AWStats** and **Webalizer** allow periodic analysis of server log files to create bar graphs; pie charts; and lists of top users, browsers, countries, and more



Third-Party Analytics

These tools are taking off in popularity, especially those offered by search engines like Google and Bing, which provide integration with other tools

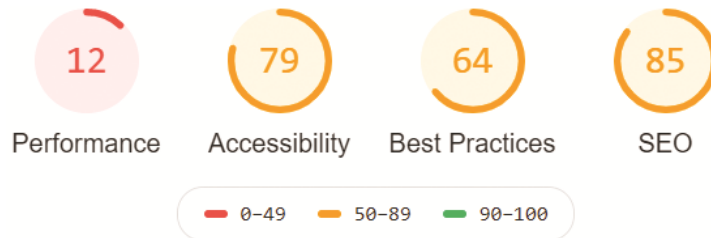
The advantage of third-party analytics is the increased power of these systems and the ease of installation. The disadvantage is the lower accuracy of data and disclosure of potentially valuable traffic information to the third party.



Performance Tuning and Rating

The **Lighthouse project** is an open source tool that provides analysis across a range of categories and makes some specific technical suggestions that are easy for the web developer to integrate. The tool is built into Chrome and can also be accessed on the web (<https://web.dev/measure/>).

Performance, **Accessibility**, **SEO** and **“best practices”** are the four categories used by the lighthouse tool, and provide a great way to evaluate and improve your website.



Speed Improvement



(a) Before Optimization



(b) After Optimization

Key Terms

Accessibility	black-hat SEO	Cost per Mille (CPM)	headless CMS
Ad Action	Bounce Rate	Click-through Rate (CTR)	hidden links
Ad Click	canonical	database engine	Indexes
ad hoc links	cloaking	Document management	input agents
Ad View	comment spamming	systems (DMSs)	interstitial ad
advertising networks	content creators	doorway pages	Keyword stuffing
Analytics	Content Management	Dynamic ads	Lighthouse project
anchor text	Systems	Email scrapers	Like button
asset management	content publishers	google	link farm
Average Visit Duration	comment spam	Google bombing	link pyramids
backlinks	Cost per Action (CPA)	Graphic ads	
best practices	Cost per Click (CPC)	Hadoop	

Key Terms (ii)

Link Spam	paid links	SEO	user role
menu control	Performance	site manager	Vulnerability Scrapers
meta tags	posts	sitemap	Web crawlers
metrics	Plugins	Social networks	web directories
navigation links	query server	stemming	What You See Is What
Newsfeeds	recurring links	super administrator	You Get (WYSIWYG)
Open Graph	reverse index	Template management	white-hat SEO
Open Graph meta tags	Robots Exclusion	Text ads	Workflow
Open Graph Objects	Standard	truncating a URL	WordPress
Page Views	Scrapers	Unique Page Views	templates
PageRank	Search engine optimization	URL Scrapers	WordPress themes
Pages	seeds	User management	

Copyright



This work is protected by United States copyright laws and is provided solely for the use of instructors in teaching their courses and assessing student learning. Dissemination or sale of any part of this work (including on the World Wide Web) will destroy the integrity of the work and is not permitted. The work and materials from it should never be made available to students except by instructors using the accompanying text in their classes. All recipients of this work are expected to abide by these restrictions and to honor the intended pedagogical purposes and the needs of other instructors who rely on these materials.