UNDERSTANDING INTERNET PRIVACY – GENERAL INFORMATION

This document explains what degree of privacy you can expect while you surf on the world-wide web and how you can control what information is given out about you. The important point to note is that you are in control — nobody can obtain personal information about you unless you explicitly allow them to.

There are various ways that a site has of obtaining information about you. When you request a page from a site, a certain amount of information is automatically disclosed in the page-request that your browser makes on your behalf. Once you've received the page, the site could ask your browser for some additional information. While you are getting the page, the site could be tracking you by taking notes about your behavior and storing those notes in an area of your hard disk (cookies) which it can read back later. And whenever you fill out and submit a form, the information on that form is sent to the site. Each of these aspects is described below in detail.

Requesting a Page

When you request a page from a site, a small amount of information about you is given to that site. In particular, the site is told the three items listed below. Beyond that, the site is unable to obtain any other information about you with out your knowledge — it does not know your e-mail address and certainly does not know your name.

1. Operating Environment

The site is told something about your operating environment such as the type of browser you are using and perhaps the operating system on which you are running. This helps the site present the page that you are requesting in a way that will best display on your screen. As an example, the site might be told that you are using the English version of Netscape 6 and are running under the Windows 98 operating system. Such information is not in any way personal so your privacy is not compromised by having it divulged.

2. Internet Address

The site is told the Internet address that you are currently using. This is sometimes referred to as your IP (or Internet Protocol) address. The site needs your IP address so that it knows where to send the page that you are requesting. IP addresses are usually registered to Internet service providers and not to individuals; each time you dial up an Internet service provider, you are assigned one of their many IP addresses at random to use for the duration of your session. So the site you are visiting can determine, for example, that an AOL member just requested a page but it cannot determine which AOL member.

Your IP address is not your e-mail address — they are two different things. Your e-mail address is the address to which your incoming e-mail is sent and uniquely identifies you in Cyberspace just as your social security number identifies you in the real world. Your IP address, on the other hand, is a temporary address that you are using for the duration of a session in order to get the pages you are requesting. It is no more a part of your identity than is the phone number of a pay telephone which you happen to be using when making a phone call.

But if you are concerned and want to block your IP address from being given out, see the section on Hiding Your Internet Address.

3. Referrer

The site is also told where you just came from. In other words, it knows which page you were reading when you clicked on the link to the page you are now requesting. This allows the site to know which other site referred you to it. Also, as you traverse the site, it allows the site to know where in the site you were most recently.

After the Page is Received

After you receive a page from a site, that page is displayed. The page might contain programs, referred to as JavaScript code, which will then execute on your machine. JavaScript code has the ability to request some information about your machine and to send such information back to the site.

If you do not want any additional information given out, you can easily prevent it. Whether or not your browser allows JavaScript code to execute is controlled by your preference settings. That preference is initially set to allow JavaScript to execute. By changing that preference, you will be preventing the site from requesting and transmitting this information.

The information that the site can request by using JavaScript code in this manner is usually not very interesting. It includes such things as the number (but not the names) of the sites you previously visited, whether or not your browser can execute programs written in a language called Java, the number and type of plugins you have installed in your browser, the height and width of the browser window, etc..

JavaScript code is normally incapable of obtaining any information about you that would seriously compromise your privacy. However, with your permission, JavaScript code can obtain much more personal information. In fact, it could even read information from arbitrary files on your hard disk and transfer that information back to the site. But you have to grant your permission before any of this can happen. You'll know when the site is attempting to use JavaScript in this manner because a box will appear asking you to grant your permission. You should not grant it unless you have absolute trust in that site. If you refuse, the JavaScript code is rendered harmless.

Downloading a File

When you are requesting a file (as opposed to a viewable page), your e-mail address might be divulged as a courtesy to the site. You know when you are requesting a file because its address starts with "ftp://" instead of the more usual "http://".

One of your preference settings determines if your e-mail address should be sent as your password when you request files. This preference is initially set to not send your e-mail address so, unless you've changed it, your e-mail address will not be divulged.

Being Tracked by Cookies

Since the site does not know who you are, it cannot possibly be collecting any information on you and has no knowledge of any previous times that you visited the site. It does not even know what you've done while on the site other than knowing where on the site you just came from.

However there are times when it would be to your advantage to allow a site to know something about your previous visits to the site. For example, if you were previously reading a long document and got as far as page 17, it would be nice if the site could take you immediately to page 17 on your next visit.

The only way a site has of remembering information that it can associate with you is to store the information onto your hard disk and to read it back each time you interact with the site. Such pieces of information are called cookies for lack of a better name. Of course the site cannot store a cookie directly but instead asks your browser to do that on its behalf. And your browser will not store a cookie without your permission (see the section on Controlling Your Cookies). Once a site has stored a cookie, it can read that cookie in the future without having to get permission from you. But the site can read only the cookies that it has stored — it cannot read the cookies that other sites have stored.

Don't be alarmed — a site cannot write to arbitrary places on your disk. The cookies that it stores go into one specific file, called your cookie file. And the site can't even write there unless you give it permission to do so. Similarly, the site can't read arbitrary information from your disk either.

If a site can store a cookie, it can keep track of all the things that you've done by simply writing these things into a cookie which it keeps updating. By this means it can build up a profile on you. This may be a good thing or a bad thing depending on what the site intends to do with the information. For example, it would be a good thing if a book-seller knew that you frequently looked for information on dogs so that it could tell you if a new dog book became available since your last visit. It would be a bad thing if it then sold that information to the local dog pound so they could cross-check for potential dog owners who do not have valid dog licenses.

Encountering Foreign Cookies

When a site stores a cookie, it is the only site that is able to read that cookie in the future. That permits a site to build up a profile on your behavior while you are at that site but not on your behavior in general while surfing the web. So at least you have some assurance that the data that is collected on you (with your permission of course) is site specific and nobody can build up a universal database on you.

But suppose that while you are visiting site sheep.com, a cookie gets stored not by sheep.com but by some marketing site called wolf.com. And sheep.com can cause that to happen very simply by having an image from wolf.com displayed on its home page. So when you visit sheep.com, you are really making a side trip to wolf.com to get the image and wolf.com can store the cookie at that time. Suppose that wolf.com has enlisted many other sites to also display its cookie-storing image. Now wolf.com will be building up a cookie that contains information about your accumulative behavior at all of these sites. And the more sites that wolf.com can entice to display its image, the more encompassing a profile it can build on you.

Such cookies that are stored by the site other than the one that you think you are visiting are called foreign cookies. If you are concerned about the privacy implications of foreign cookies but not concerned about ordinary cookies, you could give permission for sites to store ordinary cookies only but not store foreign ones.

Controlling Your Cookies

The way you give permission for a site to use (store and/or read) cookies is by your preference settings. Your preference could be that your browser should allow sites to use all (foreign as well as non-foreign) cookies, allow sites to use non-foreign cookies only, or not allow sites to use cookies. Furthermore, in your preference settings you could state that you want to be warned before your browser will store any cookie. When you first install your browser, your preferences are set to allow all sites to use all cookies with no warning given when a cookie is being stored; you will need to explicitly change your preference setting if that is not what you want.

If you don't consider cookies to be a privacy invasion and don't care who stores cookies on your machine, you would keep your preference settings unchanged. On the other hand, if you are paranoid and don't want to allow any site to store cookies, you would change your preferences to not allow sites to use cookies. But there might be a middle ground whereby you want to allow specific sites to store cookies (your brokerage house, for example, might require cookies before it can let you examine your portfolio), prohibit other specific sites (those notorious for engaging in questionable marketing practices), and be asked about all remaining sites.

You can accomplish this middle ground by setting your preferences to allow sites to use cookies but warning you first. In that case, a box will pop up each time a site attempts

to store a cookie. That box will identify the site (it might not be the site that you are currently visiting, as in the case of foreign cookies) and ask you if you want to allow the cookie to be stored. It will also ask you if you want to remember your decision on behalf of this site. If you accept the cookie and specify that you want the decision remembered, the browser will automatically grant all future cookie-storing attempts made by this particular site without giving any warning. On the other hand, if you reject the cookie and ask to have the decision remembered, the browser will automatically reject all future cookie-storing attempts from this site.

By using the Cookie Manager, you can bring up a list of cookies that have been stored on your hard disk as well as a list of sites for which you have asked to have the cookie-storing decisions remembered. And you can selectively delete any of the cookies or sites in these lists.

Evading Cookies

It should be mentioned that even if you have disabled cookies, the site still has ways of tracking you, at least while you remain at that site. Presented here is one example.

The site could store the information not in a cookie on your machine but rather in the links that it lets you fetch. Each link that it presents for you to click on contains the address of the next page to fetch. But the site could customize that link specifically for you so that it contains a bit of tracking information as well.

To make this clear, suppose that you visit a site called trackme.com. That site presents you with its home page and that page contains a link to a second page. What you see on your screen is some text describing the link (for example, "visit our second page"). In addition to the visible text, the link also contains the address of the second page, such as trackme.com/secondpage. But suppose the link on the home page doesn't contain just trackme.com/secondpage but contains something like trackme.com/secondpage?0 instead. The "?0" might be a code saying that you haven't visited the second page yet. Suppose you click on this link and view the second page. Then you click on a link on the second page that gets you back to the home page. The home page that the site presents to you this time differs from the one it sent you previously in that the link back to trackme.com/secondpage now contains trackme.com/secondpage?1. The site is now using the page itself (rather than a cookie) to keep track of where you've been and what things you've clicked on.

The good news is that this sort of tracking works only as long as you remain at the site and visit its related pages. Once you leave the site all of this information is lost. If you should then return again later you will be presented with the "trackme.com/secondpage?0" link all over again. (Of course if you bookmark a page from such a site, when you return to that page via the bookmark that tracking information will still be there.)

Submitting Information on Forms

Of course if you voluntarily chose to divulge information to the site, such as by submitting a form that the site presents to you, you are knowingly providing the site with whatever personal information you filled in. The site is then free to store that information in its data base and to use the information in any way it sees fit. For your protection, many sites are now voluntarily establishing privacy policies which dictate what they will and will not do with any information you give them. Each site determines its own privacy policy and makes that policy available for you to view.

Keep in mind that there is no policing of sites with regards to their privacy policies and they can say in it whatever they want. So when it comes right down to it, the final decision as to whether you want to voluntarily submit information to a site will depend on how much trust you have in the site. You might be inclined to believe what is said in the privacy policy of http://home.netscape.com whereas you might be justified in being dubious about any policy offered by http://www.ripoff.com

You will often find yourself entering the same information on the forms of many different sites. For example, all sites that sell you something will probably ask for your name, your shipping address, and your credit card number. It's tedious to have to type this in every time. Instead you can ask the Form Manager to save the information from a particular form and then pre-fill that information onto forms that you encounter in the future. The Form Manager saves the information on your local machine and not on any website. When the Form Manager pre-fills a form with the saved information, that information is not sent to the site until you submit the form. Once again you are in control — no information is released until you say so.

Divulging your Password

If you are like most users, you've registered for services at various sites. The registration consisted of selecting a user name and password. Each time you return to such a site, you fill out and submit a form containing the user name and password that you selected for that site. To avoid having to remember a different password for each site, especially those you don't visit often, you might have used the same password everywhere. And the same goes for your user name, providing somebody else hadn't already taken it.

So each site that you registered with has a record of two important pieces of information about you — your user name and password. And if this is the same user name and password that you always use, an unscrupulous site administrator at any one of these sites has enough information to go impersonating you by logging in to other sites at which you are registered. You might not be concerned about this because it really doesn't hurt you if somebody logged in as you at some newspaper site and read what was going on in the world. But you might be concerned if somebody managed to guess which stockbroker you used, and logged in as you and made some stock transactions.

The way to protect yourself, of course, is to use a different password at every site that you register with. But this means you have to keep track of every password that you've ever used. The Password Manager in the browser can help you out by remembering the user name and password that you used when you last logged on to a site, and then pre-filling that information onto the log-in form the next time you visit that site. You can then either submit the log-in form with these pre-filled values, or change them before submitting if they are not what you want.

The Password Manager also allows you to see which user names you have stored for which sites. And it allows you to selectively delete any of these items if you wish.

Hiding Your Internet Address

When you request to see a page from a site, your browser needs to tell the site your Internet address (IP address) so the site knows where to send the page. This is in effect your return address. Your internet service provider has many IP addresses assigned to it and it selects one for you to use each time you start a session Every time you connect to your provider you will be given a new IP address.

Some users have their own fixed IP addresses which they use every time they connect to the Internet. But these users are in the minority and if you are one of them you undoubtedly know about it. So if you have not heard anything to the contrary, you can assume that you get a new IP address for each session.

Even though it's only a temporary address, you might not want that information to be given to a site you intend to visit. But if your browser doesn't provide this information, the site won't know where to deliver the requested page. So this is the one piece of information that you can't ask your browser not to reveal.

If you really want to hide your IP address from the site, you need to use some trusted intermediate site. You go to the intermediate site and tell it the name of the site whose page you want. The intermediate site requests the page on your behalf, using its own IP address as the return address. Then, when it gets the page, it forwards it on to you. The site that supplied the page never gets to see your IP address.

There are several sites that provide such services. Use your favorite search engine to find them — try search words such as "anonymous" and "surfing".

Summary

True privacy on the Internet is hard to achieve. Given the technology today, the most practical approach is to limit the exposure of your personal information on the Internet. Be sure of who you are dealing with, and use old fashion common sense in releasing information and/or evaluating what information is gathered about you.