Stanford Technology Law Review

Cookie Viewers and the Undermining of Data-Mining: A Critical Review of the DoubleClick Settlement

TAL ZARSKY*

CITE AS: 2002 STAN. TECH. L. REV. P1
URL: http://stlr.stanford.edu/STLR/Events

¶1 In the ongoing search for solutions to growing concerns over information privacy, we have added another useless remedy to our arsenal, courtesy of the most recent DoubleClick settlement.

¶2 August 26, 2002 marked another milestone in the never-ending DoubleClick saga. DoubleClick Inc. collects and analyzes data pertaining to Internet users' online activities, utilizing this information to strategically place online advertisements on users' screens. It takes pride in "helping marketers deliver the right message, to the right person, at the right time."[1] Opponents protested that DoubleClick was invading the privacy of Internet users by wrongfully collecting and using their personal information. After surviving an FTC investigation and settling a class action suit, DoubleClick reached a settlement with a group of State Attorneys General,[2] thus concluding an inquiry into DoubleClick's collection and analysis of browsing data.

¶3 This agreement included several predictable undertakings by DoubleClick: notice requirements, restrictions on the use of sensitive information, and opportunities for users to opt-out. Yet this agreement also introduced a novel requirement: DoubleClick must develop and implement a *cookie viewer*. The viewer "will allow a consumer to see the profiled interest categories DoubleClick uses in selecting advertisements for that consumer."[3] This is a new concept, ("a first in the e-commerce world"[4]) and therefore will require the development of special software. As the DoubleClick case is leading the way in the field of information privacy, other players in the e-commerce arena may follow suit and implement such applications in their own websites

¶4 A cookie viewer application would have been appropriate in the previous world of *niche marketing* which made use of pre-determined categories to sort potential customers. However, it loses all relevance in view of the advanced data mining applications now available. Proposed legal solutions, (once again), stagger behind the advance of technology.

¶5 The original concepts of niche-marketing are predicated upon the initial analysis of vast amounts of personal information and its classification into several groups (which are usually provided with a descriptive name), each consisting of individuals who share an affinity regarding certain statistical factors.[5] Thereafter, advertisers use this taxonomy to focus on a specific class when constructing their marketing schemes. These practices are implemented in the Internet environment as well; DoubleClick (and others) sort Internet users into these pre-defined categories by analyzing their surfing habits and target each class with specifically tailored advertising banners.

---

* The author is a J.S.D. candidate at Columbia Law School and wishes to thank Professor Eben Moglen for his assistance, advice and encouragement. For a more detailed analysis of data mining applications and their affects on the use of personal information, please refer to the author's upcoming article: Tal Zarsky, *"Mine Your Own Business!" - Making the Case for the Implications of the Data Mining of Personal Information in the Forum of Public Opinion*, 5 YALE J.L. & TECH. (forthcoming 2003).

Copyright © 2002 Stanford Technology Law Review. *All Rights Reserved.*

¶6 When considering these practices, the cookie viewer is a potentially powerful tool and a step in the right direction, as it informs web-users of the pre-defined categories in which they are placed. Thus, the user will have some insight to the process of direct marketing and supposedly will be less prone to manipulation.[6] Yet these practices of niche marketing are now outdated. The emergence of *data mining* applications introduces a new and improved business model, one that leaves no place for pre-defined categories, or cookie viewers.

¶7 Data mining applications bring the analysis of personal information to a higher level by allowing it to commence with minimal guidance from the analyst, who is not required to generate an initial hypothesis. Niche-marketing schemes are now powered by *clustering* applications, which are the data-mining response to the classification task. With clustering, the entire database is broken down into several clusters of information, each one sharing an internal statistical affinity. This process is not generated by a hypothesis originating in the minds of data analysts, but stems solely from the attributes of the data itself. Thus, the clusters formed are not required to conform to the taxonomy used to divide the population in the past, or answer to catchy names such as "Pools and Patios" or "Shotguns and Pickups." The borders of such classes might be elaborate, not easily defined by a simple category name, and ever-changing.

¶8 DoubleClick (as well as other advertisers and marketers) will no doubt change its business practices in view of these enhanced abilities. Instead of targeting advertisement campaigns using pre-defined niches, it will constantly re-cluster and re-classify its database, splitting its pool of users into various groups differently for every task. Thereafter, a specific cluster will be sought and targeted for each campaign according to the attributes of the relevant product. For example, a company providing exclusive cruises will seek a cluster that includes the highest number of those who showed interest in exclusive travel by visiting related websites or clicking on similar ads. DoubleClick will locate this cluster and target advertisements to all its members.

¶9 In the world of clustering, the cookie viewer cannot offer any meaningful insights. With clustering, an individual is placed in a nameless class with flexible and elusive borders. Since there is no *ex ante* hypothesis generating the categorization, there are no clearly defined geographic or demographic boundaries to the group, (such as "suburban and over 50") only a list of statistical variables. At most, the viewer will merely inform the user that he was selected to receive a sailing-related advertisement because he belongs to the group that includes the greatest number of individuals interested in sailing or another water sport. Do we really need a sophisticated software item to figure that out?

¶10 Regulating the use of personal information is a complicated and complex objective. In the DoubleClick settlement, both parties made an admirable effort to come up with an agreement that embodies a balance between public concerns and business interests. Yet both parties failed to notice how recent technological advancements have altered existing business plans and models, thus undermining the proposed solution. When contemplating regulatory schemes regarding the use and analysis of personal information, we must always be wary of the potential *undermining of data mining*!

---

[1] *See* DoubleClick Inc. Form 10-K for the Fiscal Year Ending Dec. 31, 2000 (filed March 12, 2001).

[2] In the Matter of DoubleClick Inc., Agreement between The Attorneys General of the States of Arizona, California, Connecticut, Massachusetts, Michigan, New Jersey, New Mexico, New York, Vermont, and Washington and DoubleClick Inc., *available at* http://www.oag.state.ny.us/press/2002/aug/aug26a_02_attach.pdf (last visited Dec. 3, 2002).

[3] Press Release, Office of New York State Attorney General Eliot Spitzer, Major Online Advertiser Agrees to Privacy Standards for Online Tracking: Company to Increase Visibility and Verify Data Collection Practices, *at* http://www.oag.state.ny.us/press/2002/aug/aug26a_02.html (Aug. 26, 2002).

[4] *Id.*

[5] *See generally* DAVID SHENK, DATA SMOG 114 (1997).

[6] *See* Donna Goodison, *Web Ad Service Ends Privacy Flak; DoubleClick Settlement Changes Policy*, BOSTON HERALD, Aug. 27, 2002, at 33 (quoting Jonathan Zittrain, director of the Berkman Center for Internet and Society at Harvard Law School, who praised the cookie viewer as "a technology-based solution rather than changing the text on a privacy policy that no one reads").

Copyright © 2002 Stanford Technology Law Review. *All Rights Reserved.*