



UNITED STATES OF AMERICA
FEDERAL TRADE COMMISSION
WASHINGTON, D.C. 20580

Maneesha Mithal
Associate Director
Division of Privacy & Identity Protection

Direct Dial: 202.326.2771
Fax : 202.326.3062
E-mail: mmithal@ftc.gov

March 12, 2010

BY E-MAIL & FEDERAL EXPRESS

Reed Freeman
Morrison & Foerster LLP
2000 Pennsylvania Ave., NW
Washington, DC 20006

Dear Mr. Freeman:

On October 13, 2009, staff from the FTC's Division of Privacy and Identity Protection contacted your client, Netflix, Inc. ("Netflix"), regarding the privacy implications of Netflix's planned release of customer movie viewing data in connection with the company's efforts to improve its movie recommendation algorithm. Specifically, staff expressed concern that, despite Netflix's efforts to "anonymize" the customer data prior to its release, it would be possible to re-identify specific customers and thereby associate them with their movie viewing histories and preferences.

Staff's concerns about Netflix's planned release stemmed from research published after the company released a similar data set on October 2, 2006. According to news reports and Netflix's website, Netflix released the first data set as part of its Netflix Prize 1 contest ("Prize 1"), through which researchers competed to improve the algorithm Netflix uses to recommend movies to its subscribers. Netflix's algorithm takes into account past viewing habits and movie preferences of each of its subscribers. The Prize 1 data set represented the movies rated by over 480,000 Netflix customers and the date each rating was given. In an apparent effort by Netflix to anonymize the data, the company replaced customers' names with unique numbers and did not include addresses, phone numbers, or other direct identifiers.

Following the conclusion of Prize 1, two researchers at the University of Texas, Arvind Narayanan and Vitaly Shmatikov, published a research paper demonstrating that it is possible to re-identify particular individuals within the Prize 1 data set using a minimal amount of outside information.¹ With this minimal information, one could determine all of the Netflix movies that

¹ *Robust De-Anonymization of Large Sparse Datasets (How to Break the Anonymity of the Netflix Prize Dataset)*, 2008 IEEE SYMP. ON SECURITY AND PRIVACY 111 (Feb. 5, 2008). Specifically, the researchers used publicly-available movie reviews posted by Netflix subscribers on

a subscriber had rated for a given period of time.

Notwithstanding the research and its attendant publicity,² Netflix announced on August 6, 2009, its intention to create a second contest, Netflix Prize 2 (“Prize 2”). Prize 2 would involve the release of a data set based on Netflix customers’ movie viewing habits and preferences. In addition, the Prize 2 data set would contain certain demographic data about Netflix customers.³

In light of the Narayanan and Shmatikov research, Netflix’s intention to release a second data set – one containing a richer portfolio of consumer information – raised serious concerns about the risk that Netflix’s customers would be re-identified and associated with their potentially sensitive movie viewing histories and preferences. Due to advances in technology that allow for vast amounts of data to be collected, stored, accessed, and combined, staff encourages companies to be cautious when releasing data presumed to be “anonymous” or “not personally identifiable,” especially when those representations are made to consumers.

Consequently, in a letter to Netflix dated November 3, 2009, staff identified a number of concerns related to Prize 2. These include the risk of re-identification and the extent to which Netflix’s previous representations to its customers about disclosure of their information would raise concerns under Section 5 of the FTC Act. Following a number of productive discussions between staff and Netflix, the company sent a letter to staff stating that it intended to suspend plans for Prize 2. Further, Netflix stated that if it releases a second data set in the future, it would not do so publicly; rather, it would release such data only to researchers who contractually agree to specific limitations on its use. In addition, Netflix stated that it would implement a number of operational safeguards to prevent the data from being used to re-identify consumers. Finally, Netflix agreed to further discussion with FTC staff prior to releasing the data.

the popular site www.imdb.com to re-identify individuals in the data set. We note that, although customers may choose to publicly rate some of the movies they have seen, they may wish to keep the remainder of their movie viewing habits private.

² See *id.*; Bruce Schneier, *Why ‘Anonymous’ Data Sometimes Isn’t*, WIRED, Dec. 13, 2007, available at http://www.wired.com/politics/security/commentary/securitymatters/2007/12/securitymatters_1213; Robert Lemos, *Researchers Reverse Netflix Anonymization*, SECURITYFOCUS, Dec. 4, 2007, available at <http://www.securityfocus.com/news/11497>; see also Natasha Singer, *When 2+2 Equals a Privacy Question*, N.Y. TIMES, Oct. 17, 2009, available at <http://www.nytimes.com/2009/10/18/business/18stream.html>; Paul Ohm, *Broken Promises of Privacy: Responding to the Surprise Failure of Anonymization*, Aug. 13, 2009, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1450006.

³ See Steve Lohr, *Netflix Awards \$1 Million Prize and Starts a New Contest*, N.Y. TIMES, Sept. 21, 2009, available at <http://bits.blogs.nytimes.com/2009/09/21/netflix-awards-1-million-prize-and-starts-a-new-contest/>.

Based upon these assurances, staff has determined to close the investigation. The company's swift response and willingness to take steps to protect consumer data represents a meaningful commitment to protecting the privacy of consumers. We encourage Netflix to maintain this commitment as it develops future policies and practices involving consumer data.

The closing of this investigation is not to be construed as a determination that a violation may not have occurred, just as the pendency of an investigation should not be construed as a determination that a violation has occurred. The Commission reserves the right to take such further action as the public interest may require.

Sincerely,



Maneesha Mithal
Associate Director
Division of Privacy and Identity Protection
Federal Trade Commission

cc: David A. Hyman, General Counsel for Netflix