

Threat Level

Privacy, Crime and Security Online

privacy

Like 174 18

0

Share 5

Netflix Spilled Your *Brokeback Mountain* Secret, Lawsuit Claims

By [Ryan Singel](#)

12.17.09

4:29 PM

Follow [@rsingel](#)



An in-the-closet lesbian mother is suing Netflix for privacy invasion, alleging the movie rental company made it possible for her to be outed when it disclosed insufficiently anonymous information about nearly half-a-million customers as part of its \$1 million contest to improve its recommendation system.

The [suit known as Doe v. Netflix](#) (.pdf) was filed in federal court in California on Thursday, alleging that Netflix violated fair-trade laws and a federal privacy law protecting video rental records, when it launched its popular contest in September 2006.

The suit seeks more than \$2,500 in damages for each of more than 2 million Netflix customers. In order to get a better movie recommendation algorithm, the online DVD rental company gave more than 50,000 [Netflix Prize](#) contestants two massive datasets. The first included 100 million movie ratings, along with the date of the rating, a unique ID number for the subscriber, and the movie info. Based on this data from 480,000 customers, contestants had to come up with a recommendation algorithm that could predict 10 percent better than Netflix how those same

subscribers rated other movies.

But video records count among the most privacy protected records in the U.S. — a reaction to a reporter getting Supreme Court–nominee Robert Bork’s records from a video store. The lead attorney on the new suit, Joseph Malley, recently reached [a multimillion-dollar settlement with Facebook over its failed Beacon program](#), which drew fire in part for sharing users’ Blockbuster rentals with their friends.

Thursday’s suit argues that the information is personal data protected by Netflix’s privacy policy, and that NetFlix should have known that people would be able to identify users based on that data alone. In fact, just two months before NetFlix launched the contest, AOL released “anonymized” search-engine logs, which reporters quickly used to track down real people. So it wasn’t surprising that just weeks after the contest began, two University of Texas researchers — Arvind Narayanan and Vitaly Shmatikov — identified several NetFlix users by comparing their “anonymous” reviews in the Netflix data to ones posted on the Internet Movie Database website. Revelations included identifying their political leanings and sexual orientation. The complaint calls that the *Brokeback Mountain* factor, arguing that marketers will suck up the data, combine it with other data sets and start pigeon-holing people into marketing categories, based on assumptions about the movies they rated.

[M]ovie and rating data contains information of a more highly personal and sensitive nature. The member’s movie data exposes a Netflix member’s personal interest and/or struggles with various highly personal issues, including sexuality, mental illness, recovery from alcoholism, and victimization from incest, physical abuse, domestic violence, adultery, and rape.

The Plaintiffs’ and class members’ movie data and ratings, which were released without authorization or consent, have now become a permanent, public record on the Internet, free to be manipulated and exposed at the whim of those who have the Database.

That’s why the lesbian mom joined the lawsuit as a Jane Doe, according to the complaint, since she believes that “were her sexual orientation public knowledge, it would negatively affect her ability to pursue her livelihood and support her family and would hinder her and her children’s ability to live peaceful lives.”

The contest ended this summer when two different teams passed the 10 percent improvement mark, with the [prize money going to a team led by AT&T researchers](#).

The suit is also asking the court to stop Netflix from launching its promised [second contest](#) to improve the recommendations — this time giving out user data that includes ZIP codes, ages and gender, along with movie ratings and ID numbers substituted for user names.

That’s a foolish idea on Netflix’s part, according to [University of Colorado law professor Paul Ohm](#), who in a blog post in September called the idea “a privacy blunder that could cost millions of dollars in fines and civil damages.” Ohm, a former Justice Department lawyer, recently authored a legal paper calling into question the practice of anonymizing data, essentially finding that [if data is useful to researchers, it could also, by definition, be re-identified](#).

For instance, if a data set reveals a person’s ZIP code, birthdate and gender, there’s an 87 percent chance that the [person can be uniquely identified](#).

Ohm did not however support a lawsuit against Netflix for the original contest, arguing the company made good faith efforts to hide identities, using a data-obfuscation technique called perturbation.

A Netflix spokesman said the company could not comment since it had not yet seen the suit. Photo: Still from the 2005 film *Brokeback Mountain*, courtesy Paramount Pictures.

See Also:

[Medical Records: Stored in the Cloud, Sold on the Open Market](#)

[How the Netflix Prize Was Won](#)

[AT&T Will Donate Share of \\$1 Million Netflix Prize to Science](#)

[Lawsuit Accuses Facebook of Conspiring to Break Video-Privacy Law](#)