7
Advertising and Embedded Content

[My web history is] mine—you can't have it. If you want to use it for something, then you have to negotiate with me. I have to agree, I have to understand what I'm getting in return.[1]
—Sir Tim Berners-Lee

Publishing information is the backbone of web content, and bloggers and webmasters frequently rely on embedded content from companies such as Google to enhance the quality of their sites. Unfortunately, embedding third-party content is the equivalent of planting a web bug in web pages,[2] alerting the source of the embedded content to a user's presence on a given site and facilitating logging, profiling, and fingerprinting. More important, the source of the third-party content can aggregate these single instances and track users as they browse the web. This notion deserves restating: The simple act of web browsing across many disparate sites has the potential to generate a continuous stream of information back to the providers of third-party content. The more popular a given third-party service is, the more sites will deploy their content, and the greater the window of visibility on users' web surfing activity becomes. In the case of advertising

networks such as Google/DoubleClick and web-analytics services such as Google Analytics, the risk is large indeed.

This chapter explores the risks associated with embedded content by focusing on Google's advertising network and Google Analytics, but it also provides an overview of other forms of embedded content that present related risks, such as embedded YouTube videos, maps, and Google's Chat Back Service.

CROSS-SITE TRACKING

As mentioned in Chapter 3, "Footprints, Fingerprints, and Connections," many web sites embed third-party content in their sites. Third-party content can take the form of legitimate images and video clips, among other forms of content, but it can also be used to track users as they surf the web. Advertisers and web-analytic services give webmasters enticing analysis tools and advertising profit, simply requiring that, in exchange, webmasters add small snippets of HTML and JavaScript to their pages. Unfortunately, such third-party content is a severe privacy and web-based information-disclosure risk because the user's web browser automatically visits these third-party servers,[3] where their visit is presumably logged and their browser tagged with cookies. More important, the larger the advertising network is, the larger the window a given company has on a user's online activity. For example, if a user visits 100 different web sites, each containing advertisements from a single advertising service, that service can observe the user as he or she visits each site. Figure 7-1 depicts cross-site tracking via an advertising network. In this figure, a user visits six distinct web sites, each hosting content from a single advertiser. In turn, the user's visits create one set of log entries on each of the six legitimate servers. However, because each visit contained an advertisement from a single advertising network, the advertiser is able to log all six visits.
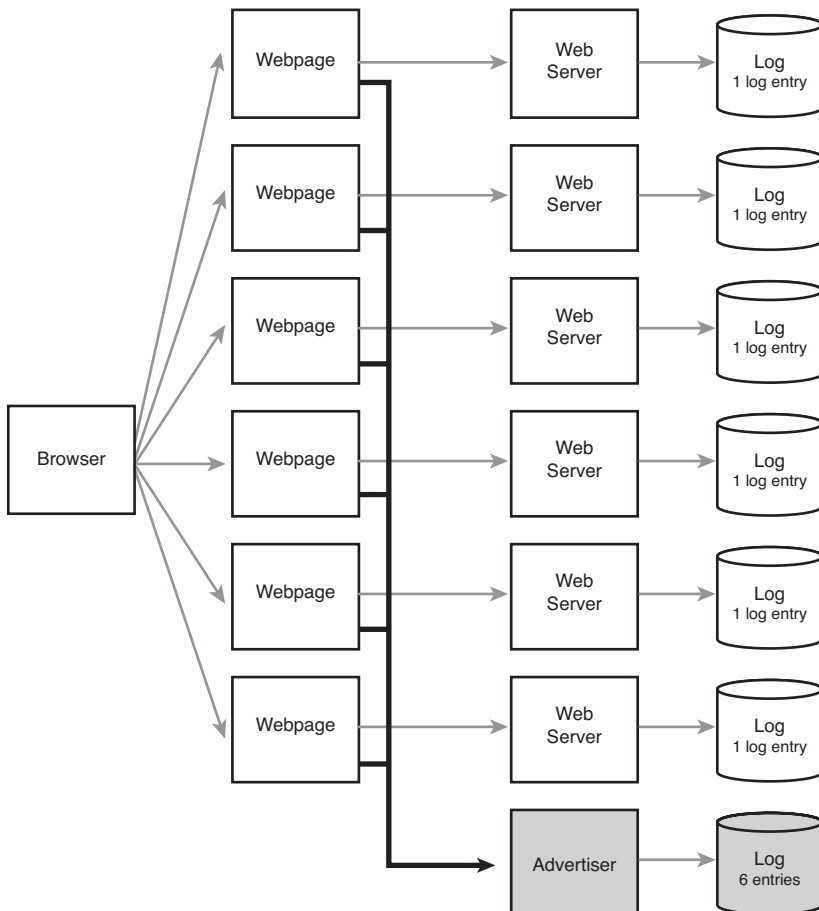
Figure 7-1 Example of cross-site tracking by an advertising network. When the user visits six distinct sites, he or she generates one set of log entries at each site. However, if each site contains advertisements from a single advertising network, the advertiser is able to record all six visits.

Let's look at a real-world example by visiting a popular web site, MSNBC (see Figure 7-2). As it turns out, the MSNBC web site is laden with third-party content.

Figure 7-2   Analyzing the MSNBC web site demonstrates that it contains a great deal of third-party content. The problem is rampant among other web sites, large and small.

In a world without third-party content, the user should simply receive content from MSNBC's domain, msnbc.msn.com. In the real world, however, the user visits 16 additional domains from 10 different companies. Two of these domains, DoubleClick and GoogleAnalytics, are owned by Google. The web browser provides the user with little assistance in detecting third-party content. The user simply sees the browser's status bar rapidly flicker as the browser contacts each new site. To provide a clearer picture, I captured the raw network activity using the Wireshark protocol-analysis tool and created Table 7-1 to detail each of the third party domains visited.

Table 7-1   Third-Party Sites Visited When Browsing the MSNBC Web Site

Domain	Notes
a365.ms.akamai.net a509.cd.akamai.net	Domain owned by Akamai.com, a mirroring service for media content
ad.3ad.doubleclick.net	Digital marketing service, acquired by Google
amch.questionmarket.com	Hosting web site where online surveys are posted
c.live.com.nsatc.net c.msn.com.nsatc.net rad.msn.com.nsatc.net	Registered to Savvis Communications, a networking and hosting provider
context3.kanoodle.com	Search-targeted sponsored links service
global.msads.net.c.footprint.net hm.sc.msn.com.c.footprint.net	Registered to Level 3 Communications, a large network provider
msnbcom.112.2o7.net	Registered to Omniture, a web analytics and online business optimization provider
prpx.service.mirror-image.net wrpx.service.mirror-image.net	Registered to Mirror Image Internet, a content delivery, streaming media, and web computing service
switch.atdmt.com view.atdmt.com	Registered to aQuantive, parent company to a family of digital marketing companies
www-google-analytics.l.google.com	Traffic measurement and interactive reporting service offered by Google

Think about it. Simply visiting a single web page from a popular news service informs *16* third-party servers of the visit, a 16-fold magnification of logging. This is not a manufactured example, but it is representative of a common practice. Embedding third-party content in web sites is ubiquitous, and so is the problem. The end result is that web surfers are frequently tracked by companies they've never even heard of. It is also worth considering that information sharing via embedded content doesn't occur only with "third parties"—sharing can also occur between ostensibly separate entities that are actually owned by the same parent company. For example, the A9 search engine (an Amazon.com company) inserts search term–related Amazon book advertisements adjacent to search results. These advertisements allow Amazon to track what A9 users search for, click through, and possibly buy online. If the user does make a purchase on Amazon.com, Amazon knows that user's real-world identity, including billing and shipping information. In the case of A9, Amazon makes clear on the A9 site that A9 is an Amazon.com company, but the important idea is that corporate ownership—and, hence, implicit information sharing—might not be obvious as users browse the web.

ADVERTISING

A famous New Yorker cartoon from 1993 showed two dogs at a computer, with one saying to the other, "On the Internet, nobody knows you're a dog." That may no longer be true.[4]
—Louise Story

Advertising is the fuel behind virtually all free online tools. Advertising is also the means for tracking your web surfing across the Internet. Anytime you visit sites that serve advertisements from a common advertising network, your activities can be logged. These logs can then be used to create precise profiles, facilitating tailored advertising. Importantly, log analysis isn't a static process. For example, advertising companies are actively developing technologies to anticipate people's next steps.[5] Based on the popularity of the largest web companies, the amount of information they can collect is staggering (see Figure 7-3).
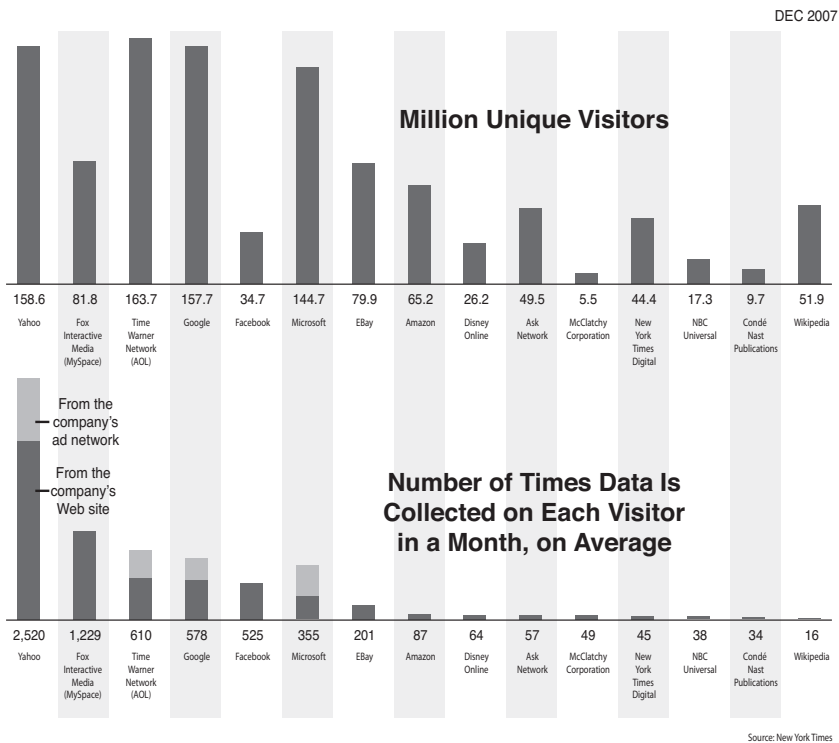
DEC 2007

Million Unique Visitors

158.6	81.8	163.7	157.7	34.7	144.7	79.9	65.2	26.2	49.5	5.5	44.4	17.3	9.7	51.9
Yahoo	Fox Interactive Media (MySpace)	Time Warner Network (AOL)	Google	Facebook	Microsoft	EBay	Amazon	Disney Online	Ask Network	McClatchy Corporation	New York Times Digital	NBC Universal	Condé Nast Publications	Wikipedia

From the company's ad network

From the company's Web site

Number of Times Data Is Collected on Each Visitor in a Month, on Average

2,520	1,229	610	578	525	355	201	87	64	57	49	45	38	34	16
Yahoo	Fox Interactive Media (MySpace)	Time Warner Network (AOL)	Google	Facebook	Microsoft	EBay	Amazon	Disney Online	Ask Network	McClatchy Corporation	New York Times Digital	NBC Universal	Condé Nast Publications	Wikipedia

Source: New York Times

Figure 7-3 *New York Times* analysis of the data collection conducted by some of the largest online companies.

Web advertisements are big business, and some of the largest services are offerings by Google, including AdSense, AdWords, and DoubleClick. However, advertisements are appearing in many other forms. Microsoft is quietly offering an ad-funded version of its Works office suite.[6] Pudding Media, a San Jose–based startup, is offering advertisement-supported phone service.[7]

Google understands the growth potential of advertising. In 2007, the company announced that it would pay $3.1 billion to acquire DoubleClick, a leading online advertiser, whose current estimated revenues at the time were $150 million.[8] A year ago, Yahoo! made a similar, smaller-scale move by acquiring the online global ad network BlueLithium for approximately $300 million.[9] At the time of this writing Yahoo! was reportedly considering an agreement to carry search advertisements from Google, amid a potential hostile takeover attempt by Microsoft.[10] Although it is impossible to know for sure at this time, such an allegiance carries the very real potential of allowing Google to gather search terms, issue cookies, and conduct other activities based on Yahoo!'s extremely large user base. Make no mistake, a key component of such acquisitions and alliances is access to user data and the power it provides.

NOTE

Online advertising is a growth industry that is eating into traditional media markets. For example, the media-analysis company Simmons found that Internet video advertisements are 47% more engaging—and, hence, more effective—than traditional television advertisements.[11]

Chapter 3 should have convinced you that users scatter significant, and often personally identifiable, information behind as they surf the web. When *New York Times* reporter Louise Story decided to determine how personally identifiable the information maintained by large web companies is, she asked four large online companies a single question: "Can you show me an advertisement with my name in it?" Story provided the following summary of the responses.[12]

Microsoft says it could use only a person's first name. AOL and Yahoo! could use a full name, but only on their sites, not the other sites on which they place ads. Google isn't sure—it probably could, but it doesn't know the names of most of its users.

Although these results are telling in their own right, keep in mind that these are official responses provided to a *New York Times* reporter. In other words, these are legal

opinions on the subject, as opposed to technical capability. What the four companies have the *capability* to do is an entirely different matter. Advertising networks are the net that allows large online companies to gather this precise information and use it for user profiling, data mining, and targeted advertising. Because of this capability to aggregate and analyze user information, advertising campaigns can follow users as they switch from independent sites. You might have encountered this technique when shopping online. For example, if you were searching for flat panel monitors on site X, when you hopped over to visit site Y, low and behold, there were advertisements for flat panel monitors.

Microsoft is openly touting an "Engagement Mapping" approach that seeks to move beyond the "outdated 'last ad clicked'" model by understanding "how each ad exposure—whether display, rich media or search, seen multiple times on multiple sites and across many channels—influenced an eventual purchase."[13] The article also states that Microsoft intends to use data on user behavior *before* clicking an ad to be able to say that an unclicked ad still made an impression. If the article is to be believed, Microsoft intends to collate that information with search queries and sites visited within the period of a day or a couple of days. Two questions immediately arise. Where does that other data come from (and is it a mere coincidence that Microsoft is seeking to acquire Yahoo! and other search companies)? Now that we are talking about keeping tabs of long-term user behavior, and hypothesizing why a user did this or that, what other kinds of data mining will this floodgate open? They are essentially saying, "We are going to dedicate a lot of resources to watching where you go and what you see on the web." The fact that Microsoft Windows is the most popular operating system on Earth[14] just magnifies the concern. The law surrounding online advertising and the collection of user data is still immature, so advertisers have a very wide lane in which to operate. For example, a U.S. Federal Court ruled that ads displayed by search engines are protected as free speech when deciding what advertisements to display.[15]

ADSENSE

Google AdSense[16], sometimes called Google Syndication, is an advertising service Google provides that allows webmasters to earn advertising revenue by hosting AdSense ads (see Figure 7-4). These revenues aren't trivial, commonly ranging from a few hundred dollars a month to $50,000 or more per year, making the service extremely popular.[17] AdSense advertisements are context-sensitive ads served by Google based on the hosting site's content. Unfortunately, merely visiting a web site hosting these advertisements informs Google of the user's IP address and gives Google the opportunity to log the user's visit and tag the user's browser with a cookie.
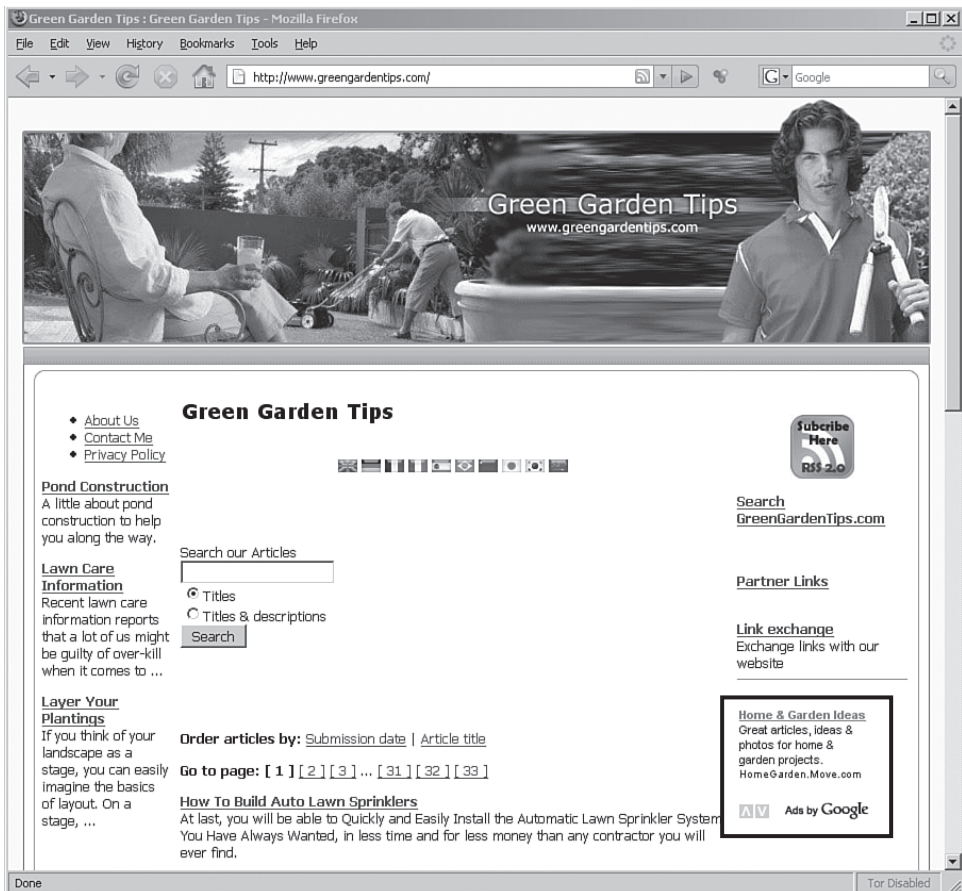
Figure 7-4 Screenshot of Google AdSense. Notice the (debatably) unobtrusive advertisement in the bottom-right corner. Embedded advertisements such as these alert the advertising network of your presence on a site.

AdSense isn't limited to textual ads on traditional web pages. Google is experimenting with AdSense for other forms of content, including RSS Feeds,[18] web site search boxes,[19] mobile content,[20] video, and Cost Per Action AdSense.[21, 22] Nor are AdSense and similar services limited to minor sites. Major online retailers also participate. For example, eBay signed deals to run ads from Google and Yahoo!.[23] Figure 7-5 shows an example with

eBay and Yahoo!. When searching for an item on eBay, Yahoo! servers provide contextual advertisements, leaving open the likelihood of Yahoo!'s logging of eBay visitors.[24]
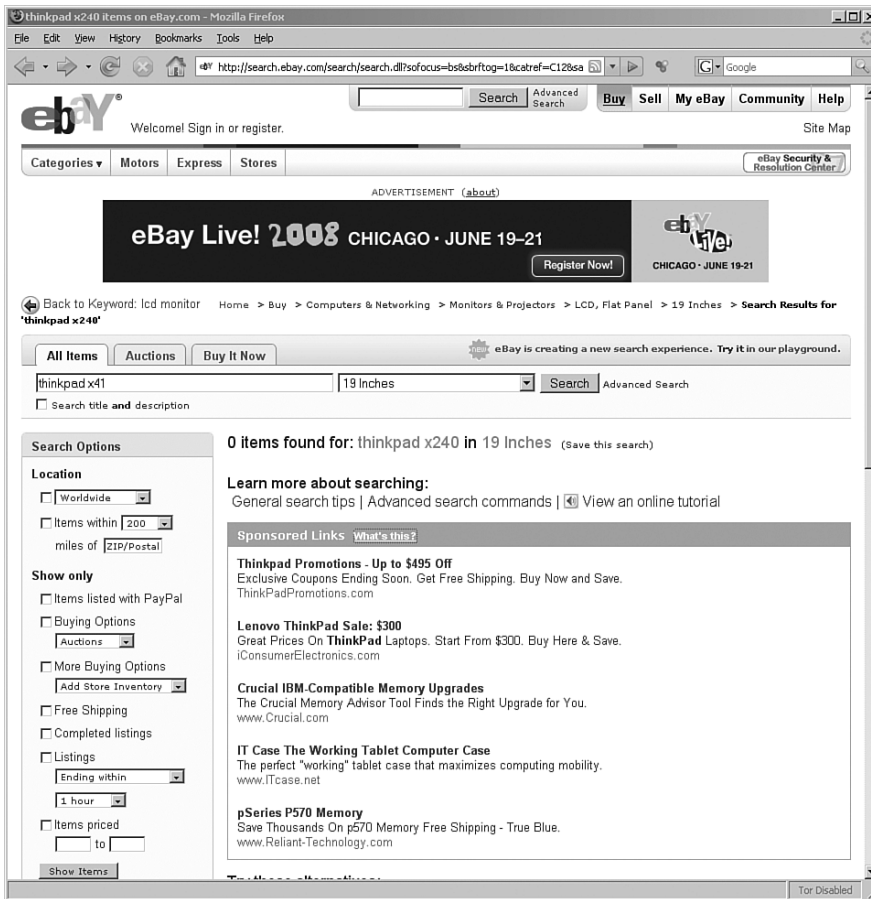
Figure 7-5   Screenshot of Yahoo! advertisements embedded in an eBay web page. Because these ads are pulled directly from Yahoo! servers, user information such as cookies and IP addresses is disclosed directly to Yahoo!.

> **WARNING**
>
> It is also important to consider the identity-disclosure risks from the perspective of the webmaster, who must use the user's registered Google account to log on to the service and use the web interface to administer his or her accounts.[25]

The future of AdSense is difficult to determine. Some analysts believe that Google is acquiring sites that will provide traffic itself instead of paying adverting fees to third-party sites.[26]

ADWORDS

AdWords is a fundamental part of Google's business model. According to the BBC, every time a user conducts a search on Google, the company makes 12¢ in revenue.[27] When you consider that Google receives more than 60 billion searches per year in the United States alone, you can see that the program generates huge profit. Google believes that AdWords "is the largest program of its kind."[28] Using AdWords, would-be advertisers bid on search terms that are displayed as part of the user's search results; the better the placement, the higher the cost. AdWords are relatively unobtrusive, but quite effective, advertisements (see the right side of Figure 7-6).[29]

AdWords poses both information-disclosure risk and other security risks. Attackers have used AdWords and similar services to misdirect users to malicious sites; see the "Malicious Ad Serving" section later in the chapter.[30] However, Google's AdWord Partners are a significant information-disclosure risk because searches from these sites can be sent to Google. According to Google those who have already joined their "growing advertising network" include AOL, Ask.com, Ask Jeeves, AT&T Worldnet, CompuServe, EarthLink, Excite, and Netscape.[31] Even third-party search engines that delete their logs locally are still at risk. Take, for example, Ask.com, who took an industry-leading position by offering AskEraser, a function that deletes search activity from Ask.com servers.[32] However, Google delivers the bulk of Ask's advertisements, so user information, including the search query and IP address, are passed back to Google each time a page is served to a visitor.[33]
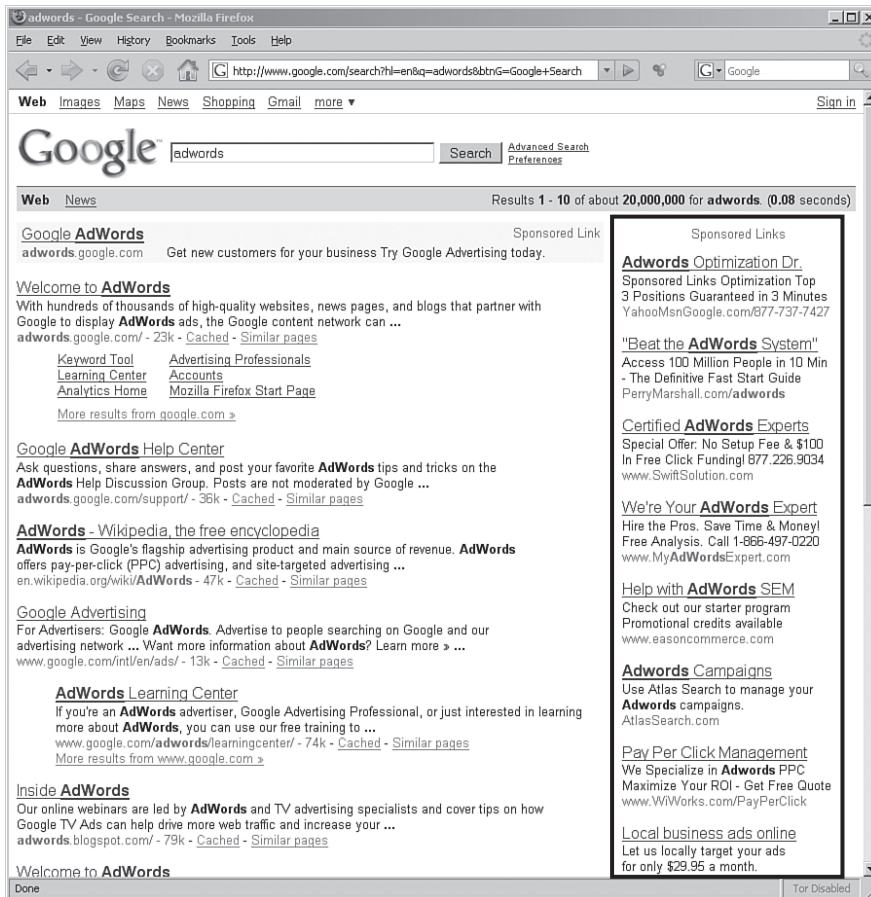
Figure 7-6 Screenshot of Google AdWords. Notice the advertisements on the right side of the image.

GOOGLE DOUBLECLICK

DoubleClick is a major online advertising service, long criticized for using cookies and IP addresses to track users as they surf the web.[34] DoubleClick is an extremely popular advertising service and counts a large number of Fortune 500 companies as clients. In 2007, Google announced a definitive agreement with DoubleClick for $3.1 billion in cash to acquire the company. The acquisition drew the attention of the U.S. Federal Trade Commission and European Regulators, who investigated antitrust and privacy

implications but eventually acquiesced.[35, 36, 37, 38] Google closed the acquisition of DoubleClick shortly thereafter.[39]

> **WARNING**
>
> DoubleClick allows users to opt out of its ad-serving and search products by issuing a special cookie. This is an elegant solution from the perspective of DoubleClick because users concerned enough to opt out might delete this cookie by accident as they remove other traditional tracking cookies. One possible solution is a browser plug-in that allows easy-to-use and fine-grained cookie control, although this approach still assumes that companies employing opt-out cookies will honor the request and not log user activity anyway.[40, 41]

The implications of a combined Google–DoubleClick dreadnaught are significant. Google excels in search advertising (AdWords) and simple textual advertisements (AdSense). On the other hand, DoubleClick excels in "display advertising," such as flashy banner ads and video advertisements, which reach between 80% and 85% of the web population.[42, 43] The end result is a broad net that permits Google to track a user's web searches and web site visits, with the potential to impact the privacy interests of more than 1.1 billion Internet users worldwide.[44] This acquisition underscores the fact that mergers and acquisitions are about data, including both existing data stockpiles and access to continued data streams.

> **NOTE**
>
> Google's acquisition of DoubleClick places significant competitive pressure on Microsoft.[45] As a response, Microsoft actively sought to acquire Yahoo! to increase its advertising reach.[46]

ADVERTISING RISKS

Advertising poses more risks than those already discussed regarding AdWords, AdSense, and DoubleClick. Attackers can exploit advertising networks to compromise end-user machines, unethical interface techniques can trick users into disclosing sensitive information, and historically unbiased network providers can insert advertisements as the web pages make their way to the user's browser.

MALICIOUS AD SERVING

Advertising networks are more than just information-disclosure risks. They also serve as a malware attack vector. Advertising services pay web site owners for publishing advertisements on their web sites. A very common technique is the banner ad we've all seen at the top of web pages. Such ads usually take the form of animated GIF files, but they now include many image and video formats. Individuals and organizations that want to advertise using such a service create a media file and pay an advertiser a fee, and the advertiser serves the image to thousands of visitors of sites that belong to its advertising network. The risks here are twofold. Attackers have created misleading advertisements as a means to draw traffic to a malware serving or other malicious web site.[47] The users' trust of the advertisement company and the hosting web site increases their trust of the advertisements, leaving web surfers more vulnerable to such an attack. Virus writers have used the Google Adwords service to serve text ads that appeared to link to legitimate destination sites, but silently infected vulnerable web surfers by routing users through an intermediate, malicious site. Attackers also have used a vulnerability in Internet Explorer to compromise visitors as they passed through the intermediate site, before ultimately arriving at the legitimate site.[48]

The ads themselves have also been used to attack the web user directly. Malformed graphical images are one common technique. For example, a banner advertisement displayed on MySpace served spyware to more than one million visitors. In this case, attackers exploited a flaw in the way Windows processed Windows Meta File (WMF) images to install a Trojan horse.[49] Because the attack occurs when the browser displays the image, the user needn't click the advertisement to be infected. Another attack, served by DoubleClick, used rich media advertisements created in Adobe Flash to exploit a similar vulnerability, with the malicious advertisements appearing on extremely popular sites, including those of *The Economist* and Major League Baseball.[50] Rich media advertisements are highly interactive and are becoming increasingly popular. Even the seemingly simple task of securing browsers against malicious images is proving difficult, and complex, rich media tools such as Flash are proving to be an even greater challenge.[51] Other complex environments, such as ads embedded in Adobe PDF files, are now being explored as ways to reach potential customers, and I expect that similar issues will arise.[52]

NOTE

Google has taken an active role in countering malicious advertisements and web sites. For example, Google warns users of potentially malicious sites by clearly labeling suspect sites in their search result listings. As another example, Google's I'm Feeling Lucky search button no longer automatically redirects a user to a suspicious site; instead, the user is presented with a list of search results instead.[53]

MALICIOUS INTERFACES

Beyond malicious ad serving, advertisers employ another concerning strategy that I call *malicious interfaces.* In the idealistic world of interface design theory, interface designers always operate in the best interests of their users. Designers carefully study user tasks and painstakingly craft interfaces and applications to help users accomplish them. However, in the world of online advertising, the exact opposite is true. Designers frequently violate design best practices to coerce or mislead users into viewing advertising. Examples abound on the web: fake hyperlinks that pop up advertisements, giant advertisements that cover the text of articles, distracting advertising videos that begin playing the moment a page is viewed, banner ads with fake buttons that appear to be a part of the interface, advertisements embedded in video clips … the list goes on. Malicious interface designers are creative; new "innovations" are coming out regularly. The only constraining factor appears to be the tolerance of the user. The invasiveness of advertisements is getting worse. You may have heard the term "banner blindness." Users quickly learned that banner ads are of little value and ignore the advertisements, to the point that they barely perceive banner ads anymore. This defense mechanism has forced advertisers to become more aggressive in capturing user attention.[54, 55]Although, I don't claim that Google employs malicious interface design, the trend is concerning; it seems that malicious interface designers are carefully seeking the sweet spot between making advertising profit and annoying the user so much that they abandon a given site altogether.

WARNING

In many cases the link displayed by an embedded advertisement is not the actual link. Nor will hovering over the link display the destination URL in the user's browser status bar. The actual link goes first to the ad server so the click can be logged. The user's browser is then redirected to the page chosen by the advertiser.

HOSTILE NETWORKS

As carriers of key components of network infrastructure, ISPs and web hosting services are flexing their muscle to place advertisements in front of users. For example, domain registrar and web hosting provider Network Solutions hijacked customers' unused sub-domains to resort to ad-laden "parking" pages.[56] As another example, bloggers Lauren Weinstein and Sarah Lai Stirland reported Canadian ISP Rogers modification of web pages en route (see Figure 7-7).
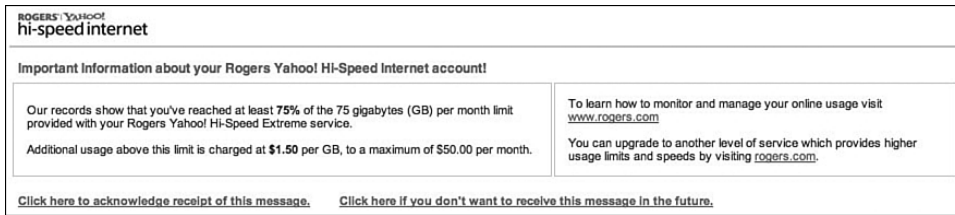
Figure 7-7   Screenshot of an ISP altering a Google web page

Inserting advertisements into web pages as they transit an ISP is potentially very big business. ISPs already have access to a tremendous amount of personal information about users' online activity. This data is a veritable gold mine if used to target online advertising. Some ISPs are reportedly selling significant amounts of user data to online marketers.[57] In the United Kingdom, three major ISPs have announced plans to use user clickstream data to insert relevant advertisements as they surf, through a new startup called Phorm.[58] ISP data contains some of the most sensitive information disclosures made by online users. If this advertising technique becomes widespread, virtually every web surfer's activities will be passed on to advertisers in some form. Fighting the issue will be difficult, and users might find that they are faced with little alternative than to accept this new status quo.

AFFILIATE SERVICES

Affiliate advertising isn't just confined to Google via its AdSense and DoubleClick programs. It is a popular marketing practice in which, in its common form, web authors embed advertisements that contain unique tracking data to identify the correct affiliate to compensate. Such advertisements can be static—that is, the advertisement exists entirely on the server of the web author. In this case, the user must click the advertisement before the advertiser is aware of the user. However, many affiliate services provide dynamic content that is pulled directly from the online company without any action by the user, immediately linking the user to the visited web site (see Figure 7-8 for an example). In addition, the online company might tag the user with a cookie or retrieve an existing cookie, opening the possibility of identifying the user by name, billing address, and shipping address.

Figure 7-8 Example of Amazon affiliate network dynamic advertisements. Some affiliate networks, such as Amazon, encourage web authors to embed these advertisements into their web content, allowing the user to be logged, tagged with a cookie, and perhaps identified by name, merely by visiting the web page containing the advertisement.[59]

FACEBOOK BEACON

Facebook's Beacon illustrates both the detailed insight that large online companies have into the activities of their user populations and the lengths some companies will go to increase profit. Facebook's Beacon service allows Facebook users to share their purchases from affiliated online companies, such as books, movies, and gifts, with their Facebook

221

friends.[60] However, when the service was first offered, participation was "opt out." The end result is that many users were infuriated and the civil action group MoveOn.org initiated a very effective online petition that rapidly gained more than 50,000 supporters.[61] MoveOn even blamed Facebook for "ruining Christmas" because the Beacon advertising system allowed users to see holiday present purchases made by friends and family.[62]

Shortly thereafter, Facebook changed the service to "opt in" by requiring explicit user permission before publishing purchases to the user's Facebook friends.

OTHER CROSS-SITE RISKS

The preceding sections illustrated the risks associated with AdWords, AdSense, and DoubleClick, but cross-site information-disclosure risks do not end with advertising. The web functions thanks to hyperlinks and embedded content. Through these vectors, Google and other large companies can gather tremendous amounts of user information and help link clusters of information to individual users, companies, and other organizations. The following examples share one common characteristic: Each relies upon third-party web masters to embed tracking (or trackable) content into their web sites. Most web masters would not add such content arbitrarily; instead, they are enticed by at least a nominal incentive for cooperation.

> **WARNING**
>
> Embedding Google or other third-party content into a web page isn't the only concern. Posting your own content on a Google-hosted service, such as Blogger, YouTube, or Orkut, allows Google to track all users to the site. In addition, you become dependent on Google and its ideas about censorship, privacy, and quality of service.

GOOGLE ANALYTICS

Google Analytics is a free tool for webmasters that provides a powerful and intuitive interface for analyzing web log data (see Figure 7-9).[63] Google Analytics is part of a class of applications that provide statistical and graphical analyses of web visitor activity based on web server log data and (optionally) on data gained via cookies placed on users' computers, web bugs, and JavaScript code. Such tools display site visitor reports (for example, geographic locations of visitors, most active visitors, and browsers used), page view

reports (for example, entry/exit pages, most popular time of day, and number of requests for each page), server reports (for example, amount of bandwidth consumed and which files were requested), and referer reports (for example, search queries and referring URLs). Other popular web analytics software includes Webalizer (www.mrunix.net/webalizer/) and WebTrends (www.webtrends.com/).
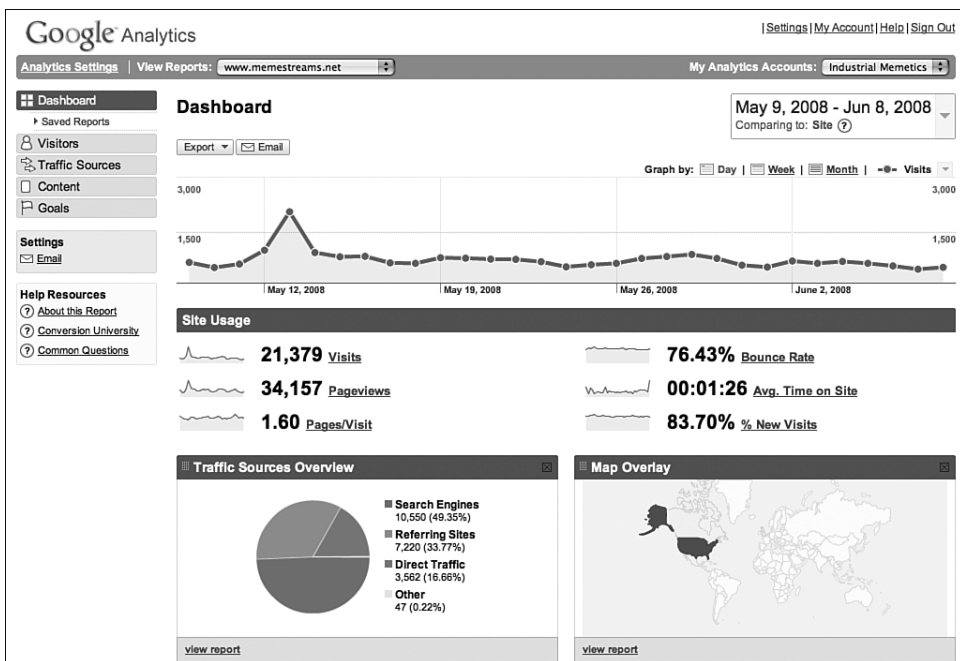
Figure 7-9 Google Analytics gives Google the capability to track users as they visit any Google Analytics member site.

Google Analytics is easy to install. Webmasters need only paste code similar to the following into web pages they want the service to track.

```
<script src="http://www.google-analytics.com/urchin.js"
type="text/javascript">
</script>

<script type="text/javascript">
_uacct = "UA-994065-1";
urchinTracker();
</script>
```

223

This code is straightforward JavaScript. It serves as a hook in each web page to contact Google whenever a page is loaded and download a JavaScript file called urchin.js. The _uacct variable stories a unique tracking code assigned to the webmaster. The script then launches the urchinTracker() function in the newly downloaded urchin.js file. Unfortunately, the code within urchin.js is far more complex and apparently obfuscated.[64, 65] The following is a short snippet:[66]

```
function urchinTracker(page) {
 if (_udl.protocol=="file:") return;
 if (_uff && (!page || page=="")) return;
 var a,b,c,xx,v,z,k,x="",s="",f=0;
 var nx=" expires="+_uNx()+";";
 var dc=_ubd.cookie;
 _udh=_uDomain();
 if (!_uVG()) return;
 _uu=Math.round(Math.random()*2147483647);
 _udt=new Date();
 _ust=Math.round(_udt.getTime()/1000);
 a=dc.indexOf("__utma="+_udh);
 b=dc.indexOf("__utmb="+_udh);
 c=dc.indexOf("__utmc="+_udh);
 if (_udn && _udn!="") { _udo=" domain="+_udn+";"; }
 if (_utimeout && _utimeout!="") {
 x=new Date(_udt.getTime()+(_utimeout*1000));
 x=" expires="+x.toGMTString()+";";
 }
```

Although many webmasters sing the praise of Google Analytics, the tool also poses a significant privacy concern for web surfers. Each time they visit a web page that contains the request to download urchin.js, the user's web browser contacts a Google server and downloads and then executes the script, leaving behind all the typical web-browsing footprints described in Chapter 3. The urchin.js script presumably discloses additional information, but Google does not provide specific details. The primary risk of Google Analytics is that it gives Google the capability to track users as they browse from web site to web site, including the use of cookies.[67] There is no official count of the number of participating web sites, but several years ago, an analyst estimated the number to be about 237,000.[68] The number now is presumably far greater. Some of the most popular sites on the web employ Google Analytics, such as Slashdot.org, which downloads Google's newer ga.js script, an equally difficult script to interpret. You can see a snippet in Figure 7-10.

```
var _gat=new Object({c:"length",p:"cookie",b:undefined,bb:function(d,a){this.wb=d;this.Hb=a},o:"__utma=",
O;h--){o=d.charCodeAt(h);a=(a<<6&268435455)+o+(o<<14);c=a&266338304;a=c!=O?a^c>>21:a}}return a},B:function
c(d)):escape(d)},z:function(d,a){var c=decodeURIComponent,h;d=d.split("+").join(" ");if(c instanceof Funct
v:function(d,a){return d.indexOf(a)},D:function(d,a){c=_gat.b==c?d[_gat.c]:c;return d.substring(a,c)},m
c("alltheweb","q"),c("gigablast","q"),c("voila","rdata"),c("virgilio","qs"),c("live","q"),c("baidu","wd").
"/";d.ha=100;d.Da="/__utm.gif";d.ta=1;d.ua=1;d.F="|";d.sa=1;d.qa=1;d.nb=1;d.f="auto";d.C=1;d.Ga=1000;d.Mc
i[s][O]+b,e))}f.Eb=function(){return n.b==g||g==f.t()};f.Ba=function(){return m?m:"-"};f.Qb=function(k){m
for(var b=O;b<j[y];b++)if(b<4&&!n.Ea(j[b]))j[b]="-"};f.tc=function(){return p};f.Ic=function(k){p=k};f.ic
b){var e=f.U,i=B.1,s;f.Ha(k);B.1=b;for(s=O;s<e[y];s++)if(!t(e[s][1]())())e[s][3]();B.1=i};f.Yb=function(){l
new Image(1,1);u.src=h.Da+r;u.onload=function(){j()})}if(1==B||2==B){var x=new Image(1,1);x.src=("https:"=
m)p.g(j.ca,new p.h.ab(r,d,a,c,h,o));else{m.Xb=r;m.Oa=d;m.K=a;m.qb=c;m.Jb=h;m.Kb=o)};_gat.h.$.prototype.Bb
n){n=new f.h.$(d,a,c,h,o,j,m,r);f.g(p.la,n)}else{n.mb=a;n.Wb=c;n.Vb=h;n.Sb=o;n.sb=j;n.Ub=m;n.vb=r}return
new ActiveXObject(n+".6");p="WIN 6,O,21,O";f.AllowScriptAccess="always";p=f.GetVariable(t)}catch(B){}if(!
a.n&&a.n.javaEnabled()?1:O;a.yb=o?j():c;a.rb=h.d(a.a.characterSet?a.a.characterSet:(a.a.charset?a.a.chars
v[z];if(f(q,t(l.wb))){x=y(x,"?").join("&");if(f(x,"&"+l.Hb+"="))(u=y(x,"&"+l.Hb+"="))[1];if(f(u,"&"))u=y(u
q=y(q,"/")[O]}if(O==m.v(q,"www."))q=m.D(q,4);return new m.k.q(p,q,p,"(referral)","referral",p,u)};j.kc=fu
"";x=j.kc(j.a.location);if(j.r.H&&q.Eb()){z=q.Ca();if(!r(z)&&!f(z,";"))(q.Ra();return"")}z=n(e,m.X+i,";")
F?1:F;q.Rb([i,j.ja,F,k,l.ka)].join(".");q.Ra();return"&utmcn=1"}else return"&utmcr=1"}};_gat.k.q=functi
a(d.ra))};_gat.k.q.prototype.zb=function(d){var a=this,c=_gat,h=function(o){return c.z(c.B.D(o,"|"))};a.u
l[j[b]];if(a.b!=e){if(k)v+=j[b];v+=x(e);k=false}else k=true}return v}function x(l){var v=[],k,b;for(b=O;b<
[l.M()],k;for(k in c)if(a.b!=c[k]&&!l.yc(k))a.g(v,k.toString()+u(c[k]));return v.join("")};d._setKey=funct
o)}};_gat.cc=function(d,a){var c=this;c.Wc=a;c.Dc=d;c._trackEvent=function(h,o,j){return a._trackEvent(c.
i){if(o(b)||o(e)||o(i))return"-";var s=r(b,c.o+a.e,e),w;if(!o(s)){w=f(s,".");w[5]=w[5]?w[5]*1+1:1;w[3]=w[
""==g.f||"none"==g.f){g.f="";return 1}q(){if(g.nb)return c.t(g.f);else return 1};a.lc=function(b,e){if(o(
return b};a.Oc=function(b){if(a.P()){var e="";if(a.j!=h&&a.j.M().length>O)e+="&utme="+c.d(a.j.M());e+=a.I
a.a.createElement("script");b.type="text/javascript";b.id="_gasojs";b.src="https://www.google.com/analyti
o(i.za())){E=x(G,"&",e);a.L=true}else{C=f(i.I(),".");s=C[O]}else if(F){if(!I||!J){E=x(b,"",e);a.L=true}el
function(){var b;if(!B){a.zc();a.e=a.rc();a.s=new c.Y(a.a,g)}if(z())a.xc();if(!B){if(z()){a.va=a.lc(a.a.r
r(b,"gaso=","&"):r(a.a[c.p],c.Sa,"");if(e[t]>=10){a.A=e;if(a.V.addEventListener)a.V.addEventListener("lo
if(g.J&&g.J[t]>O)a.Jc();a.Jc(b,a.L=false)};a._trackTrans=function(){var b=a.e,e=[],i,s,w,A;a._initData();
l(w[e]);i=f(w[e],g.F);for(s=O;s<i[t];s++)i[s]=l(i[s]);if("T"==i[O])a._addTrans(i[1],i[2],i[3],i[4],i[5],i
a.G,a.a,a.e))};a._link=function(b,e){if(g.H&&b){a._initData();a=a[n].href=a._getLinkerUrl(b,e)}};a._linkB
function(){a._initData();return new c.Z};a._sendXEvent=function(b){var e="";a._initData();if(a.P()){e+="&
false;return w};a._trackOutboundUrl=function(b){a._initData();if(a.P()){var e=new c.Z;e._setKey(6,1,b);y.
b)};a._clearIgnoredRef=function(){g.ga=[]};a.Tc=function(){return 1}{return g.da};a._setAllowHash=function(b){g.nb=
function(){return g.ta};a._setLocalGifPath=function(b){g.Da=b};a._getLocalGifPath=function(){return g.Da}
a._setCampContentKey=function(b){g.db=b};a._setCampIdKey=function(b){g.eb=b};a._setCampMediumKey=function
```

Figure 7-10 Screenshot of the Google Analytics script `ga.js` downloaded from Google when users visit Slashdot.org

The risk of being tracked across 250,000 or more web sites is concerning enough, but the true risk of Google Analytics is that the user data can be combined with web sites participating in Google's AdSense and AdWords programs, enabling the company to track users across a broad swath of the most popular portions of the web. Users see only a brief flicker in their browser's status bar as their browser contacts Google's servers. The potential of "free" web-analytics software is not lost on Google's competitors; both Yahoo! and Microsoft recently released free web-analytics tools.[69, 70]

CHAT BACK

Google's Chatback service enables web authors to embed a status indicator, a "badge," directly into their web pages. When the page is loaded, the badge (see Figure 7-11) indicates whether the user is available for communication via Google Talk. Merely visiting the page causes the user's browser to pull the Chatback badge from Google's servers, leaving behind footprints in their logs. Clicking the link can start an online conversation,

leaving open the eavesdropping and logging risks discussed in Chapter 5, "Communications." Although this is a text-based service, similar risks exist via VoIP "call-me buttons" offered by companies such as Jajah, Jangle, Jaxtr, Tringme, and Grand Central.[71, 72]

Figure 7-11   Sample Google Chatback badge. Web authors place small snippets of Google-provided code in their web pages, and visitors to the page can see whether the author is available to chat via Google Talk.

YOU TUBE VIDEOS

Embedding YouTube videos is an extremely popular practice by web authors (see Figure 7-12). When doing so, authors place code similar to the following in their web pages.[73]

```
<object height="350" width="425">
<param name="movie" value="http://www.youtube.com/v/KJukKpQDVLQ">
<param name="wmode" value="transparent">
<embed src="http://www.youtube.com/v/KJukKpQDVLQ"
 type="application/x-shockwave-flash" wmode="transparent"
 height="350" width="425">
</embed>
</object>
```

Notice that the code embeds a movie object pulled from Google's servers. Again, users need only visit a page containing an embedded YouTube video to leave themselves open to tracking by Google, even if the page is run by a third party and there are no DoubleClick or AdSense advertisements.
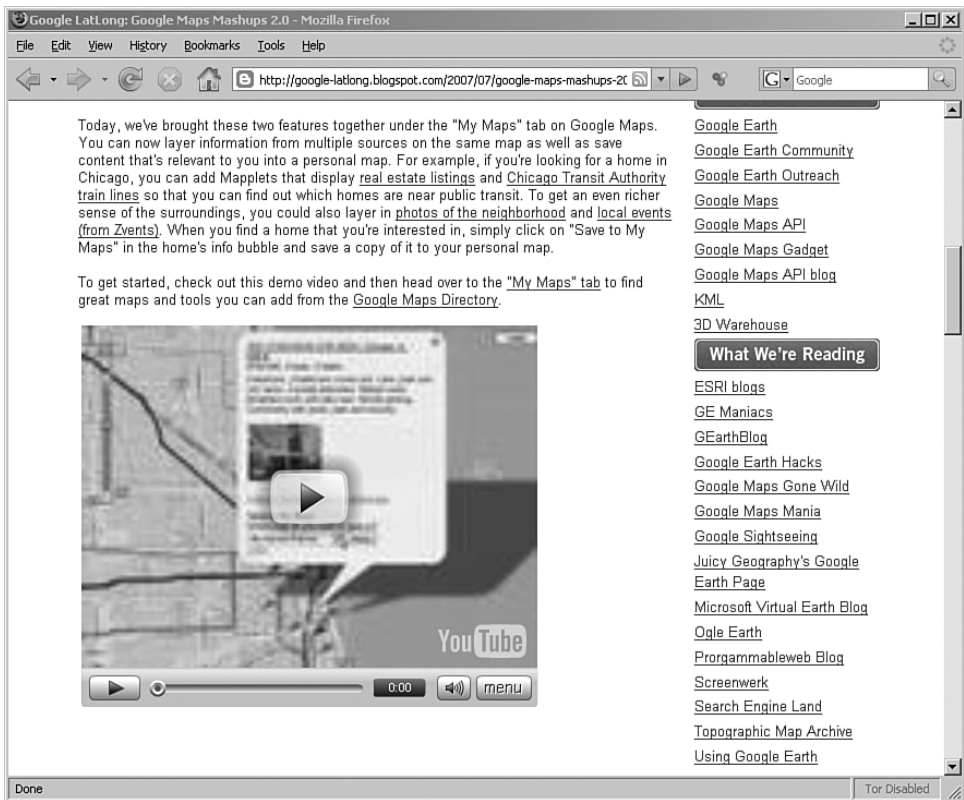
Figure 7-12 Example of a YouTube video embedded in a web page. When the image is merely displayed in the user's browser, that user can be immediately logged by YouTube.

SEARCH ON YOUR WEB PAGE

Another common practice is for web authors to include a Google Search box on their site (see Figure 7-13). Although some visitors find this useful, it also facilitates the disclosure of search queries, as well as the user's IP address and the site he or she is visiting, to

Google. In some implementations, the disclosure takes places only when the user clicks Submit, as in the following code:[74]

```
<form method="get" action="http://www.google.com/search">

<input type="text" name="q" size="31"
 maxlength="255" value="" />
<input type="submit" value="Google Search" />
<input type="radio" name="sitesearch" value="" />
 The Web
<input type="radio" name="sitesearch"
 value="askdavetaylor.com" checked /> Ask Dave Taylor<br />

</form>
```

However, note the Google logo in the image. If the webmaster includes the logo on the page, he or she can choose to download the image directly from Google; this immediately informs Google when someone visits a given site.

Google also offers AdSense for search, which helps webmasters earn revenue by creating a custom search engine for a site.[75] Along with customized search results, users see targeted advertisements.
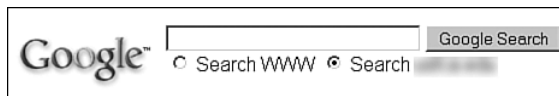

Figure 7-13 Many web sites include a Google search field, which encourages users to disclose search terms and the site they are visiting.

FRIEND CONNECT

Friend Connect is a new service offered by Google that enables web authors to add social networking facets to their sites by embedding small snippets of code. Google Friends Connect "offers a core set of social gadgets such as member management, message board, reviews, and picture sharing."[76] Figure 7-14 shows a sample site provided by Google and illustrates several concerns with Friends Connect. Visitors to the site are offered the opportunity to sign in using their existing credentials, which uniquely identifies them to Google or one of several other participating services, including Yahoo! and AOL. The sites' members, photos, and comments can be disclosed. In addition, because

this site includes an embedded YouTube video and two Friend Connect widgets, the user can be logged three times by Google's servers by merely visiting the site. Friend Connect is an interesting service that will likely be very popular. Therein lies the risk: Friends Connect and future generations of social networking applications will amplify user disclosure and facilitate uniquely identifying users.

Figure 7-14 Google's new Friends Connect service enables web authors to add social networking functions to their sites.

EMBEDDED MAPS

Another common practice, and subsequent cross-site disclosure risk, is embedding maps within web pages. From hotels to tourist attractions, to business and social events, web authors rely upon third-party mapping services such as Google Maps and MapQuest to provide easy-to-use, interactive maps for their site's visitors. Unfortunately, the practice also informs the mapping service of the IP address of the visitor, HTTP cookies, the site the user came from, and a location he or she is interested in. For example, in Figure 7-15, a web author for an academic conference directly embedded a Google map into the conference web site.[77] Thus, every potential conference attendee who browses the conference's directions page immediately informs Google, and possibly Yahoo! and MapQuest, of his or her interest in the conference and probable attendance. With thousands, perhaps millions, of embedded maps in sites across the web, this practice greatly extends the cross-site visibility of large online companies such as Google and Yahoo!. The future of information-disclosure risks associated with embedded mapping is likely to worsen. Simple mapping is giving way to *mapplets* (or *mashups*), which combine mapping with virtually any type of location-based data (think homes for sale, local coffee shops, or driving ranges). The end result is a growth in the type and quantity of information disclosed via embedded maps and their progeny.

> **NOTE**
>
> The term *mashup* applies to far more than just mapping. Mashups are a core Web 2.0 tenet and apply to web applications that combine more than one data source into single integrated tools. Beyond mapping, examples include combinations of images, videos, news, search, and shopping data.[78] Mashups increase information-disclosure risks because their use can share user information with many disparate mashup data source providers.
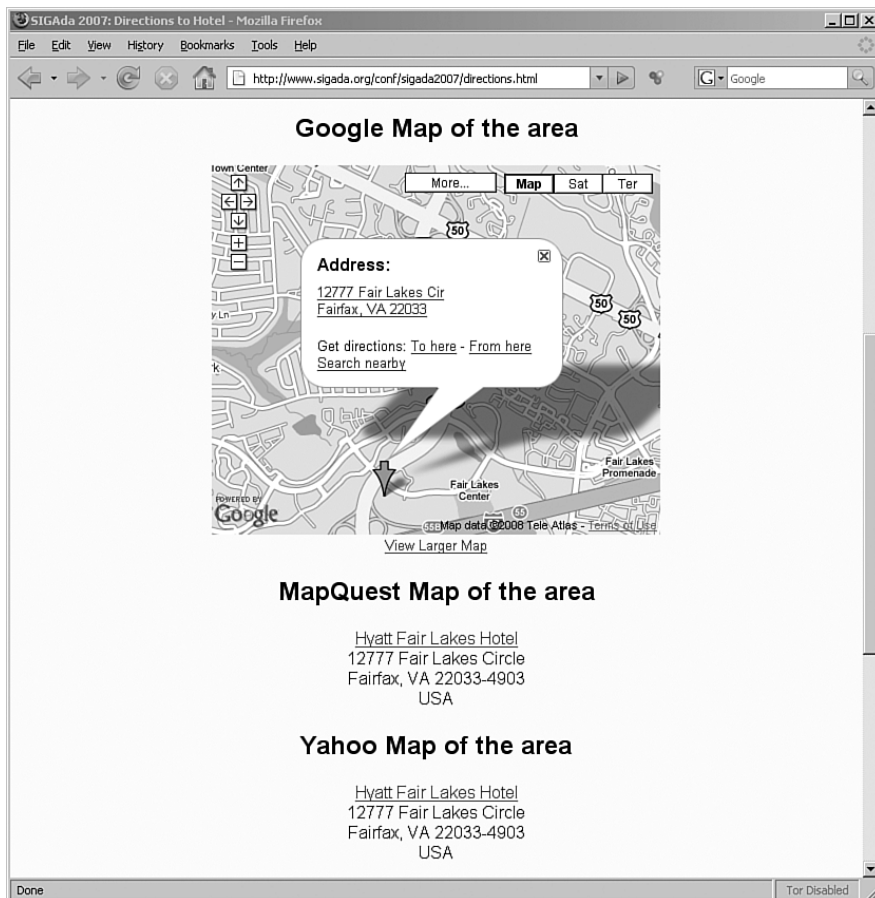
Figure 7-15 Embedded maps in web pages immediately inform the mapping service of the user's visit to a given web page, as well as that user's interest in a specific area, as is the case for this academic conference.

SUMMARY

Web browsing isn't a one-to-one conversation with a single web site. Instead, embedded content such as maps, images, videos, advertisements, web-analytics code, and social networking widgets immediately disclose each user's visit to a third party when that user merely views a page in his or her web browser. Web authors and webmasters gain a great

deal of value by embedding these "small snippets of code" in their web sites, such as gaining access to advertising revenue, free web-analytics reports, improved customer contact, and richer, more compelling web content. The true benefit is to the online companies, which gain a greatly increased field of view that isn't constrained to their own properties, but instead encompasses a major swath of the Internet. As these companies innovate and field compelling new services, expect their field of view to increase further, as webmasters and web authors across the Internet embed new and better content. A key conclusion is that embedded third-party content forces the user to accept many different privacy policies from many different companies, most likely without even being aware of it. This creates a lowest common denominator effect of privacy policies; your real privacy in terms of visiting a web site is the equivalent of the worst policy of all the sites embedded there. This is a huge issue. Consider the MSNBC example earlier in the chapter. Most users might be aware that they fall under the MSNBC privacy policy, but they likely are not aware of the information being collected by the ten other companies providing embedded content, let alone the finer points of each of these companies' privacy policies, if they even exist.

ENDNOTES

1. Rory Cellan-Jones, "Web Creator Rejects Net Tracking," BBC, 17 March 2008. http://news.bbc.co.uk/2/hi/technology/7299875.stm, last accessed 21 April 2008.

2. Sometimes third parties actually employ web bugs, typically in the form of a 1×1 transparent GIF.

3. For advertisements, these third-party servers are often called *central ad servers.*

4. Louise Story, "To Aim Ads, Web Is Keeping Closer Eye on You," *The New York Times,* 10 March 2008. www.nytimes.com/2008/03/10/technology/10privacy.html?_ r=1&oref=slogin, last accessed 2 May 2008.

5. *Ibid.*

6. Ina Fried, "Microsoft Quietly Offering Ad-Funded Works," CNET, 18 April 2008. www.news.com/8301-13860_3-9922750-56.html?tag=nefd.top, last accessed 10 May 2008.

7. Louise Story, "Company Will Monitor Phone Calls to Tailor Ads," *The New York Times,* 24 September 2007. www.nytimes.com/2007/09/24/business/media/ 24adcol.html?_r=3&ei=5065&en=5822f6a12e575488&ex=1191297600&partner=M YWAY&pagewanted=print&oref=slogin&oref=slogin&oref=slogin, last accessed 10 May 2008.

8. Catherine Holahan, "Google's DoubleClick Strategic Move," *BusinessWeek,* 14 April 2007. www.businessweek.com/technology/content/apr2007/tc20070414_675511.htm, last accessed 2 May 2008.

9. Mark Hendrickson, "Yahoo! Acquires Ad Network Blue Lithium," *TechCrunch,* 4 September 2007. www.techcrunch.com/2007/09/04/yahoo-acquires-ad-network-bluelithium/, last accessed 10 May 2008.

10. Kevin Delaney, "Yahoo!–Google Pact May Be Close," *The Wall Street Journal,* 2 May 2008. http://online.wsj.com/article/SB120968562237161201.html?mod=rss_whats_news_technology, last accessed 2 May 2008.

11. Nate Anderson, "Study: Ads in Online Shows Work Better Than Ads on TV," Ars Technica, 26 December 2007. http://arstechnica.com/news.ars/post/20071226-study-ads-in-online-video-work-better-than-ads-on-tv.html, last accessed 17 May 2008.

12. Louise Story, "Where Every Ad Knows Your Name," *The New York Times* Bits Blog, 10 March 2008. http://bits.blogs.nytimes.com/2008/03/10/where-every-ad-knows-your-name/?hp, last accessed 10 May 2008.

13. "Microsoft to Test New Measure of Web Ads," Reuters, 25 February 2008. http://www.reuters.com/article/internetNews/idUSWNAS219120080225?feedType=RSS&feedName=internetNews&pageNumber=1&virtualBrandChannel=0, last accessed 11 May 2008.

14. David Legard, "IDC: Consolidation to Windows Won't Happen," *LinuxWorld,* 27 April 2004. www.linuxworld.com.au/index.php/id;940707233;fp;2;fpid;1, last accessed 2 July 2008.

15. "Google Ads Are a Free Speech Issue," Slashdot, 28 February 2007. http://yro.slashdot.org/yro/07/02/28/0139222.shtml, last accessed 10 May 2008.

16. This section focuses on information disclosure risks and does not cover Click Fraud, the illegitimate clicking of advertisements to generate revenue or charge an advertiser. See www.businessweek.com/magazine/content/06_40/b4003001.htm for more information.

17. Jeffrey Graham, "Google's AdSense a Bonanza for Some Web Sites," *USA Today,* 10 March 2008. www.usatoday.com/tech/news/2005-03-10-google-ads-usat_x.htm, last accessed 5 May 2008.

18. "What Is AdSense for Feeds?" Google AdSense Help Center. www.google.com/adsense/support/bin/answer.py?hl=en&answer=20012, last accessed 5 May 2008.

19. "Google AdSense for Search," Google AdSense, 2008. www.google.com/adsense/static/en_US/WsOverview.html?hl=en_US, last accessed 5 May 2008.

20. "AdSense for Mobile Content," GoogleAdSense, 2008. www.google.com/adsense/www/en_US/mobile/, last accessed 5 May 2008.

21. Shamim Samadi, "AdSense for Video Now in Beta," Google Blog, 21 February 2008. http://googleblog.blogspot.com/2008/02/adsense-for-video-now-in-beta.html, last accessed 5 May 2008.

22. "Google Testing AdSense Cost Per Action (CPA)," Search Engine Roundtable, 21 June 2006. www.seroundtable.com/archives/003988.html, last accessed 17 May 2008.

23. Juan Carlos Perez, "Merchants Say eBay Ad Programs Drive Buyers Away," *ITWorld,* 8 October 2007. www.itworld.com/Tech/2403/071008ebayads/, last accessed 11 May 2008.

24. eBay also has its own contextual advertising network; see http://slashdot.org/article.pl?sid=06/06/11/013216.

25. Google AdSense sign-in page, Google AdSense, 2008. www.google.com/adsense/login/en_US/?gsessionid=JjtK8kJKBuw, last accessed 5 May 2008.

26. Mitch Ratcliffe, "How Google Falls: Unprofitable in 2009," ZDNet, Rational Rants Blog, 15 February 2007. http://blogs.zdnet.com/Ratcliffe/?p=265, last accessed 10 May 2008.

27. Mark Ward, "Searching for the Net's Big Thing," BBC News, 13 March 2006. http://news.bbc.co.uk/1/hi/technology/4780648.stm, last accessed 6 May 2008.

28. "Company Overview," Corporate Information, Google, 2008. www.google.com/intl/en/corporate/index.html, last accessed 6 May 2008.

29. Google is also experimenting with video ads on search result pages; see http://bits.blogs.nytimes.com/2008/02/14/google-tests-video-ads-on-search-results-pages/.

30. For an insightful overview of AdWords attacks, see StankDawg's Defcon 13 talk, available at www.defcon.org/html/links/defcon-media-archives.html#dc_13.

31. "Google Partners Put You in Front of More Potential Customers," Google, 2008. https://adwords.google.com/select/partner.html, last accessed 6 May 2008.

32. "About AskEraser," Ask.com, 2008. http://sp.ask.com/en/docs/about/askeraser.shtml, last accessed 5 May 2008.

33. Thomas Claburn, "Google Keeps What Ask.com Erases," *Information Week,* 13 December 2007. www.informationweek.com/news/internet/ebusiness/showArticle.jhtml;jsessionid=BQP3K401V1M1QQSNDLRSKH0CJUNN2JVN?articleID=204802233&_requestid=54316, last accessed 5 May 2008.

34. Gwendolyn Mariano, "DoubleClick Able to Settle Privacy Suits," CNET News, 21 May 2002. www.news.com/DoubleClick-able-to-settle-privacy-suits/2100-1023_3-919895.html, last accessed 16 May 2008.

35. Steve Lohr, "Google Deal Said to Bring U.S. Scrutiny," *The New York Times,* 29 May 2007. www.nytimes.com/2007/05/29/technology/29antitrust.html?_r=2&ref=business&oref=slogin&oref=slogin, last accessed 16 May 2008.

36. "E.U. Lobby Says Google, DoubleClick Merger Hurts Privacy," Reuters, 20 December 2007. www.reuters.com/article/internetNews/idUSL2051059120071220, last accessed 16 May 2008.

37. Grant Gross, "FTC Approves Google/DoubleClick Deal," *Macworld,* 20 December 2007. www.macworld.com/article/131204/2007/12/doubleclick.html, last accessed 16 May 2008.

38. "European Regulators Approve Google–DoubleClick Deal," CBC News, 11 March 2008. www.cbc.ca/technology/story/2008/03/11/tech-google-doubleclick.html, last accessed 16 May 2008.

39. "Google Closes Acquisition of DoubleClick," Google Press Release, 11 March 2008. www.google.com/intl/en/press/pressrel/20080311_doubleclick.html, last accessed 15 May 2008.

40. "DART Ad-Serving and Search Cookie Opt-Out," DoubleClick, 2008. www.doubleclick.com/privacy/dart_adserving.aspx, last accessed 16 May 2008.

41. Note that DoubleClick's opt-out cookie does not prevent ad targeting based on the user's operating system, Windows version, local time, or IP address.

42. Louise Story and Miguel Helft, "Google Buys an Online Ad Firm for $3.1 Billion," *The New York Times,* 14 April 2007. www.nytimes.com/2007/04/14/technology/14deal.html, last accessed 16 May 2008.

43. Stefanie Olsen, "Privacy Concerns Dog Google–DoubleClick Deal," CNET News, 17 April 2007. www.news.com/Privacy-concerns-dog-Google-DoubleClick-deal/2100-1024_3-6177029.html, last accessed 16 May 2008.

44. Electronic Privacy Information Center Complaint with the Federal Trade Commission, 20 April 2007. http://epic.org/privacy/ftc/google/epic_complaint.pdf, last accessed 16 May 2008.

45. Clint Boulton, "Meet Google: Search Giant, Monopolist Extraordinaire," Google Watch Blog, Eweek.com, 24 December 2007. http://googlewatch.eweek.com/content/google_vs_microsoft/meet_google_search_giant_monopolist_extraordinaire.html, last accessed 16 May 2008.

46. Frank Rose, "Microsoft's Bid for Yahoo! Is All About Big-Budget Brand Advertising," *Wired Magazine,* 24 March 2008. www.wired.com/techbiz/it/magazine/16-04/bz_microsoft_yahoo, last accessed 15 May 2008.

47. StankDawg, "Hacking Google AdWords," Defcon 13, 2005. www.defcon.org/images/defcon-13/dc13-presentations/DC_13-Stankdawg.pdf, last accessed 2 July 2008.

48. Brian Krebs, "Virus Writers Taint Google Ad Links," Security Fix Blog, WashingtonPost.com, 25 April 2007. http://blog.washingtonpost.com/securityfix/2007/04/virus_writers_taint_google_ad.html, last accessed 4 May 2008.

49. Brian Krebs, "Hacked Ad Seen on MySpace Served Spyware to a Million," Security Fix Blog, WashingtonPost.com, 19 July 2006.

50. Betsy Schiffman, "Hackers Use Banner Ads on Major Sites to Hijack Your PC," Wired.com, 15 November 2007. http://www.wired.com/techbiz/media/news/2007/11/doubleclick, last accessed 4 May 2008.

51. Nonmalicious Flash applets hosted on web pages have also been shown to be vulnerable to attack. See Dan Goodin's "Serious Flash Vulns Menace at Least 10,000 Websites," www.theregister.co.uk/2007/12/21/flash_vulnerability_menace/

52. Eric Auchard, "Adobe, Yahoo! test running ads inside PDF documents." Reuters, 28 November 2007. www.reuters.com/article/marketsNews/idUKN2754715120071129?rpc=44&sp=true, last accessed 4 May 2008.

53. Robert Freeman, "I'm Feeling Lucky," Frequency X Blog, IBM Internet Security Systems, 29 April 2008. http://blogs.iss.net/archive/FeelingLucky.html, last accessed 17 May 2008.

54. Jakob Nielsen, "Banner Blindness: Old and New Findings," Alertbox blog, 20 August 2007. www.useit.com/alertbox/banner-blindness.html, last accessed 7 May 2008.

55. An interesting exception is NCSoft's game *City of Heroes,* which makes viewing in-game advertisements optional; see http://yro.slashdot.org/article.pl?sid=08/04/06/0554230.

56. Cade Metz, "Network Solutions Hijacks Customer Subdomains for Ad Fest," *The Register,* 11 April 2008. www.theregister.co.uk/2008/04/11/network_solutions_sub_domain_parking/, last accessed 8 May 2008.

57. Henry Blodget, "Compete CEO: ISPs Sell Clickstreams for $5 a Month," Seeking Alpha, 13 March 2007. http://seekingalpha.com/article/29449-compete-ceo-isps-sell-clickstreams-for-5-a-month, last accessed 8 May 2008.

58. "U.K. ISPs to Start Tracking to Serve You Ads," *TechDirt,* 18 February 2008. http://techdirt.com/articles/20080218/024203278.shtml, last accessed 8 May 2008.

59. For example, an Amazon cookie disclosed by the user's browser might have been associated with previous Amazon purchases and, hence, tied to shipping and billing information.

60. "Facebook Beacon," Facebook Business Solutions, 2008. www.facebook.com/business/?beacon, last accessed 18 May 2008.

61. Henry Blodget, "Facebook's 'Beacon' Infuriate Users, MoveOn," Silicon Alley Insider, 21 November 2007. www.alleyinsider.com/2007/11/facebooks-beaco.html, last accessed 18 May 2008.

62. Josh Catone, "Is Facebook Really Ruining Christmas?" ReadWriteWeb, 21 November 2007. www.readwriteweb.com/archives/facebook_moveon_beacon_privacy.php, last accessed 2 July 2008.

63. Kristen Nicole, "Google Analytics Gets a Beautiful New Interface," Mashable.com, 8 May 2007. http://mashable.com/2007/05/08/google-analytics/, last accessed 15 May 2008.

64. The density of the code could also be seen as an attempt to decrease the size of the file, to improve response time.

65. Blogger Garett Rogers provides an introductory analysis to the operation of `urchin.js` at http://blogs.zdnet.com/Google/?p=39.

66. The full script is more 13 pages long. You can view it at www.google-analytics.com/urchin.js

67. For additional information on Google Analytics cookies, see www.customizegoogle.com/block-google-analytics-cookies.html.

68. Garett Rogers, "How Do I Know the Number of Google Analytics Accounts?" Googling Google ZDNet Blog, 28 November 2005. http://blogs.zdnet.com/Google/?p=42, last accessed 15 May 2008.

69. "Microsoft adCenter Analytics Registration," Microsoft Digital Advertising Solutions, 2008. http://advertising.microsoft.com/advertising/adcenter-analytics-registration, last accessed 15 May 2008.

70. Aurelie Pols, "Yahoo! Buys Indextools: 80% of the Functionality of Omniture for Free!" OX2 Web Analytics Blog, 8 April 2008. http://webanalytics.ox2.eu/2008/04/15/yahoo-buys-indextools-80-of-the-functionality-of-omniture-for-free/, last accessed 15 May 2008.

71. Nick Gonzalez, "TringMe: Phone Free Click to Call," *Tech Crunch,* 2 October 2007. www.techcrunch.com/2007/10/02/tringme-phone-free-click-to-call/, last accessed 17 May 2008.

72. Erik Schonfeld, "Google Talk Adds a Chatback Widget," *TechCrunch,* 26 February 2008. www.techcrunch.com/2008/02/26/google-talk-adds-a-chatback-widget/, last accessed 17 May 2008.

73. Thai Tran, "Google Maps Mashups 2.0," Google Lat Long Blog, 11 July 2007. http://google-latlong.blogspot.com/2007/07/google-maps-mashups-20.html, last accessed 17 May 2008.

74. Dave Taylor, "How Can I Add a Google Search Box to My Web Site?" Ask Dave Taylor, 10 December 2004. www.askdavetaylor.com/how_can_i_add_a_google_search_box_to_my_web_site.html, last accessed 19 May 2008.

75. "The Power of Google Search on Your Site," Google AdSense, 2008. www.google.com/adsense/www/en_US/afs/index.html, last accessed 19 May 2008.

76. "More Info About Google Friend Connect," Google Friend Connect BETA, 2008. www.google.com/friendconnect/home/moreinfo, last accessed 18 May 2008.

77. SIGAda2007, Association for Computing Machinery Special Interest Group on Ada Conference web site, 10 October 2007. www.sigada.org/conf/sigada2007/directions.html, last accessed 23 May 2008.

78. "Mashup (Web Application Hybrid)," Wikipedia, 2 July 2008. http://en.wikipedia.org/wiki/Mashup_(web_application_hybrid), last accessed 2 July 2008.