Privacy Issues on the Internet Caused by Corporate Colonization

Research Paper

Gonzaga University

Christian Miemiec
Gonzaga University
COML509 B2
Alexander Kuskis, PhD
April 30, 2011
Privacy Issues Caused by Corporate Colonization on the Internet

Throughout our studies in COML509 we have seen and learned that privacy and autonomy are important to users of the Internet. Unfortunately, autonomy on the Internet is often taken away from these users by corporations of all kinds. Autonomy on the Internet insinuates independence and the suggested infringement on independence is a topic of growing concern (Thurlow, Lengel, & Tomic, 2004, p. 91). Corporate colonization is the act powerful corporations take, supported by neoliberal policies, to re-create and reinforce (through the Internet) the dominant discourses and practices of consumer capitalism, marginalizing critical communication. The activity corporations take to contravene (often subtly) autonomy is often done to maximize profits. We can also see how privacy is impugned by these same forces. The forces (like media, ISP, and telecom conglomerates) behind corporate colonization create an almost insidious connection between corporations and the Internet: autonomy and privacy are two of its victims.

When looking at privacy (and privacy on the Internet) we suddenly find ourselves with a dilemma. It is a common dilemma when researching social issues. This dilemma is caused by a lack of scope, definition, and peripheral insinuations centered on the term privacy. In order to discuss salient research regarding this term that is legally and personally difficult to define, I will provide a baseline definition of privacy for use in this paper. Even if given a common framework or definition of privacy, we then face a second challenge. The sheer breadth of possible topics relating to privacy on the Internet requires a significant narrowing of scope. Using the provided definition I will present information on some important privacy issues. The issues provided are not meant to be comprehensive but instead representative of common and often hidden issues.

This information will be divided into involuntary and voluntary privacy concerns.
Privacy

When I first entertained this research topic I made a glaring oversight. I initially assumed my working definition of privacy was universal. In subsequent research into Corporate Colonization I found, through feedback, that where I saw privacy concerns others saw something different. McCullagh (2008) states, “Privacy encompasses a variety of different issues and is important for a number of reasons. Therefore, a single definition that adequately incorporates all the subtle differences that privacy evokes has so far proven impossible” (p.4). Much of the research found that deals with privacy on the Internet is prefaced on a common understanding of privacy. Surveys used in these studies ask questions about how people feel about privacy issues but many of these same studies do not define a baseline for what privacy is. This is a problem. To help us avoid this in this paper, I will use a comprehensive and useful framework offered by DeCew (1997) based on a cluster concept of privacy.

DeCew (1997) argues that privacy is, “a broad and multifaceted cluster concept”, which encapsulates “our ability to control information about ourselves, our ability to govern access to ourselves, and our ability to make self-expressive autonomous decisions free from intrusion or control by others” (pp.61-62). DeCew (1997) envisages privacy as a, “complex of three related clusters of claims concerning information about oneself, physical access to oneself, and decision making and activity that provide one with the independence needed to carve out one’s self-identity through self-expression and interpersonal relationships” (p.78). The cluster includes three aspects of privacy: (1) informational privacy; (2) accessibility privacy; and (3) expressive privacy (DeCew, 1997, pp. 75-76). I will use this concept and these aspects of privacy as a basis for the presented areas of concern regarding privacy on the Internet.
Involuntary Internet Privacy Issues

It may be a bit misleading to use the term *involuntary* in the heading of this section. I wanted to use the qualifier *kinda* after the term, but felt that level of assumed familiarity was inappropriate for a research paper. Maybe *inadvertent* is more descriptive. Most involuntary internet privacy issues do have some form of remedy; however, these precautions and utilities are often not employed effectively or at all. As was noted earlier, there are so very many avenues of risk that we cannot completely cover them and all their variants. In this section we will focus on three interrelated areas of concern; data mining, Google, and behavioral advertising.

Data collection and data mining happen continuously focusing on the Internet user. We are notified regularly of site and portal privacy terms and conditions, however, “Online privacy statements are often placed inconveniently at the bottom of the page and are often tedious, complex, and replete with legal language the average web user finds difficult to comprehend” (Fernback & Papacharissi, n.d., p. 3). Additionally, we find, “many of the security statements of e-tailers sound reassuring, but offer very little protection to the individual consumer” (Fernback & Papacharissi, n.d., p. 3). Often these statements are meant to legally protect the company and not the user (Fernback & Papacharissi, n.d., p. 3). The end result is that we give e-commerce and other portals permission to collect information about our preferences and activity without really understanding what we have agreed to or because we don’t believe we have a choice.

With their legal bases covered these online institutions create *digital profiles* of their users and use profile information in conjunction with other usage data for data harvesting (Spring, 2010, p. 11). New businesses that collect information tied to your email address have sprung up to correlate information from social media sites to other profile information. As an
example, privacy experts say that in the near future it is possible your credit card interest rates could go up based on a tweet that you were just laid off (Spring, 2010, p. 12). One survey showed nearly two-thirds (64.9%) of users indicated that they were "very" or "somewhat" worried about their personal privacy and more than 7 in 10 (71.4%) agree "strongly" or "somewhat" that people have lost all control over how personal information about them is used (Kachhi & Link, 2009, p. 77). Yet given this level of concern they continue to provide the inadvertent information cited. We will see in future sections how much information we continue to provide voluntarily.

Google is a juggernaut in the information space and on the Internet. Google combines common practices that other search and email providers use with their own technologies. Due to their ubiquity they warrant their own discussion. As of 2008, Google is estimated to account for nearly 60% of all Internet search queries in the United States—over six billion each month (more than double the next largest search engine). They are estimated to have received 76% of search revenue collected by the top three search engines. Yahoo, its top competitor, received just over 18% (Tene, 2008, p.1434). We will ignore some of the recent news items about Google, such as those regarding location tracking. Even still, there is an unbelievable amount of data tracking that a company like Google does that falls into the category of involuntary. Tene (2008) wrote:

Google’s access to and storage of vast amounts of personal information create a serious privacy problem, one that Princeton computer scientist Edward Felten recently called “perhaps the most difficult privacy [problem] in all of human history.” Every day, millions of users provide Google with unfettered access to their interests, needs, desires, fears, pleasures, and intentions. Many users do not realize that this information is logged and maintained in a form which can facilitate their identification. As John Battelle
memorably put it, “[l]ink by link, click by click, search is building possibly the most
lasting, ponderous, and significant cultural artifact in the history of humankind: the
Database of Intentions.” This “Database of Intentions,” meaning “[t]he aggregate results
of every search ever entered, every result list ever tendered, and every path taken as a
result,” constitutes a honey pot for various actors. (p.1435)

The full impact of this statement can best be brought to light by the following list of
search strings. The following information was provided by AOL and is attributed to an
anonymized user (number 1515830). This information is representative of the same type of
information Google tracks. This list will underscore how much users unintentionally reveal
about themselves when they use search engines:

{\begin{verbatim}
  chai tea calories
  calories in bananas
  aftermath of incest
  how to tell your family you're a victim of incest
  surgical help for depression
  oakland raiders comforter set
  can you adopt after a suicide attempt
  who is not allowed to adopt
  i hate men
  medication to enhance female desire
  jobs in denver colorado
  teaching positions in denver colorado
  how long will the swelling last after my tummy tuck
  divorce laws in ohio
  free remote keyloggers
  baked macaroni and cheese with sour cream
  how to deal with anger
  teaching jobs with the denver school system
  anti psychotic drugs
\end{verbatim}}

(Tene, 2008, pp.1443-1444)

From less than 20 search queries, it is easy to detect information concerning the user’s
health and mental condition, personal status, profession, geographical location, and even favorite
sports team. Now imagine the volume and depth of personal information contained in search-
query logs assembled over thousands and thousands of searches. This is the type of information
Google has stored. Add to this the information Google gains from Gmail, other Google apps (like Chrome), and key tracking of typed information through Google toolbar and you can see that Google stores and retains much more information than most Internet users are aware of.

The culmination of much of the data mining and tracked information practices presented has resulted in *behavioral advertising*. Behavioral advertising, which is one form of targeted advertising, is the practice of collecting data about an individual’s online activities for use in selecting which advertisement to display (McDonald & Cranor, 2010, p. 2). Third party cookies as well as “DART” (Dynamic, Advertising, Reporting, and Targeting) technologies are used to provide the insight needed (Tene, 2008, p.1448). The result is that users see ads that are created by correlating which sites an individual visits, ads clicked, inferences about age range and sex, and approximate physical location based on the computer’s IP address. Advertisers build profiles of that individual’s characteristics and likely interests (McDonald & Cranor, 2010, p. 3). Studies show that only 11% of users understand opt-out cookies even though 86% believe that ads are being generated based on website history and 64% find the idea invasive (McDonald & Cranor, 2010, p. 1).

The paradigm of involuntary privacy issues covered can be summed up as deceive, inveigle, and obfuscate (if you are an X-files fan you may recognize this). Massive amounts of data are collected and correlated by major Internet players, often with confusing means, and this data is used to target the users. The biggest threat to privacy is the data that the Internet players say they are not using and often don’t admit they are using but this data shows so very much about the users. All this is done with most users having a tacit understanding they have agreed to give up some privacy. They just do not realize how much.
Voluntary Internet Privacy Issues

In this next section we will look at a few privacy issues that Internet users voluntarily provide large amounts of personal data. It appears most people don’t think about the quasi-permanence of information they provide to the world. These same people do not think about the compendium of information they provide over time. Unfortunately, most people don’t think like criminals but criminals do. We will look at some privacy issues that arise from the use of Facebook, blogs, and professional sites like Monster.

Like Google is to online searches, Facebook is the 800 pound gorilla of the social networking world. There are over 500 million Facebook users. There is much discussion regarding Facebook’s privacy settings and practices, but we will focus on how users cause their own privacy issues.

Based on surveys, we find that of users under the age of 30, 76.3% of profiles were visible by friends only (Taraszow, Aristodemou, Shitta, Laouris, & Arsoy, 2010, p. 89). However, 64.1% provide their email address, 54.2% provide their hometown, and 99.2% provide their full birthday in their Facebook profiles (Taraszow, Aristodemou, Shitta, Laouris, & Arsoy, 2010, p. 91). Additionally, 89.3% of users provide a profile picture where you can determine gender and approximate age combined with 96.2% providing their real names (Taraszow, Aristodemou, Shitta, Laouris, & Arsoy, 2010, p. 90). Users do this because they enjoy the Facebook experience. Debatin, Lovejoy, Horn, and Hughes (2009) state, “It can therefore be assumed that the expected gratification motivates the users to provide and frequently update very specific personal data that most of them would immediately refuse to reveal in other contexts, such as a telephone survey” (p.88).
Considering the vast amount of data mining already being done it is easy to see how any access to the information provided on Facebook could be of significant malicious use. Debatin, Lovejoy, Horn, and Hughes (2009) highlight this aspect:

The IT security firm Sophos set up a fake profile to determine how easy it would be to data-mine Facebook for the purpose of identity theft. They found that out of 200 contacted people, 41 percent revealed personal information by either responding to the contact (and thus making their profile temporarily accessible) or immediately befriending the fake persona. The divulged information was enough “to create phishing e-mails or malware specifically targeted at individual users or businesses, to guess users’ passwords, impersonate them, or even stalk them.” (p.87)

Another possible impact to a user’s privacy is that law enforcement is free to use public information that people provide as well. In a rather benign example:

A police officer resorted to searching Facebook after witnessing a case of public urination outside a fraternity house at University of Illinois at Urbana-Champaign and the only other witness on the scene claimed not to know the lawbreaker. Once on Facebook, the officer searched the man’s friend list and the lawbreaker he was looking for. (Debatin, Lovejoy, Horn, & Hughes, 2009, p. 85)

As can be seen, of all of the privacy concerns people have regarding Facebook, the most significant one they should be aware of is their own disclosure of information. Figure 1 (Debatin, Lovejoy, Horn, & Hughes, 2009, p. 88) best summarizes the Facebook privacy paradigm:
Blogs represent another area where users provide copious amount of personal information. McCullagh (2008) states, “By their very nature, blogs raise a number of privacy issues as they are easy to produce and disseminate, resulting in large amounts of sometimes personal information being broadcast across the Internet in a persistent and cumulative manner” (p.3). We also find that of a survey of 1258 bloggers, 62.6% stated their primary reason for having a blog was to document their personal experiences and share them with others (McCullagh, 2008, p. 9). Because the primary reason for blogging is personal many do not use the anonymous option. However, when asked whether they do anything to limit who gets to read what they post, 72.3% of respondents said no (McCullagh, 2008, p. 10).

The area of voluntary privacy concern we will look at is again one caused by users providing their own information. Many people use professional sites to look for jobs, create a network of contacts, and to learn about industry standard benefits. A commonly used site to look for possible employment is Monster. Sites like Monster provide the ability for users to create public (searchable by subscribers), confidential (name and references removed but searchable),
and private (usable only by user) resumes (Monster, 2010). However, the confidential option is the least used variant.

When we look at the voluntary privacy concerns discussed we find a common thread. The most danger, from a privacy perspective, comes from the user.

**Conclusion**

There are many more Internet privacy issues than were covered here. In considering Internet privacy in a corporate managed and colonized sphere it is always important to define privacy for yourself. I used the framework defined by DeCew (1997) which speaks of privacy in three ways:

1. **Informational privacy** – control over one’s information
2. **Accessibility privacy** – control over physical access to information
3. **Expressive privacy** – control over how to choose, interact and act

From the research presented we find that informational and accessibility privacies are of primary concern. Data on users is collected and correlated from multiple locations providing an almost eerie profile of the individual. It also is evident that based on how users choose to wield their expressive privacy will have a major impact on their informational privacy. Being aware of Internet privacy concerns is the best way to make informed decisions and to protect yourself.
References


