The Privacy Market

The privacy market has three considerable segments. There is the market of customers who seek privacy in all transactions. There is the market for merchants who wish to distinguish themselves with respect to privacy practice. There is the market for merchants seeking to avoid the risks and costs of excessive data compilations.

The costs of data compilations are both direct and indirect. Direct costs include the costs of designing a web site that includes data compilation features, preparing the data for use, and storing the data for the long term. Data which are seen as free carry some costs, and in some cases significant risk given the international nature of networked commerce.

NetGain argues that to successfully track and observe consumers a company should expect that 20% of total expenditures should be on consumer data compilation and analysis. Thus any web site design that requires that consumers be tracked should budget for this order of expenditure. Of course the bulk of this expenditure is collecting the data, cleaning the data, and storing the data securely. Thus consumer privacy tends to be a cost-effective way to do business for those businesses which do not entirely depend on selling consumer attention span. In contrast, Netscape or other attention-span merchants can reasonably be expected to spend on the order of 20% of overall cost tracking consumers because they are an attention-span merchant.

In collecting data companies expend monies and undertake risk. Companies sometimes undertake risks which are best avoided by collecting information simply because it is easy to collect and affordable to keep. The list of companies who have had privacy debacles illustrates that the practice of examining the reason for data compilations is rare. Recently companies have been creating an executive position, the Chief Privacy Officer, to examine exactly these risks and costs. Yet the role, strategic oversight, and responsibilities of the Chief Privacy Officer remain uncertain. In some companies the CFO reports to the Chief Information Officer or the Chief Technology Officer, in others the CFO reports to the Chief Executive Officer. In few cases is the power of the CPO position commensurate with its responsibilities, given that every consumer data flow would be of interest.

These practices can fairly be termed data pathologies or information pathologies. Similar forces were at work when American companies found comfort in large inventories, only to see the just-in-time manufacturing of Asia cut margins and grab market share. While large data inventories may provide some form of comfort in the world of click commerce, where customers are hard to grab, those data are costly.

Data have cost in three significant areas: collection, cleaning, and storage. First, data are costly to collect. Additional cookies and unnecessary data transfers slow page loads and decrease response time. Processing load on the server increases with the intensity of data surveillance. When using advanced caching networks, such as provided by Akamai, advanced data surveillance can decrease the usefulness of the cache by requiring additional communications. Sites can decrease data compilation and storage when overloaded, but few do so.

After the data have been compiled the second significant cost is in data hygiene. For a database to be useful it must be intermittently cleaned. This requires understanding the value or purpose of the data generated, so cleaning is sometimes simply skipped. The result is that many databases are filled with incorrect, useless and outdated data. As David Banisar of the Electronic Freedom Foundation states, "Databases are filled with garbage. And consumers like it that way." In fact, not only do consumers lie to merchants on such trivial queries as age and sex, the Coalition for Patient Privacy reports that a majority of Americans have lied to their doctors. Thus data pulled from the web must be cleaned to be useful.
The third the cost of data is in storage. In this dimension security again plays a role. A most common privacy and security violation is in making credit card numbers publicly available. This is usually the result of an incorrect setting by the database administrator. Certainly enhanced consumer privacy would reduce this risk, thereby making any privacy-protecting site less of a target for hackers.

Similarly data storage is a clear target for commercial espionage. Any company determined to view the customer profiles of an on-line competitor can do so. The inability of Microsoft to keep the servers on which it stores its source code secure is an argument for the difficulty of computer security. In fact, the more complex burglary of voice mail systems was implemented for the purpose of commercial espionage as early as 1990.

On the day that the Federal Trade Commission announced that Geocities had substantially violated the privacy of its customers the value of Geocities stock plummeted. The value of Geocities fell nearly $1,000,000 for each minute that the stock market remained open. The value of Geocities was in large part the value of its customer relationships. An FTC agreement allowed Geocities to maintain business at the low price of a privacy policy which was both posted and followed. Geocities stated that the settlement would in no way harmed their business practices.

Similarly Intel learned the value of privacy in the release of its 1999 commerce-enhancing user identifier in the Pentium III chip. Intel released a chip with a unique identifier, which could not be disabled by the user. Some three hours after a boycott of Intel was announced by electronic civil liberties groups Intel posted software enabling consumers to remove the identifier.

In both the Geocities and Intel cases the merchant was requesting something of value from the user that was little or no value to the merchant. Intel had no plans to profit form the identifier, it was simply easy to add. Geocities' success did not depend on the success of marketing the secondary information optionally provided by customers. In both cases the merchant took risks to enable increased data compilations which were of little or no use in the daily business of the merchant.

In summary, data are expensive. The collection and compilation of data are expensive. The potential for legal action or customer reaction to errors in privacy calculus can be considerable. Furthermore the Federal Trade Commission continues to investigate companies with respect to the existence of privacy policies and compliance with those policies.

Of course the United States is not the only nation on the Internet, and thus privacy policies across the globe may come into play. In particular, the Europeans have a considerably different philosophical and practical approach to privacy. The Europeans view privacy not as a conceptual right but rather as a set of practical controls on data.

The European Data Directive creates a very real risk for merchants on the Internet. The United States does not have privacy protection adequate for European consumers. In October of 1999 there was a risk that this policy difference would disrupt data flow between the United States and Europe. In order to resolve these differences the European and American negotiators developed safe harbor regulations such that any American business which follows these regulations will not be at legal risk in serving European customers.

European Safe Harbor regulations require that a company have a clear privacy policy and that this privacy policy be strictly followed. However, in any organization doing so is less simple than it appears. Designers may want increased customer data to evaluate a new web site design or implement web maintenance, or they may obtain detailed browsing information in order to
optimize a system. Managers do not always know the details of their privacy policy. If the privacy policy is not implemented in the database, it does not provide a safe harbor.

The conditions on the privacy policy are less than those required by the Directive. The European Safe Harbor provides legal protection for a company in any nation which is a member of the European Union.

Thus the ideal situation is to ensure data protection with an automated mechanism. While European customers can ensure privacy with customer-targeted software it is not incumbent upon the European customer to do so.

The ideal merchant technology would ensure merchants' access to information used for customization (e.g. screen size, previous purchases) while letting consumers evaluate the information.

Privacy protecting technology will ensure that companies do follow privacy policies. It is extremely easy for privacy policies to slip given today's rapid change in personnel and business relationships. A privacy policy implemented with one web site may not be obtained for an upgrade.

The European Union and the Federal Trade Commission have made it clear that it is not incumbent upon the consumer to protect his or her privacy. It is the responsibility of the merchant. Yet the majority of privacy products in the market today are consumer-centered.

This is a rational response to the increased consumer concern for privacy. The percentage of consumers who are concerned about privacy has risen for the past two decades. Consumers are interested in the convenience of Internet commerce in the near term, but customer retention requires an appreciation of the balance between the need for data and the need to avoid the risk which comes with data.

Consumer information management technology comes in two basic forms: concentration of trust and distribution of trust. Under systems that concentrate trust the consumer places all data in that system. Examples of systems which concentrate trust are the Microsoft wallet and the P3P technologies. The P3P or Platform for Privacy Preferences, has users enter all data into a single form. Those data are then stored on the user's machine. The user selects a privacy policy which he or she can accept, and then the data are set automatically in response to a request by any site with a matching privacy policy.

P3P asks the consumer to provide information and trust the machine to evaluate the veracity of selected sites. Electronic civil liberties organizations have argued that this standard decreases, rather than increases, consumer privacy by preventing the consumer from directly evaluating information accuracy, site trustworthiness, and transactional context.

An alternative is to distribute trust. This is most commonly seen in techniques for advanced anonymous publishing (e.g. Publius) or anonymous routing (so-called onion routing). In these cases information is distributed so that cooperation is needed between the parties in order to expose information. Merchants that choose to have private transactions will not see consumer data, and thereby avoid all the risks described above. Merchants that choose the secure alternative (where the billing information is protected but identity is not) are subject to a subset of the risks described here. Distributed trust enables the merchant to complete transactions, target information needed for business, and dispose of all costs and risks associated with information tangential to the transaction.
Privacy is rarely expensive for a merchant. Yet violations of privacy can prove quite costly.

There are issues of legal and market liability. It cannot be determined what the market value of privacy will be, yet it is certain that errors in privacy calculus can be costly. There is a definite need for merchants to limit their downside risk by examining data compilation practices and adopting privacy-enhancing software.