KNOWPRIVACY

PROJECT MEMBERS: JOSHUA GOMEZ, TRAVIS PINNICK, ASHKAN SOLTANI

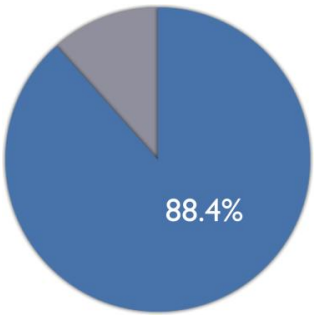
PROJECT ADVISER: BRIAN CARVER

EXECUTIVE SUMMARY

Online privacy and behavioral profiling are of growing concern among both consumers and government officials. In this report, we examine both the data handling practices of popular websites and the concerns of consumers in an effort to identify problematic practices. We conclude by offering potential solutions to realign privacy practices with consumers' expectations.

The data for this report were pulled from six domains, three regarding actual website practices and three regarding user expectations. We analyzed the policies of the 50 most visited websites to better understand disclosures about the types of data collected about users, how that information is used, and with whom it is shared. We also looked at specific practices such as sharing information with affiliates and third-party tracking. To understand user concerns and knowledge of data collection we looked at surveys and polls conducted by previous privacy researchers. We looked at records of complaints and inquiries filed with privacy watchdog organizations such as the Privacy Rights Clearinghouse (PRC), The California Office of Privacy Protection (COPP), and TRUSTe. Through several Freedom of Information Act requests, we also received records of complaints directly from the Federal Trade Commission (FTC). Finally, to gain some insight into what aspects of data collection users are being made aware of, we looked at news articles from three major newspapers for topics related to Internet privacy.

Our analysis of web bugs revealed that they are ubiquitous on the web; this is troubling because users are less likely to know of web bugs, and effective controls for this tracking technology are lacking. All of the top 50 websites contained at least one web bug at some point in a one month time period. Some had as many as 100. Of greater note was the depth of coverage that some tracking companies have. Several of the tracking companies had a web bug on the majority of the top 100 sites. Google in particular had extensive coverage. It had a web bug on 92 of the top 100 sites, and on 88% of the total domains reported in the data set of almost 400,000 unique domains.



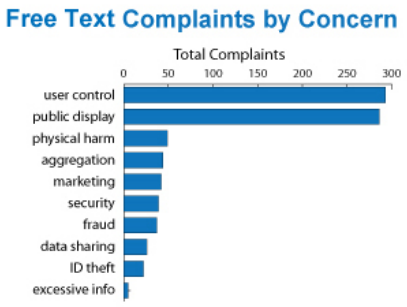
Combined coverage of Google trackers, March 2009. 348,059 out of 393,829 unique domains reported. March 2009

Our survey of privacy policies revealed that most of the top 50 websites collect information about users and use it for customized advertising. Beyond that, however, most contained unclear statements (or lacked any statement) about data retention, purchase of data about users from other sources, or the fate of user data in the event of a company merger or bankruptcy.

Sharing of information presents particular problems. While most policies stated that information would not be shared with third parties, many of these sites allowed third-party tracking through web bugs. We believe that this practice contravenes users' expectations; it makes little sense to disclaim formal information sharing, but allow functionally equivalent tracking with third parties.

Most policies also stated that information could be shared with affiliates. Thus, we requested lists of affiliates with whom data are shared from website operators, but received none. To get a general impression of the number of companies each website in the top 50 could potentially share data with under its current privacy policy, we looked up the parent company of each website and the list of subsidiaries those companies own. The average number of subsidiaries was 297, with a median of 93. This is a second key disconnect between user expectations and website practices. Users do not know and cannot learn the full range of affiliates with which websites may share information.

From the previous surveys and polls we found several consistent conclusions: 1) users are concerned about websites collecting information about them and using it for behavioral advertising, 2) users desire control over the collection and use of information about them, and 3) users lack knowledge and understanding about data collection practices and policies.



Categorization of sample of complaints about Internet privacy made to the FTC and PRC, 2004-2008.

The survey results were reflected in our analysis of consumer complaints. Qualitative analysis of the complaints suggested that users were mostly concerned with a lack of control over data collection and public display of personal information. Indeed, nearly 40% of the complaints in the data set from the FTC were concerned with opt-out control. Data from the FTC, PRC and COPP all contained a significant number of complaints about data broker portals, particularly ZabaSearch, a free search engine that provides data from public records. Data from TRUSTe showed growing concern with unauthorized use of personal data, particularly to set up public profiles on websites. While the FTC has framed online privacy issues in terms of "harm," consumers' complaints focus on lack of control over personal information.

Analysis of news articles suggests that users are made aware of behavioral profiling and other data collection topics in general. However, discussion of some practices was non-existent. Almost no mention was made of the practice of sharing data with affiliates or of the use of web bugs.

Finally, even the act of complaining about privacy is frustrated by a lack of clarity. Consumers do not know where to complain, in part because privacy policies do not provide clear information about remedies.

Based on our findings we recommend that website operators and third-party trackers disclose to users all the information that has been collected about them and with whom it has been shared.

Given the size of affiliate networks, we recommend that users be allowed to choose whether or not websites can share information about them with corporate affiliates.

Given the invisible nature of third-party tracking, we recommend that browser developers provide functionality that makes users aware of web bugs.

We recommend that all privacy policies include a link to the FTC's online complaint form.

We recommend that website privacy policies be written in a more readable form without contradictory statements about third-party sharing. Websites that allow third-party tracking cannot fairly state that they do not share information with third parties.