

Note: The following piece dates from 2000-01, at which time the publication was known as the Yale Symposium on Law and Technology. Page numbering, editorial style, and citation format may differ from that of the Yale Journal of Law and Technology.

Internet Privacy: Who Makes the Rules[†]

Richard M. Smith[‡]

Abstract: Richard M. Smith, Chief Technology Officer of the Privacy Foundation, discusses the ways emerging technology infringes consumer privacy on the Internet. He believes the widespread use of cookies and the growing use of online profiling by companies like DoubleClick create serious privacy problems for people who use the Internet. The solution lies in combining the efforts of programmers, who can find ways to eliminate these profiling mechanisms online, and of lawyers, who can structure legal rules to proscribe information misuse.

Cite as: 4 YALE SYMP. L. & TECH. 2 (2001)

I. INTRODUCTION

¶1 In my talk today, I would like to introduce some of the privacy concerns currently making news on the Internet. I will first look at who the players in the field are. I will explore the roles, sometimes competing and sometimes complementary, of lawmakers and programmers in resolving privacy issues. Second, I would like to show how the unique environment of the Internet has driven the development of privacy issues, notably online profiling, mostly through enabling technologies. Third, in an effort to help make some of these issues more tangible, I will discuss the role of DoubleClick, both as a source of trouble and as a would-be solver of problems. Finally, given the trajectory of development as it exists today, I will explore some possible developments in the near future of online profiling.

II. WHO ARE THE PLAYERS?

¶2 Last September, I took a sabbatical from my company and, as sort of a full-time avocation, began looking into this issue of how companies spy. I

[†] Edited transcript of remarks delivered to the Yale Law and Technology Society on September 12, 2000.

[‡] Chief Technology Officer of the Privacy Foundation

have come up with a lot of interesting stories. One of the things that has happened since I got involved with this is that I have met a lot of media folks, because they're interested in this issue. Since early 1999, there has been a lot of press attention on the issue of Internet privacy. I also happened to meet a lot of lawyers since many of them have taken an interest in this issue.

¶3 In my previous life, my dealings with lawyers were limited mostly to contracts (you get into a lot of licensing agreements running a software business). Now, with my new hobby, I have discovered a new aspect of things. This is what I want to talk about today. The title of my talk is "Privacy: Who Gets to Make the Rules, The Programmers or the Lawyers?" I chose that title because over the last year, I have come to realize while talking with lawyers that programmers think they get to make the rules since they create the products. One of the games programmers play when they create the products is that they create their own little world by making and determining the rules of that world. Then, there are the lawyers who actually make the laws. If you go to a legislative body, you will meet a lot of lawyers who are in the business of making rules. For example, at the Federal Trade Commission (FTC) the people there are all lawyers too (or at least they all seem to be lawyers) and they are in the business of making rules. I have also talked to what I usually call ambulance chasers (but they call themselves class action lawyers) who like to make rules by setting precedents. Last but not least, I have chatted with a number of state attorney generals' offices where they like to make rules too about Internet privacy issues.

¶4 This creates an interesting situation in which programmers (my own background) have built software systems and really have not interacted with lawyers. If we go back five or six years, most programmers were limited to worrying about copy infringement. They did not want to steal code too well, if you will, from somebody else and get into copyright issues. Even six or seven years ago there was not a lot of patent activity; there was some activity on software patents, but not a lot like there is today on the Internet. The marketing department was always in charge of naming the product, so programmers did not have to deal with trademarks either.

¶5 Things are changing with the Internet. Lawyers and programmers are going to have to communicate with each other. Ten years ago, most computers stood alone. The limit of connectivity was the local area network within the building with a file server and e-mail before Netscape and web browsing became big. Once all of these computers connected to the Internet and communicated with each other, things became a lot more interesting.

¶6 I have been examining the tension that resulted when programmers began building interesting computer systems that work on the Internet and people did not like what they saw. They said, "Wait a minute, there is a little bit too much surveillance or tracking, or spying going on here." The types of lawyers I described earlier started to get wind of this and said, "Wait a minute, those are the issues that we worry about." The essence of the conflict resulted in this manner.

¶7 When we talk to programmers about these issues, they say, "Well, that is inevitable, that is the way the technology works, there is nothing we can do about it." It is like a car manufacturer saying, "All cars have to pollute and all cars have to be dangerous on the road." This the wrong answer, but it is given a lot. What the programmers are really saying is, "We do not consider things like privacy, or security, or liability important enough to worry about. If we did, if we found them interesting, then we would be saying something different." A lot of programmers leave the issue of privacy in software out of the discussion because they just do not consider it important. Part of the resolution to the problems of Internet privacy and spying is going to be educating programmers that they need to worry about these issues, because if they do not, their businesses are going to be in trouble financially.

III. HOW THE INTERNET CREATES THE PROBLEM

¶8 The fundamental problem with the Internet is that tracking does not work intuitively. The best analogy I can give is buying a *New York Times* at the newsstand. In our trade, we call this the "offline" version. I pay my fifty cents at the newsstand and walk away. It is an anonymous transaction. Nobody records the fact that I bought the *New York Times* for today, and if I paid with cash, there is no record that I made the purchase. If I subscribe to the *New York Times* and have it come to my house, the *New York Times* knows that I am a subscriber but that is the limit of what they know. *The New York Times* knows that I get the paper every day and can assume that I read it, but that is the limit of what they know about me.

¶9 On the Internet, online transactions are much more trackable and much less anonymous than their offline counterparts. This is where the Internet gets interesting, and why the law begins to take an interest in it. If I read the *New York Times* online, I have to register. I have to give them my name and e-mail address. They do not ask for much; they could ask for my address and phone number, but all they ask for is the little things. When I go to their site, however, each time I read an article, they know what articles I have read because their computers supplied them to my computer. They know me because I provided my name, my e-mail address, and my ZIP

code. They know what articles I like to read, and they know for how long I like to read them. This information is a valuable commodity for marketers.

¶10 Internet business models are based on the fact that people can be tracked to see what they are interested in, and they can be sold more stuff because businesses know more about them. A look at initial public offerings for dotcoms illustrates the following message: "The advantage that we have in the online world over offline counterparts is that we can spy on people, know what they want, and sell more stuff. This is the business model." It is an interesting business model, and it is clearly one designed by programmers or technical people because they see this opportunity to watch people and profile them. It is not clear whether this is a valid business model, and I think we will see over the next few years that it is an extremely questionable one. However, this is the game going on out there.

A. Tracking: How Is It Done?

¶11 I brought here today a program I wrote on a server at my old company. It echoes back what a person's browser sends out every time he/she goes to a web page. One of the things a browser sends out is an Internet protocol (IP) address, which is basically the telephone number for a computer. Every computer on the Internet now has a unique IP address. Looking at this IP address, somebody may not know who I am, but they do kind of know from where I am. This is true because every IP address maps into some kind of name. This is an arbitrary name that gives you the organization that owns that IP address. If someone at a law firm is investigating General Motors because he/she is going to sue them, General Motors can monitor this individual when he/she visits the website and see what pages he/she is looking at. Other less important information, like what operating system or browser is being used, can also be learned. Someone can also learn what web page a person has come from. Every time a person is out on the Internet, he/she leaves footsteps in the sand.

¶12 Web browser cookies might be one of the most controversial features on the Internet today. Cookies are simple mechanisms introduced around 1994 by Netscape for the purpose of tracking people. A cookie is just a customer identification (ID) number at a company or at a web site. The first time a person comes to the web site, he/she is assigned a customer ID number, the next number in the queue. If a person is the millionth person to visit a website, his/her customer ID number will be 1,000,000. The cookie is stored on the person's hard drive, and each time you go back to that website, the cookie is sent back to them. Accompanying the cookie is other information, like what pages a person has been reading. Cookies enable websites to track people over time. This is where the issue gets

controversial. The *New York Times* could have a complete record of every article a person has read on their website in the past year.

¶13 Cookies were originally invented in order to enable applications like shopping carts on e-commerce sites. If a person goes to an e-commerce site and wants to buy something and is moving from one page to the next, the cookie is needed just to keep track of the person's shopping cart. However, marketers discovered cookies, and now all websites use them for long-term tracking of what people are doing at their websites.

¶14 Basically, a cookie is a spying mechanism. People use different words like "tracking," "spying," "snooping," but it is all the same. It is just a matter of who says it. Cookies were created by programmers to solve a problem; they needed to make shopping carts work, so they said, "Ah, we will use cookies, that seems like a good idea."

¶15 Interestingly, the programmer who invented cookies and Netscape had some consciousness of potential problems in the future. Instead of sending out a single customer ID number to every website which would enable different websites to put their profiling data together and compare notes about each individual person, the Netscape people had the common sense not to recreate the Social Security Number problem. However, other people who came along and built products similar to web browsers, like Real Jukebox from Real Networks, did not have the same common sense. They assigned each customer one ID number upon installation of the software and this number is shipped out to every website where that customer listens to an audio clip or views a short video. Another example is the Sprint PCS webphones, which made the same mistake. Sprint PCS sent people's phone numbers out to the websites visited from the phone. Imagine a scenario in which a person checks his/her stocks from their phone and ten minutes later, a broker calls with some stock tips. At least in the browser world, people are a little better off than in the world of Real Jukebox or cell phones.

B. *Offline Examples*

¶16 Another offline example to consider is the electronic toll transponder systems like EZ Pass and FastLane. Instead of spending fifty cents every time, or getting a toll ticket, a transponder on a person's windshield transmits a serial number associated with an account in an individual's name and the toll is automatically deducted from his/her balance. The interesting consideration is: for what other applications can the transponder be used? There is already talk of possibly charging gasoline purchases to it or food purchases at McDonald's from the drive-through window. Another application might be to put up transmitters on the side of the road and see how fast

people are going. This is an example of the much larger problem of tracking. Paying tolls used to be an anonymous transaction - with the FastLane, your serial number is recorded, plus two cameras record your license plate number in case of disputes. A lot more tracking is going on here than in the cash-based world of tolls.

¶17 Another quick offline example is the Stop & Shop card I signed up for in order to get discounts. There is a serial number associated with my card, like a cookie, and I cannot use it at Star Market. This is why the Stop & Shop card is a really good analogy: every time I go back to a particular website, it's like giving my Stop & Shop card number to the Stop & Shop people. So there are plenty of mechanisms in the offline world that mirror the way we are being tracked in the online world.

IV. THE PROBLEM WITH DOUBLECLICK

¶18 DoubleClick has become the poster child for bad behavior in the privacy arena. What they actually did was probably not all that awful and not that different from what other people are thinking about doing. However, the way that DoubleClick has bumbled about just irritated people.

¶19 DoubleClick provides banner ads on websites. It is not the most interesting and exciting business; who cares about showing those little banner ads? But they track people as well as websites. Since they put banners on thousands of websites, they can track a particular user across many sites. They can see what a particular person is interested in at many different places. They have cookies too, but their cookies are supercookies because they work across sites. Netscape's original intention was that this should not be allowed to happen, but the implementation of the browser allowed companies like DoubleClick who place content-like ads on many sites to track a single user across many sites. For the last three or four years, people have been trying to correct these "third-party cookies" as a violation of the spirit of whatever Netscape had in mind. This is also why lawyers are beginning to get involved in this area of privacy.

¶20 Although DoubleClick's business is sort of boring, they have managed to make it pretty exciting for people who like to make rules. DoubleClick was originally just in the business of providing banner ads and providing tracking services for companies on the Internet. Then, they decided it would be fashionable to have a database with personally identifiable information that they could use in an online world to provide even more targeted banner ads. So they bought a company called Abacus Direct, a direct marketing company in the Denver, CO, area with a very interesting database. Probably everybody in this room here has felt its effect but never knew why.

¶21 Abacus Direct works with all the catalog vendors. Almost every time somebody buys something through a mail-order catalog his/her name, address, phone number and items purchased are sent to Abacus Direct. As a result, Abacus Direct has a database containing about ninety percent of U.S. mail order purchases (amounting to about 2 billion transactions).

¶22 This is something of which people are not really aware. People might become aware of this if they buy something from a mail order catalog and a few months later, they suddenly receive a whole bunch of other catalogs that seem related to their previous purchases. Abacus Direct is in the business of helping catalog merchants market their products. Each merchant contributes a lot of data and then pulls names out to whom they can send new catalogs. This is a cheap way of acquiring new customers. But in order to do this, they need to have a lot of information about consumers. Furthermore, most people are probably not aware that there is a big old computer room somewhere on the outskirts of Denver with this kind of information in it. However, it is perfectly normal in the offline world to build this kind of database.

¶23 DoubleClick's fatal mistake was to decide that it would buy Abacus Direct and try to use the offline database online. There is a group of rulemakers involved here who are driven by the culture of their community, and the culture of the online community is that people are not identified. The unwritten rule is that they cannot try and figure out who people are. There may be ways to do it, but people are not allowed to do it.

¶24 DoubleClick broke this unwritten rule and spent the last year trying to make up for it. They may ultimately succeed, but it is going to be at a very large cost. They focused too much attention on the direct marketing business, and that was a bad business decision for the whole industry. It let the cat out of the bag about what people are doing.

¶25 Let's take a look at what DoubleClick does for a living.

¶26 If we go to my favorite search engine, Alta Vista, and type in a search for "sports cars," we get a banner ad that pops up for a "free new convertible." It is no accident that this ad appeared here. Even though we are on AltaVista, the banner ad is provided by our friends at DoubleClick. The business relationship between DoubleClick and AltaVista goes back about four years. AltaVista made DoubleClick what it is today because AltaVista is a very popular search engine and was able to drive a lot of business Double Click's way. This ad for the convertible is what we call a 'targeted ad.' This ad comes up for the search string we typed in.

¶27 On a month-to-month basis, a company who wants to advertise can buy search strings at the search engine companies so that every time the search strings are typed in, their ad will pop up. The advertiser can buy all the inventory of a given string, or half, or any other amount (there are all sorts of combinations available). This kind of ad is great for the advertiser because it is very targeted. Targeted ads cost about twice as much as normal, untargeted banner ads.

¶28 From a privacy perspective, these companies, the ones doing the targeting, argue that it does not make sense to show a Tums ad to somebody who is not interested in indigestion. One can make a strong case here that this is not so terrible. It may be a little eerie if a person has not noticed it before, but it is not that bad from a privacy perspective as long as the search string is not "remembered." If there is not a database somewhere showing or remembering all the search strings somebody has typed in, I can live with this sort of privacy thing.

¶29 Every time somebody searches for something at AltaVista, DoubleClick is told what that person searched for. It gets interesting when we start wondering whether DoubleClick remembers this information. If I search for a strange word, like "querty next" for whom nobody has paid the advertising rights, DoubleClick generates more revenue if they provide an ad targeted to me than a random ad that would be the likely result for "querty next." Why not show me an ad that is relevant to what I have searched for in the past? The line is crossed here, because we begin remembering things about people.

¶30 This is what online profiling is about. As it turns out, DoubleClick got an incredible amount of negative publicity and press about the issue of online profiling. They had to appear before Congress and dragged me along to say, among other things, how bad they were. It turns out that they do not actually do online profiling. Even though they were enraged and admitting, "We are big guys, we violate privacy left, right, and center," they have not implemented profiling yet. It is quite amusing. They have been sued, the lawyers are screaming at them, but they have not done the really bad stuff yet. They want to, and for competitive reasons they have to say they are bad, but they really are not. All they are doing is generating targeted ads.

¶31 However, they have plans to store information like what keywords we search for in a database. This interests the lawyers. The question is whether we want people spying on each other.

¶32 What I do for a living is talk about how the spying happens, so that people can make the right decision. I do not know the right answer. If a

person talks to DoubleClick, he/she will explain that online profiling is harmless because most of the time they do not know who that individual is, and even if they find out who he/she is by using the Abacus database, this is necessary to save the free Internet as you know it. Targeted advertising is necessary for the sustenance of the free Internet.

¶33 There are a lot of suppositions in DoubleClick's position about whether this online profiling really works or not. However, the legal question here is whether or not it is permissible to spy on people in this manner.

V. THE FUTURE OF ONLINE PROFILING

¶34 Ultimately, the way online profiling will work is that in addition to search engine strings, DoubleClick will look at articles people read and the keywords in the articles that they are reading. If a person reads a story about serial killers, they will pick out "serial killer" and this data will go in his/her profile. They do not actually save all of these keywords. What they will do is build an interest profile, in which each person is rated on about a thousand items. Each item is rated from 0 to 100: 0 means a person is not interested in a particular topic, and 100 means that person is very interested. This information is used to figure out what ads to show a person.

¶35 Some categories are off-limits because they realize that most people do not want certain types of data to be collected. These categories include medical conditions, sex, and anything related to children. Everything else is fair game. For example, information about what types of vacations a person likes to go on and what sports a person plays are all fair game.

¶36 The advertising business is not just about showing ads anymore. Typically, in the offline world, ads are targeted by content. For example, golf ads appear on the sports pages of the *New York Times* or in golf magazines. In the online world, direct marketing is about targeting by people. The assumption is that if a person is interested in golf, and is shown a golf ad, he/she is more likely to respond to the ad. Increased targeting is necessary because click-through rates on banner ads are dropping dramatically. Click-through rates are the rate of response to ads - if a person sees a banner ad and clicks on it, then this is a click-through. Click-through rates used to be two to three percent, now they are under half a percent. I happen to think click-through rates are dropping because the number of ads is going up tenfold, not because people are getting tired. Internet surfers have a limited amount of time to look at ads, and they are being bombarded by many more than they used to be.

¶137 Online profiling has piqued the curiosity of many people. In particular, the FTC has spent three years investigating this. What is interesting is that nobody is really doing it yet, so I find it amusing that the FTC is on the leading edge here. All companies have been doing is talking about it. In the software business, this is what we call 'slideware.' Slideware is a product that does not exist; only the PowerPoint slides for it exist. There have been three years of slideware about online profiling, and the FTC has had three years of meetings about it.

¶138 The real concern about privacy is that if a person starts building databases about people, then that information can be used by the real evil people-the lawyers. Lawyers can subpoena that information and use it against people. The problem with excess collection of data is that it ends up in court when people least expect it. For example if I were involved in a divorce trial, somebody could ask me, "Why were you in Lynn, Massachusetts on such and such a day when you were supposed to be at work?" This information could be obtained from my FastLane pass usage records.

¶139 Excess collection of data is the concern in the privacy community. It is unclear whether my FastLane example constitutes this problem. However, the more data collected, the more people get worried about such privacy issues. The FTC had hearings about this last November. There were 200 people including myself, mostly lawyers, in a room talking about it. This was the first time I learned about this. I have only been involved with privacy for about less than a year, but these hearings sparked my interest. The class action lawyers entered the act when DoubleClick bought Abacus. Then the state attorney generals joined in the act too, claiming they could not change the rules on data treatment. So there are a lot of people jockeying around for this.

¶140 From a programmer's point of view, I ask myself, "How might we prevent DoubleClick from even considering profiling - how can we stop data from getting to DoubleClick so that it is not even possible for them to do profiling?" This is the other angle. We can have laws that deal with this. We can say, as a country, or as a group of countries, "We think online profiling crosses the line of what is acceptable about spying on people. There is nothing like this in the offline world, in which what articles people read help advertisers determine what ads to show them." We can have laws that deal with the issue. But we can also program so that it is impossible for this to be done.

A. Who Will The Rulemakers Be, Lawyers or Programmers?

¶41 The game here is to try and decide who is going to make the rules: Is it going to be the lawyers or the programmers? Last February, DoubleClick realized they had made a huge public relations mistake when they unilaterally announced they were going to do online profiling. Together with their colleagues in the Internet marketing business, DoubleClick said, "We need to come up with rules on how this is going to work. " So with all their lawyers working together, they assembled an industry association to compile rules about the right way to do online profiling. Next, they approached the FTC and said, "We want you to bless these rules. We do not want laws to be made. The last thing in the world that we want is Congress legislating these issues. We just want an understanding of how this should work, so if you bless this, we will go off and do it." The FTC said, "Sure, this means a lot less work for us." So they haggled for three or four months and came up with a 100-page doent drawn up by lawyers outlining how online profiling should operate.

¶42 Behind the scenes, there were other people who said, "Wait a minute. There are actually a couple of people on this planet who can solve this problem for us. One of them is named Bill Gates and the other one is named Steve Case. Their companies make web browsers that support third-party cookies. If the advertising companies do not have third-party cookies, they will not be able to profile. So, why not just ask these web browser companies to not provide third-party cookies?" Twelve attorneys general from various states approached Microsoft and Netscape and suggested they get rid of third-party cookies. About two months ago, Microsoft issued a press release announcing that they had come up with a solution to the online profiling problem.

¶43 Even if the advertising companies did not have obvious public reactions to this, there was a lot of lawyer discussion about antitrust actions against Microsoft and how the advertising companies were going to be run out of business by Microsoft. The advertisers pointed to the deal they had hammered out with the FTC and said, "This is the right way to do this. We do not need this Microsoft patch. We do not need the richest man in the world to determine the privacy practices of the world. This should be done in our back room, not their back room." Microsoft came under a lot of pressure from the advertisers and the FTC to get rid of the patch.

¶44 This is an example of programmers being one up on lawyers, and then the lawyers turning around and getting one up on the programmers with the FTC deal. This drama is still going on. I do not know how it all will end, but it illustrates very well the tension between programmers and lawyers and the contest over who gets to make the rules. It is the best example so far, but it is only one example of a much wider phenomenon. Programmers like to be

in charge and they like the challenge of being one up on the lawyers. I am sure lawyers feel the same.

¶45 The privacy issue in particular is very interesting because of the role Microsoft played. Microsoft is very dominant. Furthermore, Bill Gates can make the decision, "This is the way it is going to be." In this case, there were other people outside of Microsoft who said, "You can't do that," and brought up reasons why he couldn't. But a lot of these issues are really his decision.

¶46 In summary, the way to achieve privacy on the Internet is to send out less data from our computers. The business I am in right now is to find out what data is being sent out and to point out that too much is being sent out. When I deal with lawyers, they want to know what data is being sent out, so I give them some demonstrations about what can be done and how it is done.

B. New Developments – Web Bugs

¶47 We have already talked about targeted advertising. There is another technology that I will call "web bugs." They first got written about last November in *The Washington Post*.

¶48 The example I will show is a web site run by Johnson & Johnson aimed at teenage girls, called "It's My Body." If a person looks at the site, he/she will not see any banner ads. However, if a person looks at the source code, he/she will find our old friend DoubleClick delivering a graphic to this "It's My Body" page that is one pixel by one pixel. To illustrate how small one pixel is, a period is four pixels. Obviously, somebody has gone to some trouble to hide this graphic. The web bug allows them to track people at the web site. The people running the site have hired DoubleClick to provide them with information for marketing purposes. Because DoubleClick operates across many websites, they can see what other sites visitors to the "It's My Body" site have visited, and what they search for at search engines.

¶49 DoubleClick, beyond being in the business of advertising, is in the business of stalking people and monitoring their activity on the Internet. Here, they do it for the best of reasons, to help marketers figure out where to put their ads. DoubleClick has about twenty to thirty thousand of these monitors planted around the Internet.

¶50 Tomorrow, the Privacy Foundation will announce an effort to encourage the Internet marketing industry to confess to tracking with "spotlight tags," as DoubleClick calls them. The newspapers have already

caught on to the use of web bugs, and there are already five or ten related lawsuits. The state of Michigan will file ten notices of attendant action tomorrow.

¶51 The Privacy Foundation will issue a statement about recommended practices, like making these "spotlight tags" visible. They should be little icons that you can push and get information about. My guess is that if they have to confess to tracking and come clean about usage of web bugs, the actual usage will be reduced because companies would just rather not have the publicity. The business benefit that they obtain from this marketing is not worth being caught and exposed by the newspapers. I guarantee that DoubleClick is going to be embarrassed when news of this particular "tag" on the Johnson and Johnson site hits the streets - this is a site for *teenage girls*.

¶52 If a person explores the "It's My Body" site, he/she can see that about half the pages on the site are "bugged." So, people's activity across the site can be monitored. They probably have a pretty non-nefarious purpose, which is, as I mentioned before, just trying to figure out where the banner ads should be placed. The odd thing about the practice is that it is very undisclosed.

¶53 To further demonstrate the lack of appreciation of what is going to come out of it, other sites that have done this are sites such as ProCrit (a drug for treating AIDS), which bugged the home page as well as pages for each of the four main diseases the drug can treat. From this data, DoubleClick can see which disease interests you. The bugs on the ProCrit site were actually taken down, but it took a lot of insistence. Santa.com, a site targeted at children, is another bugged site.

¶54 DoubleClick is not the only company that engages in bugging sites. Their competitors have the same sorts of bugs too. The Privacy Foundation wants to force these companies to confess to their activities - now, if a person reads DoubleClick's privacy policy where they are supposed to disclose what kind of tracking they do, he/she will not find any discussion of "spotlight tags" at all.

C. The Tables are Turning

¶55 What is fascinating about the Internet is that even though companies get to spy on people, we get to spy on the spies. If we take the offline example of Abacus Direct, there is no real way for people to know what Abacus Direct is up to - the only evidence that there is some sort of data exchange occurring is the catalogs we get in the mail. We do not have much knowledge about what Abacus Direct is up to because it is all being done out

of our minds and off our property. However, spying on the Internet is done through our browsers. This is a danger the spies have not yet fully realized.

¶56 The reason we hear today about Internet privacy is that it is possible to see who is spying on whom. From the spies' perspective, that is the downside of the Internet. It is very open from a spying standpoint, but the spies themselves are exposed too.

¶57 I have one final example of how companies like DoubleClick can get into trouble when they do not think through the implications of their actions. Back in June, if a person went to a search engine and typed in the search string, "growing pot," he/she might get a banner from a company called Free Vibe. If you then checked out who Free Vibe was, you would have realized that it was an anti-drug site run by the White House. They had purchased the advertising rights to the "growing pot" string as well as about twenty other drug-related things.

¶58 I brought up this example at a Commerce committee meeting I was at, and a reporter jumped all over it, saying that there had been some previous monkey business involving the drug office trying to influence public opinion using inappropriate ways. One could argue that maybe this is true - it is like Big Brother. This reporter went and checked out the Free Vibe site and called me up and said, "They are using cookies. They are using cookies." I said, "Big deal. Most sites use cookies."

¶59 He had not explained himself very well, but it came out in an article he wrote that DoubleClick had the Free Vibe web site bugged and was providing tracking services for the White House. DoubleClick said what they always say, that they were measuring the site's advertising effectiveness. To say the least, it did not look too good. It did not show the White House in the best light, to see that they were somehow in the monitoring game. Two days after this drama went down, the White House issued a directive banning all cookies at all federal websites because of this one little incident.

¶60 DoubleClick surrounds itself with controversy, and this has an effect on the industry because the lawyers get involved and say, "This smells. This does not feel good. It is not right. We cannot put our finger on it, but we just do not like it." This is the drama going on out there.

VI. CONCLUSION

¶61 The privacy battles are only beginning. As soon as we find a way to plug one hole, whether through legal or technical means, the marketers will find a new hole through which to pull out our personal data. In the end, any

real solution will probably be a patchwork of technical and legal solutions, as well as public pressure.

QUESTION/ANSWER

Question: Are cookie monsters that regularly delete cookies worthwhile?

Answer: Deleting cookies is a pretty good way to go. Once you delete your cookies, you cannot be reassigned the same cookie value. I really hope that Microsoft wins the third-party-cookie battle, because this would solve a lot of problems. Just this week, however, we found a new kind of cookie Microsoft invented and didn't bother to tell anybody about. Microsoft did not provide a way to turn this new cookie off. One of the problems we are going to have here is that even if third-party cookies are turned off, the marketing companies will look for a substitute, and I think they will find it. Still, I am hoping for a technical solution. The cookie blocker program I use, "I Decide," seems pretty good for Internet Explorer.

Question: One of the things you made very clear is that there are both technical and legal solutions. What is your opinion of the need for legal solutions, since technical solutions only take us so far?

Answer: I think nice people should not snoop. I think the business DoubleClick is in is snooping and I think it is fairly disgusting. I wish they would just decide that it just is not worth it to track people on the Internet, and just go back to the business of delivering ads. Another way to go is public pressure. In the long run, we will need legislation and regulation. When will we need it? I don't know. I think we are still sorting it out. I was disappointed when the FTC caved in on the deal they cut with the advertising companies. I think it is going to be a combination of all of the above - a little bit of technical, a little bit of legal, a little bit of public pressure, you know, just lifting the rock up and shining a light under it and seeing what's there.

Question: What do you think about government forces who are snooping, and who actually want more of it, such as the FBI's Carnivore, and the UK's RIP surveillance squad?

Answer: When we think a lot about this stuff, we always bring up Big Brother and Orwell, and all of this, but what we are seeing today is a lot more commercial. Rep. Markie from Massachusetts put it pretty well when he said that "What we're seeing is the government hiring out the private sector to do the snooping."

I don't have a lot of background in wiretapping, but what Carnivore is, is a method of wiretapping for e-mail. This country allows wiretapping, especially to go after organized crime and whatever else they think is important. Obviously, wiretapping can be very invasive, but my understanding is that it is fairly measured. About 1200-2000 wiretaps are allowed a year, not millions of them, and there are rules about when you can do it, and how much information you can get. There are different levels of wiretapping.

It seems to me that if we as a society allow wiretaps on telephones, then the Internet will also be a target for wiretapping. The problem I have with Carnivore is that the FBI interpreted the wiretapping laws for themselves and designed an approach to the Internet on their own. Then the White House came out and said that there needed to be rules that specifically regulate e-mail wiretapping. I think this is a better approach. I don't think that the FBI should do it until they are authorized by Congress. They shouldn't force-fit wiretap law.

Carnivore is a little black box installed on an Ethernet link at the Internet service provider that can watch all of the traffic in and out of the mail server. It can select messages bound in or out for one particular e-mail address. It can look at everything. This is the way Ethernet works. It isn't encrypted. We can probably monitor a good chunk of what people are doing right here in this building. It is fairly easy for anybody to do this kind of monitoring. The concern here is that the FBI is going to look for a terrorist bomb or something like that, mentioned in an e-mail, and this is not what Carnivore is for. It is meant to collect mail going to one account. As long as it is properly done, and we have decided that tapping phones is permissible, then I see tapping e-mail as a logical extension, but it needs specific regulation.

The RIP question in England is much more scary. MI-5, the intelligence service, has built a building in central London to which they have run wires from every Internet service provider in the country. They get a look at all of the TCP/IP traffic in England. They have broader authorizations than there are in this country to look at this kind of stuff. I think this is much more Orwell than Carnivore.

Question: What about the government's incentive to oppose technological solutions to ensuring privacy because they might interfere with government surveillance efforts?

The government has always fought encryption. The underlying structure of the Internet is not conducive to privacy because it was originally designed for a local area network where everybody knew all other people in the building. How law enforcement is going to oppose efforts to address privacy

flaws with encryption, and to what extent is unclear to me. Obviously they are going to stick their nose in it, there is no doubt about that.

There are a lot of things we need to do to the Internet to make it more private. I think we will end up with a more private Internet, with things like IP address blocking, but there will still be things the police can get to. I think that is going to be the tradeoff.

Question: Are DoubleClick and similar companies involved in chat rooms, or other kinds of targeted activities?

Answer: I am sure there are message boards where DoubleClick ads appear, and if that is the case, they would be in a position to know what chat room it is and what topic it is about. It would have to be a web-based chat room, though. Another example of ad placement for the purpose of figuring out consumer interests is eBay, where they display ads so they can figure out what consumers are interested in buying. This kind of information can go in your online profile. There is, however, a whole list of words they will never profile on, an off-limits dictionary of sensitive words, if you will.

Question: Could you talk a little about how the Internet gives you access to historically offline, and theoretically public, information just by typing in somebody's name?

Answer: This is a really big problem. I haven't looked at it much myself, but it is very interesting that when people talk about privacy, they will say that this data is available in the offline world, so what is the big deal? Your question hit on the problem, which is that when it is all online, it is much easier to access it, aggregate it and put it together. In Brookline, MA, where I am from, you can now get the real estate transaction records from the last 10 years online. It used to be that you would have to go to the library, but now you can just download a spreadsheet. There is a real danger here, and I am not sure what the answer to the problem will be.

The question is whether we will put up with it. I think that like a lot of other things, we will end up putting up with it. The kind of thing that will drive solutions to this problem is when a specific person has a specific problem, like when Judge Bork's video rental records were revealed when he was up for nomination.

It is going to get worse and worse, I guarantee you. People will make a business out of aggregating this stuff.

Question: A lot of what you have spoken about so far deals with the privacy problems from a theoretical perspective and a programming perspective. What are one or two of the worst examples of this technology being used to harm people?

Answer: I get asked this question all the time, but I don't really have an answer to this. I can't come up with somebody who has been harmed by DoubleClick. What I can come up with is something a little bit different, and which is a little harder to talk about. There was a case two or three weeks ago, in which somebody sent out a fake press release about Emulex and how the CEO was going to resign and how they would need to restate results. The stock immediately went down 60% in fifteen minutes. If you were smart, and knew how to play that, you could make a lot of money. It was quickly discovered that it was a fake press release. The question, then, was, "Who did it?"

They knew within two hours where he had sent the e-mail from; it was from a computer in a community college near where he worked. They investigated employees at a company called Internet Wire where the press release originated, and discovered that this guy went to this school. They checked what stock purchases he had made and found out that on Monday, the day of the stock meltdown, he was in Mandalay Bay in Las Vegas buying the stock when it bottomed out. They also found out that he had lost \$100,000 the day before he sent out the fake press release. They got all this information through the tracking mechanisms that I've talked about. He was arrested on Thursday.

The thing is, if you are going to commit a crime on the Internet, you are going to be tracked using things like IP addresses, possibly cookies, or noticing things you do with cookies. Companies like DoubleClick would rather not hear about this aspect of the tracking business. This tracking stuff can hurt people, but it is probably the people we want to hurt, the people who do bad things. Over the long haul, the concern is that data collected about people can be used in legal proceedings, like in a divorce. What if all of your search strings over the last six months ended up in the other lawyer's pants? Has this happened yet? Not that I am aware of. Lawyers are mostly interested in e-mails people get, and things like that.

Question: Is there any connection between online profiling and Echelon?

Answer: Not really, except that Echelon is a variation of online profiling. Echelon is a system of NASA satellite scanners listening in on satellite traffic, like faxes and other kinds of electronic doentation, to see what people are up to.

Question: Aren't online privacy issues a subset of the much larger question of how much data is being collected, how it's kept, who guards it, and whether there is a periodic destruction of it?

Answer: One of the problems we have in the privacy arena is something called Moore's Law. Usually, Moore's Law is not considered a problem. Gordon Moore, a founder of Intel, theorized that the number of transistors that can be put on a chip doubles every eighteen months. If you take a ten-year period, well, you can do the math - computers have gotten a lot better than they were 20 years ago because of Moore's Law. The same thing happens with hard disk drives. Ten years ago, I bought a 340MB hard drive for \$400. Today, I can buy a 20GB drive for \$100.

One of the problems is that it becomes cheaper and cheaper and cheaper to store all of this data that we are collecting. We have built this thing called the Internet, tying in phones and other electronic devices, to feed those hard drives. People are pack rats. Programmers in particular are pack rats. They like to save all of this stuff. That is where the concern is, that we are building much more data about ourselves than we ever did before.

As you said, it is a very large problem. It is not just about the Internet, but about everything we do that becomes computerized. There are a lot more records out there.