

Chapter 1

The Basics of Building Web Pages and Sites

If you have a completely solid, totally clear idea of exactly what you want to do on the Web, you can safely skip this chapter — maybe.

Are you absolutely certain that you have considered all the details? How does your site compare to other similar ones? Where does it fit in? What are the differences between them? Are those differences enough to distinguish your Web site, to attract visitors away from competing sites? Have you covered everything from setting a budget to grasping current Internet demographics? Has it occurred to you that you might want to avoid some apparently useful technologies?

If your answer to any of these questions isn't a resounding "yes," you should start right here and read this chapter before you do anything else.

Starting here really means stepping back and taking a look at the origins of the technologies we use today and that ubiquitously govern our communications, our transports, our trade, our military, our education, our medicine, our understanding of ourselves, and the direction of our society to come. These are astounding transformations, and we must spend a bit of time looking at the roots and vines to better understand the fruit of the Information Age. The future cannot be predicated from a present that has no past.

Therefore, this chapter begins with a discussion of the origin of Web browsers so that you may understand why browsers do what they do, which will allow you, the user, to determine what is the best browser for you. Similarly, when building your site, you must understand not only its infrastructure but also the workings of the audience (clientele) that will be utilizing your Web site so that you will be in a better position to efficiently

IN THIS CHAPTER

Web browsers

Web pages

Web sites

Analyzing Web site types

Understanding Internet demographics

Determining your approach

Dos and don'ts of Web site development

Part I Laying the Foundation

deliver services. A better understanding of how to structure the delivery method allows more effective delivery of the content itself. This chapter discusses the structure of a Web page using HTML to help you understand what coding a Web page is like.

Web Browsers

In the last few decades of the last century of the last millennium, small groups of men and women of genius and vision working at European and American research and academic locales developed the technologies that would result in the radical growth of the Internet worldwide. Although the development and growth of the Internet will be covered in some detail in Chapter 3, a bit of the background of the Internet is provided here in the context of the development of the Web browser.

The Web browser is a technological tool that allows access to networked Web pages. You may think that all the knowledge of the world lies just a few finger strokes away, but a good portion of all the Web pages created are inaccessible because of restrictions on their access. By one estimate, only 20 percent of all Web pages are freely accessed via the Internet. In short, a *Web browser* is a software application used to locate and display Web pages. Not all of these pages are on the Internet. Most corporations have extensive intranets with their own Web pages created and/or maintained by employees and departments.

The earliest graphical user interface (GUI) Web browser (actually called *WorldWideWeb*) was developed by Sir Tim Berners-Lee in 1990 at *Organisation européenne pour la recherche nucléaire* (CERN). The World Wide Web was so named by Berners-Lee after a lunchtime discussion with another CERN colleague, Robert Cailliau, who was a systems engineer and a strong proponent of a hypertext transfer program.

NOTE CERN was established in 1954 as an initial consortium of 11 European nations, to do fundamental research in physics. Today, about half of the fundamental research experiments are conducted at CERN. Its association with the Internet arose because, one of its researchers, Berners-Lee, sought (and created) a way to access hypertext documents on other sites. This discovery led him to create the first Web browser. Almost by default, he became creator of the first Web server. He also managed to create the protocol by which Web page content is sent across the Web (Hypertext Transfer Protocol, or HTTP). Of course, he also unavoidably created the first version of Hypertext Markup Language (HTML).

NOTE A graphical user interface (GUI) is icon-driven. Prior to GUIs, users would use a text-based process to gain access to their desired files, folders, and Web locations by typing in the destination. Today, we still have text-based Web browsers such as Lynx. An advantage of text-based browsers is that they allow much faster loading of Web page content over slow connections.

Berners-Lee also developed an addressing scheme (described later in this chapter) so that Web content could be reached, retrieved, and rendered. This first Web browser ran on a NeXT computer platform. The content of this first browser, *WorldWideWeb*, was rendered in grayscale and, because the art of placing images inline had not yet been mastered, images and icons were

The Basics of Building Web Pages and Sites

1

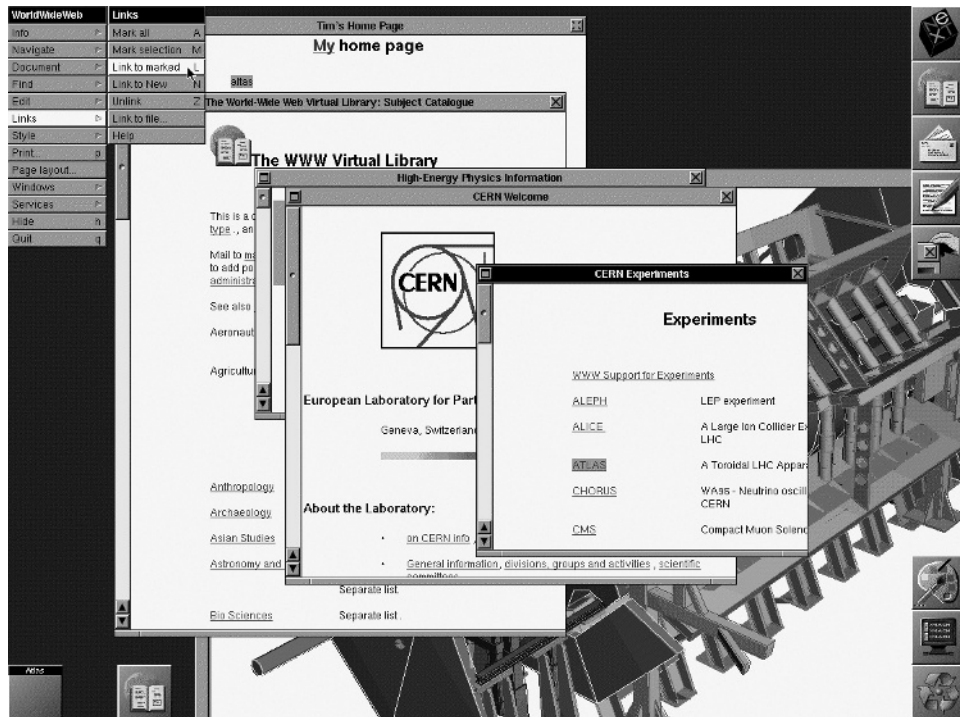
shown in separate windows. Figure 1-1 shows the first graphical browser created by Berners-Lee at CERN.

WorldWideWeb had significant font-manipulation capabilities. Users could adjust the font size, select from a range of font types, align elements, work with tables, lists, and so forth, and generally implement the features associated with HTML 1.0 — which was the dominant version of HTML from 1989 to approximately 1994. WorldWideWeb was the first browser to incorporate a basic HTML editor to allow users to work with and edit Web pages.

At almost the same time, a graduate student at University of California, Berkley named Pei Wei had independently developed a Web browser that was, in many respects, almost identical to WorldWideWeb. It also possessed a few features that were more advanced than the Berners-Lee model. It incorporated such advanced features as inline graphics, embedded stylesheets, tables, and scripts, as shown in Figure 1-2. It was, in the words of James Gillies and Robert Cailliau in their book, *How the Web Was Born: The Story of the World Wide Web* (Oxford University Press, 2000), “the most feature rich browser at that time” and “was to set the standard for everything to follow it. . . .” Certainly the *violaWWW* browser was good enough for CERN, which used it as the main browser for its lab units until the development of Mosaic.

FIGURE 1-1

WorldWideWeb was the first graphical browser.

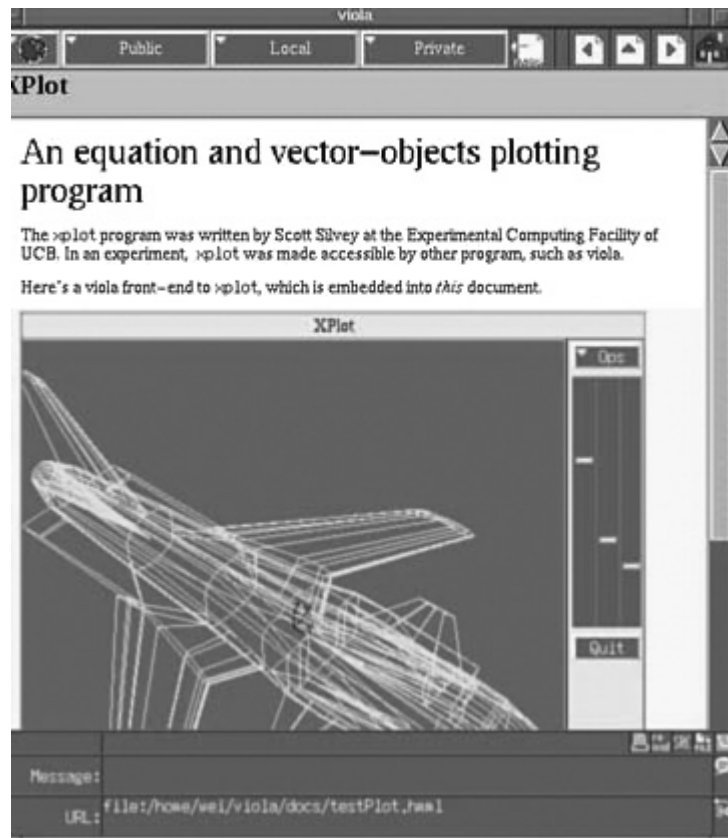


Part I

Laying the Foundation

FIGURE 1-2

The *violaWWW* browser was used by CERN before the arrival of *Mosaic*.

**NOTE**

Perhaps it was limited by only running on *Unix X-Windows* when the world was facing a looming showdown of Mac versus PC, or perhaps it was just too advanced for its time. Regardless of the reason, in this same time period, there was a mad rush to cash in on the Berners-Lee open source Web browser, *WorldWideWeb*, and commercialize it.

The *Mosaic* browser, shown in Figure 1-3, may be considered the first great commercial and popular success. Marc Andreessen and Eric Bina of the University of Illinois' National Center for Supercomputer Applications (NCSA) created and then heavily promoted this new commercially available portable browser. *Mosaic* improved upon the open source code of the noncommercial Web browser in a number of significant ways:

- It was user-focused, being relatively easy to install and use.
- It was cross-platform compatible with Unix and Apple (Macintosh).

The Basics of Building Web Pages and Sites

1

- It was the first commonly available browser to have inline images.
- It offered support for a number of graphics formats.
- It was able to render both video and audio formats.
- It introduced as add-ons the features of history (which allowed the tracking of sites visited) and bookmarks.

The Mosaic browser was aggressively marketed with a hitherto unknown 24-hour technical support for its client base. Problems were efficiently monitored, tagged, and resolved. In a matter of months, Bina improved Mosaic, and it became known as *Netscape*. Ultimately, NCSA asserted its rights over Mosaic and the innovations developed in Mosaic were licensed to Microsoft (among several other companies). Hence, the *Internet Explorer (IE)* browser was originally built around the Mosaic core. Development work on Mosaic officially ceased in 1997.

FIGURE 1-3

The Mosaic browser introduced a number of new features.



Part I

Laying the Foundation

Other browsers were also being rapidly developed by other academic or government institutions, using different operating systems and using different ways of developing and executing what was evolving as the standard core of browser features:

- Fetching
- Decoding
- Faithfully rendering Web pages, history, and bookmarks
- Cross-platform compatibilities
- Scripts
- Embedded objects
- Interactive forms

NOTE

For more in-depth information on various browsers, see the support tables at http://en.wikipedia.org/wiki/Comparison_of_web_browsers.

These days, the search for the best (the fastest, most secure, and most feature-rich) browser continues, and some of the newer players (upon close inspection) seem eerily familiar. For example, Firefox (2004) from Mozilla (see Figure 1-4) is actually a direct descendant of open source Netscape (2002) — known an even earlier time as Mosaic.

FIGURE 1-4

The Mozilla browser is a descendant of Netscape.



The Basics of Building Web Pages and Sites

1

NOTE

There is, as Paul Harvey would say, the rest of the story. It may concern an Easter egg. Those who have Firefox let them see...about:Mozilla. *Book of Mozilla* 7:15.

As for the best of the best, about two years ago extensive tests were run by a very committed, highly organized, and methodical programmer named Mark Wilton-Jones in the United Kingdom. His subsequent tables of results were endlessly copied in both print and electronic media. If you are curious (and it is quite interesting material), you should visit his Web site at www.howtcreate.co.uk/browserSpeed.html to see his rigorously tested results.

Web Pages

A *Web page* is a document that is written in (or convertible to) HTML and that is accessible to Web browsers so that they can upload or download content. The Web page, if online, is found by an address. The address is in the format of <http://www.goto.com/goinggoing/gone> and is frequently referred to as a *Uniform Resource Locator* (URL).

NOTE

This chapter does not delve into a technical discussion regarding the differences between *URL*, *URI*, and *URN*. Berners-Lee (the creator of the above address schema), uses universal resource identifier (URI) when referring to an Internet address. The reference to the address of the Web page is URL for academics and for the rest of us. For more information on the differences in terminology, see <http://gbiv.com/protocols/uri/rfc/rfc3986.html#RFC1630>, which provides a thorough explanation of the reasoning behind the development of specific features of Web addresses.

URLs (or addresses) are broken down into three parts:

- *Protocol* — This is the name of the “language” that is used to transport the data across the network from source address to destination address. The language used to transfer document content is Hypertext Transport Protocol (HTTP). Recall that the format of the Web page is Hypertext Markup Language (HTML), so it is easy enough to associate the transport method (HTTP) with what is being transported (HTML), since both are *hypertext*. The two forward slashes following the protocol essentially indicate to the Web browser, “this is where you start your business and ignore anything to the left of the //.”
- *Domain* — This is a two-part name separated by a dot (.) that identifies a particular institution or group. The first section before the dot is some specific corporate entity, and the part following the dot is a suffix that defines the top-level domain of the cited entity. For example, yahoo.com indicates an organization referenced by the domain name. Yahoo belongs to the top-level domain, and [com](http://yahoo.com) further defines [yahoo](http://yahoo.com) as a commercial business. Other top-level domains with which you are probably already familiar include gov for government sites, org for nonprofits, mil for military sites, net for network organizations, mobi for mobile, and edu for educational sites.
- *Web page (file or folder)* — After the domain name comes the forward slash (/) or multiple slashes that define the path to the specific Web page. The addressing scheme is hierarchical, and the address goes from the general to the specific. The Web page the

Part I Laying the Foundation

user is fetching could be located directly on the domain, or within a folder within the domain. For example, if you have a Web page called Crowder, stored directly on the wiley site, the format of the address would be `http://wiley.com/Crowder`. If the Web page, however, were stored within a folder called, say, CWSB, the file (or Web page) reference would be: `http://www.wiley.com/CWSB/Crowder`. The Web page Crowder is stored within the folder CWSB, which itself is part of the wiley.com domain.

The `www` prefixing the domain name indicates that the user is referencing a networked Internet server that is hosting a specially formatted document — in other words, a Web page. This page is formatted in the markup language, HTML, a simplified subset of *Standard Generalized Markup Language* (SGML). HTML is a (more or less) standardized system for manipulating, labeling, and “accessorizing” document content and structure. Chapter 4 goes into more detail about the structure of the Web page using element tags.

IP addresses and domain names

Actually, the addressing scheme is a tad more complicated than the way it is represented here because all addresses are actually numbers and not names at all. Therefore, when you type in `yahoo.com`, this is converted by a *domain name server* (DNS) into its numeric equivalent of 209.191.93.52, which would most likely be the address of one of Yahoo’s numerous servers.

IP addresses, also known as *network addresses*, are used by the Internet to locate and retrieve data from specific hosts within a specified network. Most Internet users access the Internet through an Internet service provider (ISP) that has a range of IP address available. These are assigned by the nonprofit corporation *Internet Corporation for Assigned Names and Numbers* (ICANN). These addresses are portals or doorways to the Internet, and are used by multiple users for varying lengths of time. If they checked their IP addresses, users might note that addresses dynamically change each time they log off and log on to the Internet. If their IP number is, for example, a 172.16.x.x network number, that address is internal to the ISP and not a routable Internet address — even though, through it, users have access to the Internet features to which they have subscribed. `run ipconfig` is one of several ways to check the user IP address and default gateway. The *subnet mask*, which also appears in the data about IP addresses, is for another book.

NOTE

Another great way to learn your IP address on the Internet, rather than your potentially internal address, is www.whatismyip.com.

Use of HTTP and FTP

There may be some confusion regarding the use of HTTP and File Transfer Protocol (FTP) as Web *transport protocols* because both transfer files. The distinction would be that HTTP is used to transfer Web page content to a browser from a Web server and the service is one way: from server

continued

The Basics of Building Web Pages and Sites

1

continued

to browser. FTP is used to upload and download files to and from a workstation to a file server. Another way to remember the distinction is that HTTP uses port 80, whereas FTP uses ports 21 and 20 (command and data).

A Web page is written in HTML or *Extensible HTML (XHTML)*. There are various versions of HTML, each of which gives greater flexibility to the developer of the Web site to add features such as interactive forms, embedded scripts, stylesheets, accessibility features and so on. HTML v1.0 provided support for inline images and font manipulation. HTML v2.0 gave interactive elements to forms (for example, submit, exit, and clear buttons; text boxes or areas; labels; check boxes; radio buttons; and so on). Most browsers have this level of Web page support. HTML v3.0 features were never widely used, and the latest version, v4.02, added support for scripting, multimedia, and stylesheets.

The common features of a simple Web page consist of the following:

- *Head* — This section holds information about the Web page. Comment lines within this section typically provide information on the programmer and program purpose. These also contain metatags that are used by search engines. These are used to set the title, load page scripts, and set special tags (such as META tags) to help other browsers and systems use the page. Though there may be executable calls to *functions* contained in the head, none of the content of the head section appears directly on the Web page.
- *Body* — This is where the content is formatted, and it is from the body that the content is displayed. Typical tags in the body would be `<p>` for paragraph (where text and inline images would appear), `` for a list that is unordered (“ul” meaning “unordered list”) with a line-by-line column of each item in the list, `` for a horizontal line to attractively divide the sections of the Web page, and `<para>` for an address.

CROSS-REF Chapter 4 provides much more detail about HTML and its basic elements. This discussion provides an introductory examination.

Elements are colloquially referred to as *tags*, (as in *tagging* an item such as clothes for sale). Note that all elements (except for two) appear in pairs, indicating an opening tag and a closing tag:

- `<html></html>`
- `<p></p>`
- ``
- ``
- `<p></p>`

Within each tag pair is the content to be displayed on the Web page.

`
` and `<hr />` are both examples of what are called *empty tags* because they relate to the appearance or format of the Web page but include no content themselves.

Part I Laying the Foundation

A simple setup for a Web page would be as follows:

```
<HTML>
<HEAd>
<title> A simple web page</title>
</head>
<body>
<P>This is paragraph one.</p>
<p>This is paragraph two.</p>
<p>and this is paragraph three.</p>
<ul>
<li>Item 1</li>
<li>Item 2</li>
<li>Item 3</li>
</ul>
<br>
<hr>
<address>street, city, state</address>
</body>
</html>
```

Note that HTML is not case-sensitive and does not care what is uppercase and/or lowercase in these elements. To the thousands of users challenged by pressing the Shift key correctly every time, count your blessings. However, the content in between the tags won't be automatically corrected, so make sure that you capitalize words where necessary in the content you want to display to the user of your Web page. In more advanced coding that deals with attributes, you will learn that case does occasionally matter.

You will also note that every element in this example appears in a straight line against the left margin. This formatting may provide for easier reading, but it is not necessary to produce effective HTML code. Because HTML is totally insensitive to whitespaces, you could also enter the previous example as follows:

```
<HTML><HEAd><title>A simple web page</title></head><body>
<P>This is paragraph one.</p><p>This is paragraph two</p>
<p>and this is paragraph three.</p>
<ul><li>Item 1</li><li>Item 2</li><li>Item 3</li></ul>
<br><hr>

<address>street, city, state</address></body>
</html>
```

NOTE

Many Web browsers will fix small HTML errors automatically, since there are so many different ways of interpreting HTML and all its different versions. This is sometimes known as “promiscuous” or “dirty rendering” mode.

You can see how annoying this would be trying to read, however, especially if you were trying to find where you made a typo. So, by convention, HTML is printed nicely, line by line, one command per line, and sometimes indented to distinguish related sections (such as when using nesting in if-else command structures).

The Basics of Building Web Pages and Sites

1

You'll see many examples of this (and do it yourself) later in this book as you are developing your own pages.

XHTML (which came out in 2001) was designed to provide a strict structure currently lacking in HTML coding. The reason for some of the strict structure may be obvious, but some of it may be needlessly complicated.

NOTE

One reason for XHTML is so that we know how to detect a properly formatted page. By requiring end tags, and setting more-strict formatting requirements, the browsers can better render the pages.

Back in the “good old days,” capitalization of elements didn't matter — `<head>`, `hEaD>`, `<Head>`, `<head>`, and `<HEAD>` were all the same. Size now matters, and tags must all be lowercase. The format of the single-sided tags has changed, as has the closing of tags. Some elements have absolute requirements that were optional in the past. Some commands (in the interest of consistency to the object-oriented paradigm) have been made more attribute-heavy. The command for embedding a video clip, for example, changed from the simple `<embed src = "url">` to `<object data = "url" type = "media spec" width = "x" height = "y"></object>`. Another issue is that a number of commonly used tags are no longer compatible with strict XHTML (such as `label`, `map`, `span`, `input`, `img`, `fieldset`, `iframe`, `button`, `cite`, `em`, and `form`). You should be aware that there would be compatibility issues with earlier versions of HTML.

It is very important to also be aware that a Web page will appear differently on different computers. Some computers may have different screen resolutions than the one used by the designer. A user may have a browser (remember that there are easily 50 or more of them available) that may not have the capacity to render all features of the Web page. The latest version of *Cascading Style Sheets* (CSS), the range of acceptable colors, or fonts may not be available. All this boils down to is this: a Web page developer must test the Web page on different browsers to verify the correctness of their appearances, and the designer may also want to design with a version of HTML that is acceptable to the largest possible audience. This might even mean forgoing some multimedia or interactive features.

CROSS-REF

See Chapter 13 for more information on Cascading Style Sheets.

Web Sites

There is no single accepted typology of Web sites. Web sites may be divided into several types and subtypes, such as the following:

- Static and dynamic
- Individual or group
- Domain types (such as `gov`, `mil`, `co`, `com`, `net`, `edu`, `mob`) grouped by shared interests (such as download or humor sites, review sites, news sites, or search sites)

Part I Laying the Foundation

And there are numerous extended classifications. Later in this chapter, you will see one breakdown of Web sites, and others will be presented throughout this book as appropriate to the topic. This section offers a brief overview of links, which are the single most critical shared feature of all Web sites. Other important technologies for Web site functions (including scripts, multimedia tools, and stylesheets) are examined in separate chapters throughout the book.

A *Web site* consists of a number of associated Web pages for an entity that share a common theme. The most critical part of a Web site is its linking structure. A Web site cannot exist without a way of linking together the Web pages of which it consists. These linked Web pages are, hopefully, related in a logical fashion that allows intuitive and speedy access to any of the other pages in the Web site from any location within the site. Depending on the number of linked pages and the number of levels of pages in each of those links on the home page, Web pages may be linked *linearly*, *hierarchically*, or a mixture of both (sometimes called an *augmented linear structure*).

NOTE

Hidebound traditionalists prefer links the plain, old-fashioned way: in blue and underlined. A designer might even go so far as to change colors for a visited link. But the truth is that you can do all sorts of funky stuff with links: you can change colors, or sizes, or fonts; you can make them look like ordinary text in the Web page; you can use images as links; you can even use multiple links within an image (known as *hotspots*). The details of how to do all this are reserved for later chapters in this book, but be aware that what you can do with links is only limited by your sense of the creative.

Sites with a large number of pages (for example, seven or more) use a full or modified hierarchical structure that incorporates the linear structure at specific levels. A hierarchical structure moves in layers from the general to the specific. It is very much like an inverted tree structure in which the branches spring from the trunk, and from them spring smaller limbs and then twigs and then leaves. This allows the user to get to an area of interest by using a few mouse clicks, or to move rapidly to another area or areas within the site.

At the top would be the single point of entry into the Web site: the *home page*. The purpose of the home page is to provide an overview of the organization or individual presenting the site, generate consumer interest, and allow quick and useful access into the site. There are several formats for offering links, and numerous good examples exist on the Web. One of the most helpful navigation features of a home page is a drop-down list of most frequently accessed links (*quick links*). Navigation within the site should be intuitive, and entry in and out of a page should be easy. Fundamentally, the successful home page allows access to site info in small structured bites.

Figure 1-5 shows an example of such a well-designed home page for Emory and Henry College. In this example, the home screen acts as a portal to its various divisions.

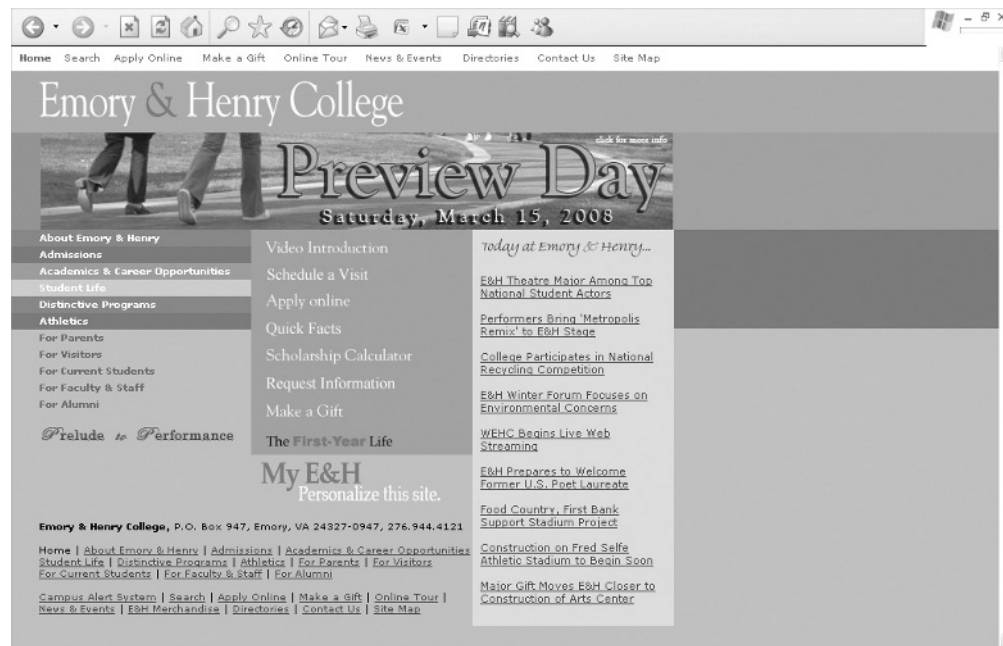
Click on any of the sections and you will be conveyed to a Web page on that selection, and the Web page will have a quite similar layout to the home page. Consider the pages of a printed book, where each page is formatted exactly like every other page, and only what is written (the content) distinguishes the pages. (The page number would be the navigation tool for getting back and forth.) Similarly, for a Web site, the formatting would be a faithful copy of the main (or home) screen. The background would be the same, as would the font styles and colors, the layout of the screen, and its dimensions.

The Basics of Building Web Pages and Sites

1

FIGURE 1-5

Home page of Emory and Henry College.



CROSS-REF Chapter 4 provides more information about designing the structure of the Web site. At this point, however, you should at least appreciate the criticality of sitting down and drawing out your site before doing the fun stuff (the content).

Though it is tempting to try to create the latest “gee whiz” site utilizing all sorts of bells and whistles to dazzle your peers, the plain truth is that Web sites are not only as good as their designers; they are also only as good as their users’ connection to the Internet, their users’ Web browsers, their users’ screen resolution, and their users’ general level of comfort with navigating the Web.

So, plan wisely — and well.

Analyzing Web Site Types

You could argue endlessly about how to categorize the overwhelming variety of sites on the World Wide Web. For all practical purposes, however, there are only five basic types of sites:

- Personal
- Informational

Part I**Laying the Foundation**

- Organizational
- Political
- Commercial

As with any attempt at *taxonomy* (that is, the art of lumping disparate entities into categories based on similarities), there are gray areas, fuzzy situations, and downright overlaps. What do you make of a personal home page that's largely political? In the system discussed here, it's still a personal home page. What about a genealogy site? Is that an extension of a personal home page, or is it an informational site? We vote for it being informational, and labeling a site "personal" only if its focus is limited to an individual and his or her immediate family and friends.

To categorize sites, the basic method we use is to try to determine the intent of the Web site. If it's to make money, then it's a commercial site, no matter what other angles there are to it. If it's to promote a particular candidate in an election, then it's a political site, even if it's made by one individual.

Personal home pages

Personal home pages exist for the purpose of introducing individuals to the world. Despite the phrase "personal home page" (which has become generically accepted as a description of this type of Web site), some personal sites involve many different Web pages and can be as elaborate as a corporate Web site. Generally speaking, the information in them is somewhat limited in scope, appealing mainly to the close friends and immediate family of the person involved. This isn't to say that some of them aren't very interesting, indeed. As with any other type of site, a carefully designed and well-written production is a real joy to explore and experience.

Just as when talking to strangers at a party you'd find some people are fascinating and have great stories to tell, whereas others send you running from the room (yawning from the boredom), personal Web sites can be a similar experience. Some home pages grab you, and others make you wish you had never clicked that link.

When putting together a personal site, it's a good idea to open up yourself as much as possible within your comfort level. The best personal sites enable you to get to know the people who created them. By the time you're done browsing these sites, you feel as if you could sit down to dinner with their authors and talk as if old friends. This isn't to say that you must spill your guts about every detail of your personal life, but the site is about you — and if *you* aren't really in it, what's the point of creating it?

Informational sites

Sites that are dedicated to providing information on a particular topic are common on the Web. If they weren't, it would be a pretty boring place. In fact, most people use the Web primarily to find the answer to some sort of question. Many of the informational Web sites are run by public-minded organizations whose goal is simply to make the public aware of particular issues.

The Basics of Building Web Pages and Sites

1

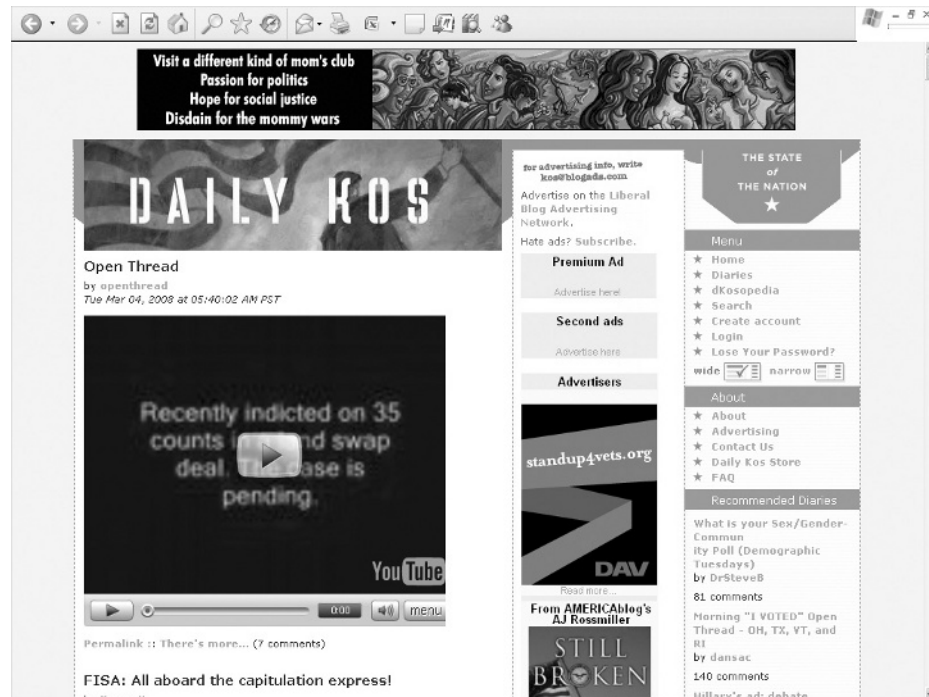
The Daily Kos political blog, for example, provides much needed clarification about the confusing political situation within the Democratic party and the state of our nation at www.dailykos.com.

Search engines, on the other hand, are a perfect example of the kind of gray area discussed earlier, where it's difficult to slot many of them into one category. The vast majority of them are commercial operations designed to generate advertising revenue. Some of them even rank the results by how much they're paid instead of by true relevance. But the only interest that Web surfers have in them is to find information. The same is true of news sites.

Another type of informational site is one that offers a limited amount of information at no charge but that provides a greater amount for a fee. Spanish Learning Resources (www.studyspanish.com) is a prime example of this. The site, shown in Figure 1-7, provides about as much training as you would get in a typical first-year Spanish language course for free. If you simply need to understand such niceties as why *ser* and *estar* both mean “to be” or need a basic grasp of Spanish vocabulary and grammar, visit this site at www.studyspanish.com. (You should still pick up a copy of *Spanish for Dummies*, though — you never know when your Internet connection might go down.)

FIGURE 1-6

Political blogs can increase public awareness on civic concerns.



Part I

Laying the Foundation

FIGURE 1-7

Spanish Learning Resources provides both free and premium information services.

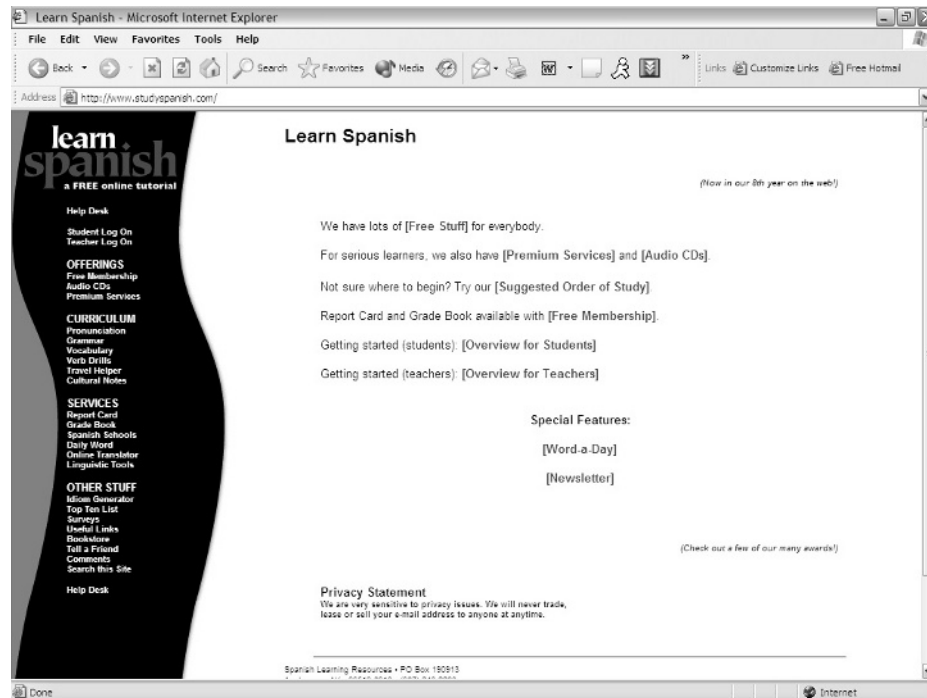


Table 1-1 lists the URLs of several informational sites that you can browse and study.

Organizational sites

Organizational Web sites are concerned with presenting information about — you guessed it — a particular organization. The reason these are distinguished from pure informational sites is their narrow focus. Their sole function is to describe the work of a particular organization, its structure, its personnel, and history. These sites don't diverge from their function at any point — no sidetracks to outside issues or topics exist, except for occasional links to other sites that partner with or support the goals and purposes of the organization.

As shown in Figure 1-8, the Internet Engineering Task Force (IETF) site, located at ietf.cnri.reston.va.us, is a classic example of an organizational site. It covers everything from the structure of the organization to the process for joining it.

Table 1-2 gives the addresses of some organizational sites on the World Wide Web.

The Basics of Building Web Pages and Sites

1

TABLE 1-1

Informational Web Sites

Site	URL
AcqWeb's Directory of Book Reviews on the Web	acqweb.library.vanderbilt.edu/acqweb/bookrev.html
AnyWho	www.tollfree.att.net
Catalog of Solar Eclipses	sunearth.gsfc.nasa.gov/eclipse/SEcat/SEcatalog.html
CIA World Factbook	odci.gov/cia/publications/factbook
CIAC	ciac.llnl.gov/ciac/index.html
Dead People Server	www.dpsinfo.com/dps/
Indiana Department of Agriculture	ai.org/oqa
Internet Traffic Report	InternetTrafficReport.com
King Arthur and the Matter of Britain	legends.dm.net/kingarthur
Letters and Dispatches of Horatio Nelson	wtj.com/archives/nelson
Sir Francis Drake	www.mcn.org/2/oseeler/drake.htm
Veterinary Information Network	vin.com

Political sites

Sites that have a political agenda to push aren't limited to dealing with particular candidates for certain offices in election years. Politics isn't all about elections. In our definition, such sites also include those with any social agenda that would be furthered by the passing of appropriate legislation. Thus, Web sites covering issues such as gun control, the medical usage of marijuana, development of wetland areas, or the control of personal information on the Internet — regardless of whether the site's editorial thrust is for or against — would qualify as political sites.

Figure 1-9 shows the Web site for the Democratic National Committee (DNC) at www.democrats.org. It could be argued that it is, in fact, an organizational Web site, representing as it does the DNC, its views, and activities. Because those activities are strictly political, however, it belongs here.

Table 1-3 provides the URLs of several political Web sites.

Commercial sites

This is the one simple, unambiguous area in the taxonomy — if the site is created strictly out of a profit motive, then it's a commercial site. Period. Other sites — particularly political sites — often

Part I Laying the Foundation

engage in some kind of fund-raising activity that is secondary to their main purpose. With purely commercial sites, there is never any doubt — the main thrust is simply “Buy our product or service” or “Support our sponsors.”

The Barnes & Noble Web site (www.bn.com), shown in Figure 1-10, goes straight to the point. It's there for one thing and one thing only — to enable you to find and buy books and music.

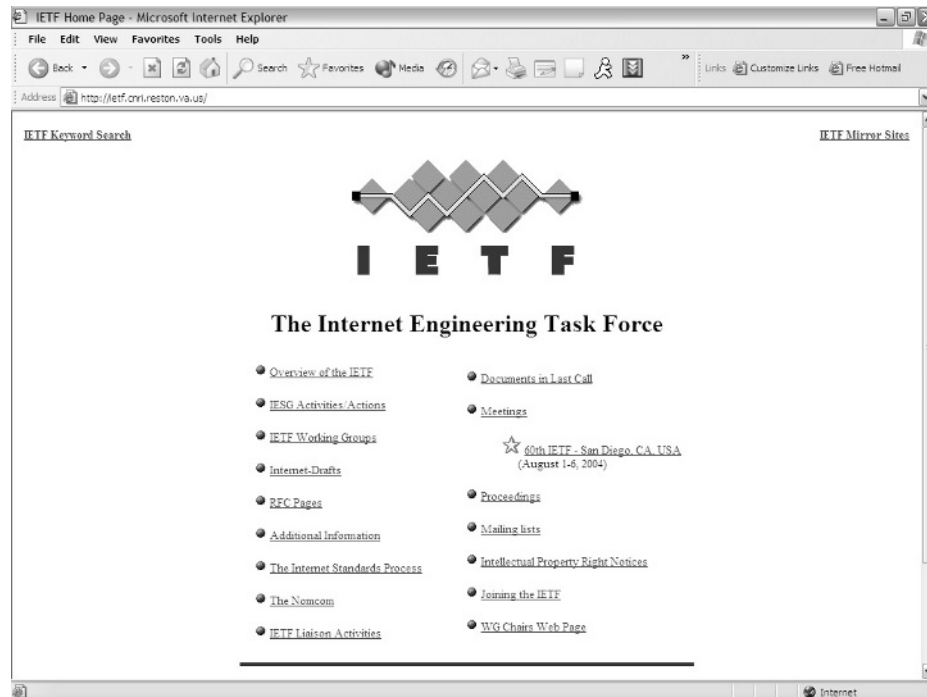
Simply accepting advertising or tossing in a few links to a bookstore doesn't magically transform another kind of site into a commercial Web site. There are few sites in any category that do not attempt to raise some money (often unsuccessfully) by a half-hearted foray into banner ads.

Sometimes, especially for sponsored sites or those which are participants in an affiliates program, the “buy” message is so subtle that the commercial nature of the site isn't obvious to the casual visitor. Many times, there is no mention at all of any sales intent, but the site's links lead to sponsored pages.

CROSS-REF See Chapter 24 for more information on affiliates programs and advertising.

FIGURE 1-8

The Internet Engineering Task Force (IETF) Web site has a full explanation of the structure, goals, and functions of the organization.



The Basics of Building Web Pages and Sites

1

TABLE 1-2

Organizational Web Sites

Site	URL
International Center for Reiki Training	reiki.org
JustLinux	justlinux.com
National Weather Service	www.nws.noaa.gov
Nautical Research Guild	naut-res-guild.org
Nicholas Roerich Museum	roerich.org
Small Business Administration	www.sba.gov
State Library of New South Wales	www.sl.nsw.gov.au
The Naval Dockyards Society	www.hants.gov.uk/navaldockyard
The Royal Society	www.royalsoc.ac.uk
The Society for Creative Anachronism	sca.org
World Wide Web Consortium	w3.org

Table 1-4 guides you to a number of commercial Web sites.

We have just reviewed the five generally accepted categories of Web sites. The success of a Web site is dependent upon how well it defines its potential clientele. There is no monolithic client base, and much research has allowed potential users to be divided into so called B2B (business to business), B2C (business to consumer), C2C (consumer to consumer), and C2B (consumer to business) business categories. Ninety-seven percent of the online market is B2B.

The purpose of the following section is to introduce you to how these online markets are defined so that the categories of Web sites may match the services required of each particular category of user. For example, B2B would be likely to handle services and products that go into the production of a final product. Let's say the construction of this book is an illustration of B2B. We have an author who provides the raw material. We have a development editor and technical editor(s) who refine the raw material and process it into something. The general editor reviews our work and adds the polish. All of that is B2B. The book, when it finally appears at bookstores, is a B2C.

The way you relate to (or market) B2B would be different than the way the book would be presented to you by a bookseller. You structure your online Web site according to the type of clientele you market to.

That is why you need to understand the demographics of the online marketplace.

Part I

Laying the Foundation

FIGURE 1-9

The Democratic National Committee site is a prime example of a political Web site.

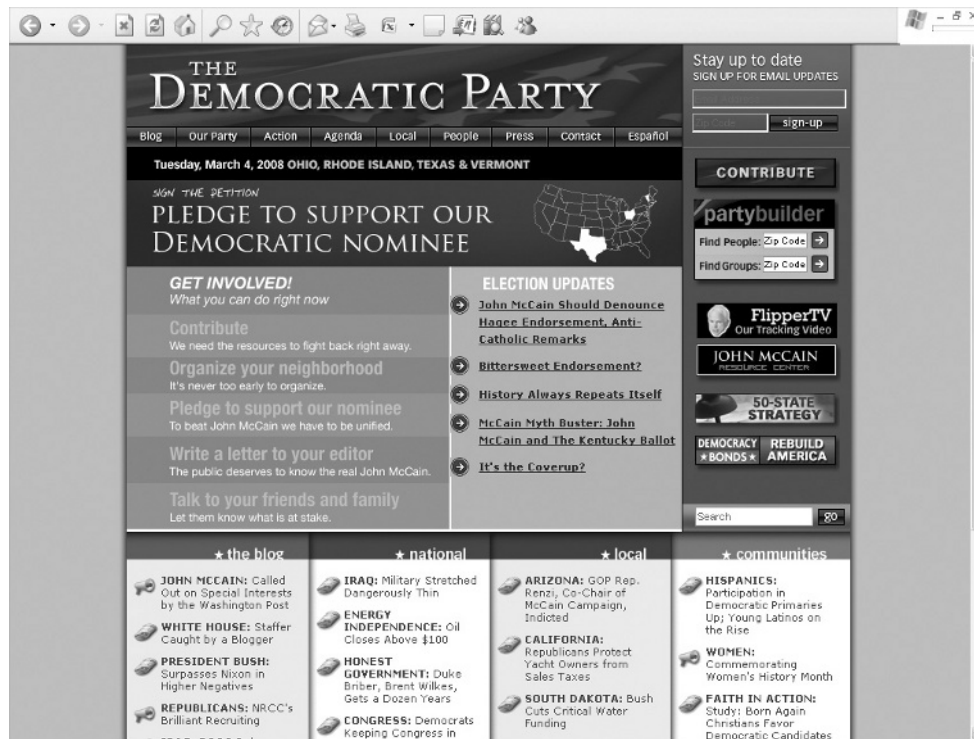


TABLE 1-3

Political Web Sites

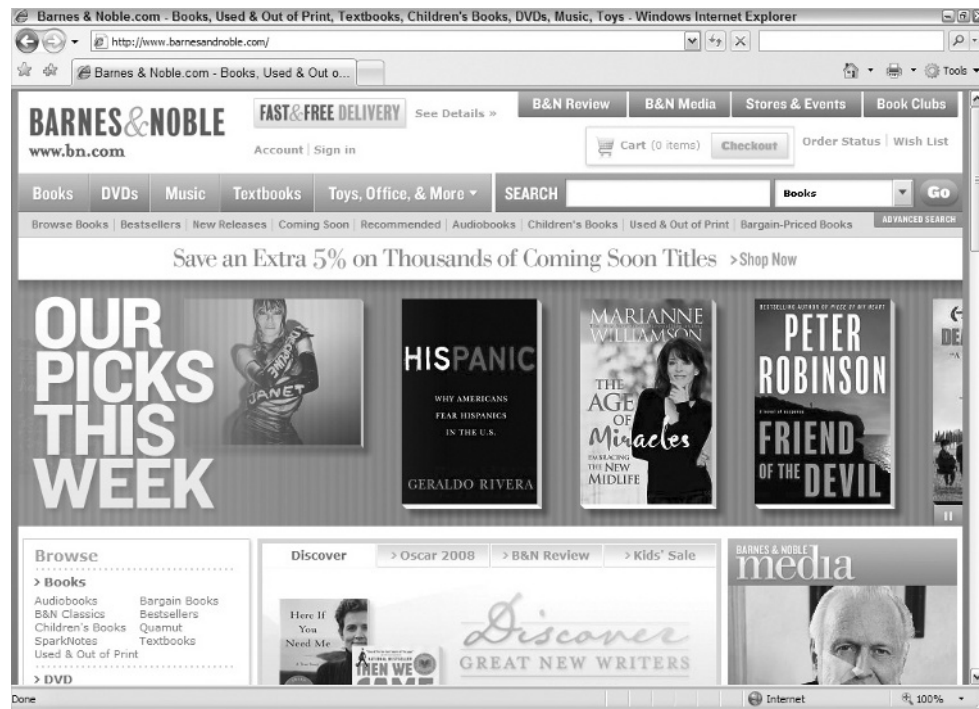
Site	URL
Australian Politics Online	www.ozpolitics.com.au
Chinese Politics Online	www.politics.unimelb.edu.au/cpo
Conservative Party Website	www.conservatives.com/home.cfm
Labour Party Website	www.labour.org.uk
Republican National Committee	rnc.org
Southeast European Politics Online	www.seep.ceu.hu
Washington Center for Politics & Journalism	wcpj.org

The Basics of Building Web Pages and Sites

1

FIGURE 1-10

The Barnes & Noble Web site is a prime example of a well-designed and successful commercial Web site.



Understanding Internet Demographics

There are tons of “facts” available about the Internet. Unfortunately, most of these statistics can be classified by one of Mark Twain’s most famous quips, “There are lies, damned lies, and statistics.” As with so many of the greatest American writer’s sayings, the truth of this rides into our present and, doubtless, our future as well.

Finding good information

It’s easy to find anything you want to know about the Internet — just look on the Internet. But how do you know if you can trust what you find? One of the simplest ways is also one of the best — know the reputation of the source. If someone’s Web site says that there are 8 billion people with Internet access, you might want to know who the source is and how many planets the source thinks are hooked up to the Internet.

Part I

Laying the Foundation

TABLE 1-4

Commercial Web Sites

Site	URL
Australian Bush Flower Essences	www.ausflowers.com.au
Gibson Research Corporation	grc.com
Jagex Software	jagex.com
Janes Information Group	www.janes.com
Magix Entertainment Software	www.magix.net
Ptaah	www.ptaah.com/home.html
Staples Office Supplies	staples.com
The History Channel	www.historychannel.com/ontv/index.html
TV Guide	tvguide.com

The same is true, however, of even less obviously suspect figures. News stories are, unfortunately, one of the least reliable of all statistical sources, because they are generally written in a hurry by people who do not really understand statistics. They also often fall prey to the fault of accepting statistics from yet other news stories. This can easily lead to the kind of situation where one reporter gets something wrong, but the story is published in a major publication. Because another reporter believes that anything in *The Wall Street Journal* or *The New York Times* must be true, the figure is incorporated into yet another article, which is taken as gospel by someone else, creating a chain that perpetuates a number that is far removed from its original source.

We'd be inclined to trust (without more than a reflexive professional doubt) a report from companies such as comScore (see Figure 1-11), Nielsen, or WebSideStory, all of which are professional outfits that have proved reliable in the past.

Table 1-5 lists the URLs of some of the top Internet research firms.

You can never really trust any statistics, no matter what the source, unless you know how those figures were arrived at — not if you're going to bet your future on them. Any reputable research company is glad to provide you with the details of the methodology they used in arriving at certain figures.

NOTE

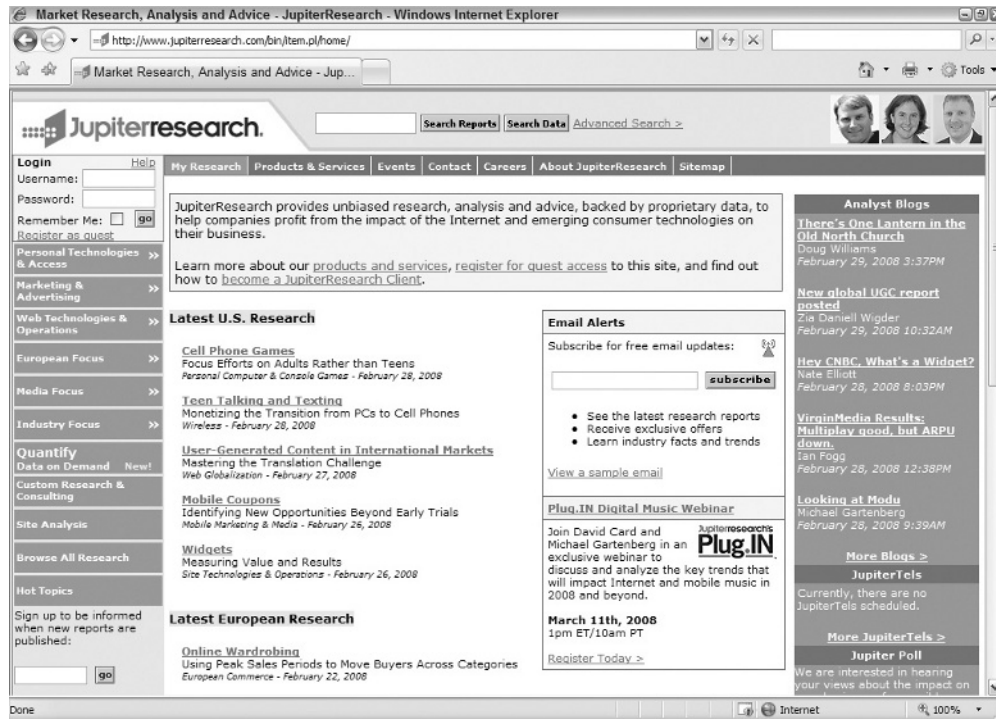
Neither the numbers nor the methodology is necessarily free. Most of these companies do issue some public reports, but they make their livings conducting custom research projects or selling the results of their own initiated studies. Often, the publicly released figures are only the tip of the iceberg. You can sometimes cut down on the cost of conducting a study by participating in an *omnibus study*, in which a group of companies each pay part of the costs, while all share in the results.

The Basics of Building Web Pages and Sites

1

FIGURE 1-11

comScore is one of the best sources of Internet facts.



Surveying site visitors

When it comes to learning about your particular site's visitors — or what they like — there's not much of a substitute for simply asking them. There are drawbacks, however, to the survey concept. The biggest problem is that people don't always respond truthfully for a variety of reasons, sometimes with the best of intentions, sometimes out of plain vanity. If you ask people whether they read tabloids or intellectual magazines, for example, the vast majority of them say that they don't read tabloids. While this has proved to be true in many different surveys, tabloids outsell intellectual magazines by millions of copies, so someone's not telling the truth. Remember, when devising your survey questions, that any time you ask people to admit something that's less than flattering to his or her self-image, you are asking those people to respond less than truthfully.

This phenomenon is widely recognized in the consumer research field, and a common (and simple) way around it was developed long ago. The trick is to ask other questions that help qualify the responses. The confirmatory questions should not be close to the sensitive one, but later, and scattered about within the survey. For example, you might ask visitors to choose their favorite

Part I

Laying the Foundation

TABLE 1-5

Internet Research Firms

Company	Web Address
comScore Networks	www.comscore.com
Forrester Research	www.forrester.com
Harris Interactive	www.harrisinteractive.com
Jupiter Communications	jupitercommunications.com
Nielsen//NetRatings	nielsenetratings.com
Nua Internet Surveys	www.nua.ie/surveys
Statistical Research, Inc.	statisticalresearch.com
WebSideStory	websidestory.com
Zandl Group	zandlgroup.com

columnists from a list. If they fail to select a single one from the intellectual magazines, you have a red flag on their previous answer — why would they read those magazines if they didn't enjoy any of their writers? In that case, you should probably exclude that respondent's answers from your tally.

When designing a form for your Web site, however, you should bear in mind that you don't want to make it too long. People — especially Web surfers — do not have endless patience for answering zillions of questions, which limits your ability to mix in confirmatory questions.

CAUTION If at all possible, resist the tremendous temptation to force your site visitors to fill out endless forms in order to access your site. You will drive away more people than you'll attract, if you don't avoid this temptation. You may think that you have the upper hand because you can withhold access, but visitors have the upper hand because they can refuse to participate.

Determining Your Approach

Before you take the first step toward actually putting your Web site together, you have to determine how you are going to approach your intended audience. You must first set your goals, decide how complex your site's technology will be, and, armed with that knowledge, set a realistic budget.

The Basics of Building Web Pages and Sites

1

Focusing on goals

Every Web site, in order to be successful, must fulfill some sort of need for its visitors. The word “need” doesn’t have to be interpreted too strictly. After all, how many people really *need* a \$400 leather jacket? Or a handcrafted miniature car? Yet, there are sites that are successfully selling such items.

Your goals must be set by one key rule: *The average person visiting your site should walk away happy.* The key to making your visitors happy — and making them want to come back again and again — is to satisfy people’s desires. Most Web site designers make the mistake of assuming that people only come for what they really have to have — or, more accurately, what the designer *thinks* people really need. That’s a fatal error to make. On the Internet, people are still people and, as in the outside world, people are driven more by emotions than by logic.

Put yourself in the shoes of your potential site visitors and take a look at what you’re offering. Does your site solve their problems (or, more accurately, their *perceived* problems)? Does it provide them with the information or products they’re looking for? Does it do it all that easily, without making your visitors work too hard to get what they want?

Deciding on complexity levels

How deeply are you going to get involved in Internet technology? When we’re talking “technology” here, we mean both the hardware and the software — not just on your end, but on your site visitors’ end, too.

Living and dying on the edge

Internet technology is constantly changing, and this means that you need to make one specific choice — whether or not to keep your Web site on the leading edge of that change. The leading edge is often sardonically referred to as “the bleeding edge,” because lots of people and companies have suffered tremendously by trying to stay up with it.

The reasons for this are twofold. First, leading-edge technology has much higher expenses. Second, no one can guarantee which of several competing approaches will end up becoming the standard. What looks as if it’s the most promising approach to accomplishing something can turn out to be unpopular in the end, regardless of its technical merits.

Our best recommendation is to hang back a little bit and not join the latest rush to the New Hot Internet Thing. But just for a little while — just long enough to be sure it’s the real thing, and not just some fad that won’t pan out in the end. As soon as you’re sure that the change is permanent and the technology is here to stay, move ahead with it. Let some other company do the bleeding, then move in fast enough to reap the rewards — if there are any to be had.

Part I**Laying the Foundation****Staying deliberately obsolete**

That said, there are good and solid reasons to deliberately ignore even a well-established Internet technology in favor of still older approaches. And those reasons are stability and demographic reach. As new standards for Web design and applications emerge, anyone on the leading edge of implementing them will quickly find that the latest Web browsers (despite anything their makers say) only partially support the standards when they are first introduced.

To further confuse the situation, the average Web surfer doesn't keep up to date with the latest Web browser. Thus, even if the newest versions of browsers finally get around to supporting the latest advance in its entirety, you'll still lose the audience share that uses old browsers if your site uses the latest ideas. If you have committed to following hot on the heels of the standards committees, your site had better be an experimental one and not meant to serve any practical purpose.

The basic question you need to consider is whether or not you want to reach the vast majority of Web surfers — those who are using Web browsers that are, at least to some degree, obsolete. If so, you should consider using the lowest level of Internet technology that still achieves your basic goals.

On the other hand, if your target audience is users of higher-level technology — you run a Web site that has nothing but tutorials on Macromedia Flash, for example — you can confidently expect that the only low-tech visitors to your site will have wandered in by accident.

Establishing a budget

How much are you willing to spend on your site? This question (more than any other one) sets the practical limitations on what you're able to do.

If you're whipping up a personal site, the odds are pretty good that all you need to invest is nothing more than the amount that you have already put into getting online in the first place. All your needs are probably thoroughly covered by the basic amount of Web space that your ISP supplies as a basic part of their service. If the amount of Web space your ISP provides is inadequate, you can doubtless find plenty more available for little or no further investment.

You need to sit down and figure out exactly how much money you're willing to commit to your project if you're going for anything beyond that rock-bottom minimum. If the money's going to come out of your own pocket, you have to take a hard look at how much you're willing to lose, because every Internet startup is a gamble. Unless you're very rich or planning on a really small operation, you may have to put a substantial percentage of your wealth into the first year of your site's operation. Are you willing to risk giving up all your disposable income to try to make your dream come true? Half of it? Ten percent?

Whether or not your site is designed to generate an income, it still costs money to run. And, while there are exceptions to the rule, the best sites tend to cost the most money. Are you passionate enough about what you want to do on the Web to divert that money from all the other things you could do with it? If the answer is "yes," then you stand a better chance of success. But if you're hesitant about committing, you may want to think twice — maybe even thrice.

Dos and Don'ts of Web Site Development

The reasons for good Web site design are obvious. You want visitors to your site to have a pleasant productive experience when they do business with you. You not only want them to like what you have to offer but also to appreciate the ease and convenience with which their online shopping is done. You want customers to come back.

Tips for good Web sites

Here are ways to develop customer-friendly Web sites:

- *Take a minimalist approach* — Customers come to your site for the content, not the clutter. Limit the distracting eye candy.
- *Make navigation easy to use and understand* — This means providing rapid access to desired content within a few clicks.
- *Make the Web site fast-loading* — Many people still use a dial-up connection to the Internet. Too much animation can slow down the rendering of a Web page.
- *Skip the sticky advertisements that follow the cursor around like a lost pup* — They interfere with the reading of content and break the user's train of thought. Many users may soon remember those sites and avoid them. There are always others.
- *Skip the counters* — Why broadcast this information? What do they add to your site? In fact, some amounts could be quite small and convince the visitor that your site is not popular.
- *Keep the information on your site current* — This is not a Web-specific issue, but it is implicit in the Information Age that information should be accurate and up to date.
- *Be grammatically correct* — There are potential customers out there who believe that if you can't spell, you can't be very good at your business, either. They will go elsewhere.
- *Do not assume your Web designers can spell or write a grammatically correct sentence* — Check your Web sites. Without being specific, there was actually an academic institution's Web site in Statesville, North Carolina (state-funded, of course), that had mindless grammatical and spelling errors.
- *Be very sparing of sound effects or music* — Not all people have the same tastes. Not all people have the same audio cards or speakers. Also, keep in mind your visitors may be browsing to your site at work and may annoy their coworkers.
- *Use thumbnails as a good aid to site content and navigation* — Make certain that all thumbnails are clickable, that they're relevant to the content, and that the images they link to load fast. Do not clutter the page with them.
- *Make contact info obvious* — Even on a home page, include contact info such as business name, address, e-mail address, fax number, and phone number — maybe even two or three, such as generic office number, technical support number, and customer support number. This saves clients a lot of time, and they do not need to guess which

Part I

Laying the Foundation

of three or more possible links actually contains the phone number they are so desperately seeking.

- *Make fonts readable* — Have high contrast with the background and, if you use them at all, do not put significant amounts of light-colored fonts against dark backgrounds.
- *Be conscious of backgrounds* — Consider white to be the background of choice. Notice the use of white backgrounds in some very high-traffic sites such as Google, Yahoo! and Drudge Report. White backgrounds make the content easier to read and enable graphics to stand out.
- *Be conscious of your user's screen resolution* — Screen resolutions can run from 640 x 480 on up. Therefore, when developing sites, be sure to make table widths flexible, not fixed.
- *Field-test the site using real, live, honest-to-goodness people who know nothing about your business* — Of course, people who designed the site or use it daily are not going to find it user-unfriendly.

The following are some sources for good Web-site design, advice, and techniques:

- www.webreference.com
- websitesforgood.com
- www.lib.berkeley.edu/TeachingLib/Guides/Internet/Evaluate.html

Top ways to have bad sites

On the other side of the coin, here are ways to develop customer-unfriendly (that is, bad) Web sites:

- Have links that don't work.
- Have messages that a Flash Macromedia advertisement is loading when you click on a company's URL and omit the Skip Intro button.
- Have a form completely clear if a customer makes any mistake anywhere in inputting information into the form.
- Have shipping charges appear on the very last screen of an online order.
- Do not specify the date, Social Security card, or phone number formats required for input.
- Have rapid, repetitive animation or video shorts, and use lots of them.
- Have short segments of loud annoying music — and loop them.
- Use lots of different colors, themes, and unrelated images and motifs on your home page and throughout the site.
- Use font colors that are practically indistinguishable from the background.
- Have your home page cluttered with countless links using different fonts and backgrounds.

The Basics of Building Web Pages and Sites

1

- Have links unrelated to the main purpose(s) of the Web site.
- Do not use a spellchecker.

All of these faux pas have occurred, countless times, on real Web sites. Several examples of bad Web sites may currently be found at such locations as:

- www.webpagesthatsuck.com
- www.worstoftheweb.com
- <http://www.bad-websites.com/>

Summary

This chapter has explained the background and development of Web browsers and reviewed their various features. This chapter also touched upon Web protocols and gave a very basic outline of a Web page using HTML tags.

The purpose of a Web browser is to fetch, decode, and display content from hypertext documents. As the content of Web sites became more sophisticated as the result of the development of more complex audio and graphics programs, the Web browser technologies become equally sophisticated to be able to fetch, decode, and display the contents rapidly, accurately, and securely. The development of one fed upon the other technologies.

Chapter 2 briefly reviews some popular Web design applications, discusses designing for the mobile Web, and makes suggestions for additional freeware/shareware graphics applications.

