Public Statement from Outlier Software
to the Federal Trade Commission

The Information Marketplace:
Merging and Exchanging Consumer Data

Jim Berry, President
Outlier Software, LLC

Separating Static and Surveillance Data Collection

There is a subtle, but significant, difference between looking up a person's address in the telephone book versus following that person home from work. Both yield information on where the person lives, but one is Static Data Collection and the other is surveillance. This distinction is important to note when it comes to privacy, because while computers have allowed for the accumulation and integration of static data from many sources (public records, warranty cards, etc.), the Internet has enabled data collection through methods that can only be defined as surveillance. While there are significant privacy issues with both Static and Surveillance Data Collection, addressing these two methods separately would give the Federal Trade Commission and Congress more flexibility to ensure that the interests of personal privacy and business are both recognized.

Surveillance Data Defined

Businesses need information about their customers in order to make informed decisions about what products to offer and how to present those products in the marketplace. This is not a new phenomenon, but with the advent of the computer and more specifically the Internet, businesses have been given the tools (cookies, clear-gifs, etc.) to collect and utilize more detailed records of consumers' habits than ever before. An individual can walk around a neighborhood mall and shop with near anonymity. Traditional stores may be able to track a purchase through a credit card or check in the physical world, but they do not keep a written log of what items a customer looked at, for how long, and what other stores that customer visited. The Internet has made this kind of tracking not only possible, but a reality.

This type of tracking and "profile-building" is what can be defined as Surveillance Data Collection. The use of information gathering techniques that are practically invisible to the average consumer can be considered nothing other than acts of surveillance. Data of this type is collected indirectly from observing the actions and

behaviors of an individual. While information collected in this manner can be stored in aggregate formats, the initial collection point is always individually identifiable. Surveillance Data Collection in this paper is defined as information collected via non-transparent means that uses individually identifiable data collection points.

This definition does not include data that is collected generically, with no tie directly back to an individual (such as the counting on a web server of the number of specific pages served, raw page referrer data (the last page a visitor was viewing), or browser type information). Surveillance Data Collection is limited to the tagging of an individual (person or computer) in a uniquely identifiable way and using that as a means to observe and record the individual's actions through time. This tagging is currently accomplished through the use of cookies, static IP address tracking, and unique URL identifiers (often related to sign-on sites). Currently, Surveillance Data Collection is primarily confined to the Internet.

Static Data Collection, on the contrary, is utilized by "brick and mortar" as well as Internet businesses. Static data is gathered in a question/answer format, that lets the individual know what data points are being collected, when and by whom. Warranty cards or online purchase forms are two methods of Static Data Collection. The business or organization outlines what information they want in order to perform a service. The individual may choose whether to provide the information correctly or not. This decision affects the ability to provide a product or the quality of the service rendered by the business or organization, but the individual is implicitly being given notice and the choice of exactly what information is being collected. The issues of notice and choice are the primary differences between static and surveillance means of data collection and seriously impact online privacy concerns.

Surveillance's Privacy Shortcomings

Data collection without the express knowledge of the consumer is not a method that instills confidence in the Internet. Information gathered through surveillance techniques also has a greater chance of being incorrect or misinterpreted. If an individual lets a friend use his/her personal computer to research skydiving or spelunking, that information could be tracked and mistakenly assigned to the computer owners "profile", flagging that person as a high-risk individual. Additionally, because surveillance data is sold and exchanged between entities without individual's knowing what was collected and when, there is more of a chance that this information could be taken out of context. These issues are very sensitive, some consumers refrain from visiting some sites (such as health and social issue sites) for fear that they may be tracked and that information used in an unknown manner at a later date. What data is being collected and how it will be used are basics tenets of privacy notice and necessary for an individual to retain any control over their own privacy.

Privacy is a comfort issue. Privacy can be defined as a state of being free from unsanctioned intrusion. It is the ability of an individual to selectively communicate or hold secret that information which is sensitive to that individual. Information in one setting may not be considered highly personal, but in another could be extremely

sensitive. Privacy is up to individual interpretation and is above all things a personal issue. Some people are more private than others and some information might be sensitive to one person though not to another. So why is it that privacy on the Internet is a standard applied by web sites on visitors as a collective, void of individuality or personalization? The consumer trust-gap is created by a loss of control over personal information on the Internet, a loss of privacy.

Surveillance Data Collection is not inherently a violation of privacy, but in current practice it is a direct violation of how privacy is defined. Data collection can only be considered individually permitted if the consumer knows specifically what data is being collected and for what purpose. Without this information, how can a consumer be expected to make an informed choice as to whether or not to participate and allow the collection? This would require that web businesses disclose exactly what information they are collecting, not just the fact that they collect data through cookies or clear-gifs. This disclosure, combined with a legally binding explanation of how the data will be used or distributed, creates proper notice. Surveillance Data Collection by a web site, without disclosure and an explanation of intent, runs contrary to notions of privacy and choice for consumers.

This problem is only compounded by the use of surveillance tracking mechanisms by third party actors across multiple web sites. Notice is not currently served by these third-party actors at the point of data collection. A consumer shopping for a book on an online retailer may be watched through surveillance techniques by not only the retailer who owns the site but from multiple ad services, site partners, and even the retailer's vendors. Currently, most sites privacy policies only refer to the data collected by the site itself. Notice, to be an effective component of consumer Internet privacy, must not only contain information on what data is being collected and how that data is to be used by the main web page, but by all of the other entities collecting data on that web page as well. Notice should be included at all points of data collection by all of the entities collecting the information. Only an informed consumer exercises a real choice when deciding what to protect as private and what to reveal.

Surveillance and P3P

This type and extent of notice is necessary to allow consumers informed choice when collecting data through surveillance techniques, however, this type of notice is exactly what is not effective in reaching customers on the Internet. Consumers on the Internet rarely read long detailed privacy notices, primarily because they are obstacles to the ease and speed of use that makes eCommerce a viable alternative to conventional shopping. The proposed P3P standard works to give notice in a simplified form, but this simplicity lacks most of the details needed to create proper notice in the case of Surveillance Data Collection.

The P3P standard should work very well to provide notice and choice to Static Data Collection. Static data is comprised of information a consumer purposefully transmits to another entity. This information might be transferred for many reasons: to ensure better service, initiate communication, or to complete a purchase transaction. On

the Internet, this could be information submitted on a registration form, purchase agreement or membership sign-up. The key difference between static and surveillance data is that the consumer knows what information is being collected during Static Data Collection. This difference allows for the P3P standard to possibly offer consumers informed choices for Static Data Collection, but not for Surveillance Data Collection.

Surveillance Standards

The technology itself that allows Surveillance Data Collection should not be discarded. It is the ways in which this technology is utilized that must be addressed. Surveillance Data Collection should be held to a higher standard of notice, in order to provide a true choice to an informed consumer. These higher standards should include:

1. Notice at the point of collection for all parties. If more than one party collects surveillance information from a specific page there should be more than one notice. If a banner ad accesses a cookie, notice should be present on or adjacent to the ad location.
2. A detailed description of what data is being collected.
3. Contact information for the organization collecting surveillance data.
4. How the data will be used? This should answer questions like: Will the data be sold or shared and with whom? Compiled into aggregate data? Combined with other data sources?
5. Combining surveillance and static data must be disclosed and detailed.
6. Choice as to whether to allow the Surveillance Data Collection or not.

This last standard of choice would preferably be based on an "Opt-in" model, but an "Opt-out" model might be acceptable if all of the other standards were met. The Opt-out model would be acceptable because an individual would be able to opt-out of any surveillance tracking form the point of collection. These standards would allow for Surveillance Data Collection while requiring that the consumer be better informed and be given a real choice as to that collection. The responsibility to present notice in an easily accessible fashion should be on the collector of surveillance data not on the individual consumer to hunt and find that notice.

Separating Surveillance and Static Privacy Issues

Currently Surveillance Data Collection is mostly confined to the Internet. This should allow the Government and industry to act by setting down baseline standards for utilizing surveillance data collecting methods without impacting traditional business practices. Separating Static and Surveillance Data Collection would allow for the issues and concerns consumers have about cookies and clear-GIFs to be addressed without having to "ban" the underlying technology and having to forgo all of the benefits these technologies add to the Internet. The dynamic and personal nature of the Internet does not necessarily have to suffer as new methods for using these technologies could offer many of the same services without being tied to a unique identity.

This would also allow the Government to address the privacy issues concerning the content, storage and exchange of personal information data without regard for the method of data was collection (static or surveillance). Addressing Surveillance Data

Collection separately would further allow personal data privacy protections to be applied consistently to offline and online organizations. Yes, stricter standards for Surveillance Data Collection would have a cost to implement and new approaches to online personalization might have to be implemented; but consumers and Internet businesses are already paying a high cost due to privacy concerns and will continue to do so, until the trust-gap is closed. Separating the methods of data collection on the Internet will assist government and industry to approach the privacy in a manner that will respect the privacy desires of consumers as well as meet the information needs of business.

Jim Berry, President
Outlier Software, LLC
518 N Tampa St.
Tampa, Fl. 33602
(813)-334-6437
www.Outliersoftware.com
jimberry@OutlierSoftware.com