Privacy and Online Data Collection

Călin Gurău
GSCM – Montpellier Business School, France

INTRODUCTION

Online *privacy* represents a controversial subject for Internet users and online companies alike. Most Internet-active enterprises are using cookies or subscription forms to collect demographic and behavioral data about the Internet users that visit their sites. In exchange, these companies are promising the *personalization* of online interaction between company and customer, and therefore better value for clients. In addition to these *benefits,* many firms promise in their *privacy* disclaimer to use the collected data only for purposes specifically accepted by clients.

Studies have shown that most Internet customers are concerned about their online *privacy* (Kim & Montalto, 2002; Kuanchin & Rea, 2004; Malhotra, Sung, & Agarwal, 2004; Sheehan, 2002). They feel that despite the strict *privacy* policies published on the Net by firms, they have no control over the use of their personal data once collected by an online enterprise. In other cases, they fear that companies may use covert data-collecting methods that are not disclosed to the Internet users. According to Westin (2001), consumers' concerns in online *privacy* revolve around "intrusions, manipulation, and discrimination; on special concerns about third parties capturing self-revelations users are making on the Internet; and on concerns about identity theft and stalking though capture of personal information." These negative perceptions highlight the need to approach online *data collection* from an ethical perspective.

This article attempts to identify and analyze the perceptions of online customers related to the benefits and perils of personal data collection on the Internet. Analyzing the primary data collected through a questionnaire survey of 300 UK Internet users, the article presents and discusses the customers' evaluation of the privacy policy applied by the commercial sites most frequently accessed, as well as the level of *personalization* offered by these Web sites. The main sources of perceived *risks* and benefits are identified and analyzed in order to identify their effect on the perception of online customers.

On the basis of these results, the article proposes a graphical model that classifies the commercial Web sites into four main categories, based on the balance between the perceived *risks* and the perceived *benefits* of online *data collection*. This article concludes with practical propositions addressed to commercial Web sites concerning the actions they can take to improve the perception of their customers regarding the benefits of online data collection, and to develop the popularity of their Web sites.

BACKGROUND

The evolution of information technology applications in the last 10 years has opened new possibilities for distant commercial interactions, related with the acquisition and exchange of information, products, or services. The exponential growth of online commercial transactions was related with an intensive competition among virtual enterprises for market shares and customer loyalty. In order to achieve a competitive advantage in the online market, many firms have implemented advanced e-CRM applications, which collect relevant data about online clients, analyze and evaluate the consumer profiles, and identify the higher value customers for the firm (Ragins & Greco, 2003).

One of the prerequisites of an effective e-CRM strategy is the collection of historical data about the interaction between the firm and its customers (Ragins & Greco, 2003). However, the collecting, archiving, and processing of personal data creates significant online *privacy* concerns. For the individual user, the privacy threats fall into two main categories:

1. **Web tracking devices that collect information about the online behavior of the user (e.g., cookies):** A company can use cookies for various valid reasons: security, personalization, marketing, customer service, and so forth. However, there is an important distinction between cookies which are active only within a specific Web site, and

the ones that can track the user's activity across unrelated Web sites. Recently some aggregator networks have deployed hidden 'pixel beacon' technology that allows ad-serving companies to connect unrelated sites and overcome the site-specific nature of traditional cookies (Mabley, 2000). Additionally, some companies are now connecting this aggregated data with offline demographic and credit card data. Eventually, these resulting databases can be used or sold as powerful marketing tools.

2. **The misuse of users' personal information in exchange for specific benefits, including increased personalization, Web group membership, and so forth:** The misuse of personal information includes unsolicited promotional e-mails, the integration of data in databases that can be sold to third parties without the consent of Internet users, or even credit card fraud.

The databases, intelligent agents, and tracking devices are surrounding the Internet users with a web of surveillance, which is often hidden and unknown to the subjects. The surveillance is initiated by the simple act of presence on the Internet. Specialized software applications such as 'cookies' are tracking the online behavior of Internet users and feeding the data into databases, which create and permanently update a profile of online consumers. These profiles are then used for segmenting the market and targeting the most profitable consumers.

Exercising control of information, after it was voluntarily released, presents another critical problem. The misuse of personal information—which can be defined as any use that is not explicitly defined in the company's *privacy* disclaimer or is not approved by the informed customer—covers many possible aspects. The customers' concerns focus on people reading private e-mails, tracking clickstream patterns to learn where people surf, compiling profiles of Net use for marketing purposes, and collecting information about children for marketing purposes without parental consent (Westin, 2001). For example, in 2000, Toysurus.com was subject to intense debate and controversy when it was discovered that shoppers' personal information was transferred through an unmarked Internet channel to a little-known data processing firm for analysis and aggregation. This operation was not disclosed in the company's privacy disclaimer, and therefore online customers were not aware of it.

Another main concern of online customers is the quality and reliability of online privacy policies (OPPs) (Westin, 2001). A series of studies (Freehills, 2000; Lichtenstein, Swatman, & Babu, 2003; NPP, 2000) indicated that many American and Australian OPPs do not comply with recognized ethical principles concerning the use of customers' data, use unclear terms, and are not consistent with the real practices of the online firms that publish them.

Regulators and legislators have addressed the controversial *privacy* issue quite differently across the world (Nakra, 2001). The United States, the largest world's financial and Internet market, has not yet adopted a national, standard-setting online *privacy* law (Jarvis, 2001). U.S. privacy statutes have primarily focused so far on protecting only specific areas of consumer privacy, such as financial data, health information, and children's personal information (Desai, Richard, & Desai, 2003; Frye, 2001; Rombel, 2001). In comparison with the American official opinion that online privacy protection is a matter of voluntary self-regulation by market-driven companies, the Europeans consider that it is more effective to enforce specific legislation regarding this issue.

The current European approach is based on three basic tenets (DTI, 2003):

1. individuals have the right to access any data relating to them and have it kept accurate and up-to-date;
2. data cannot be retained for longer than the purposes for which it was obtained, nor used or disclosed in a matter incompatible with that purpose, and must be kept only for lawful purposes; and
3. those who control data have a special duty of care in relation to the individuals whose data they keep.

Data commissioners oversee these rights in each European country and require most data controllers to register with them to track what information is being collected and where. They are charged also with investigating all complaints from citizens.

These principles have been incorporated in the European Data Directive, which came into effect in 1998, and more recently, in the European Directive on