# How to Regulate Harmful Inferences

**Author :** Paul Ohm

**Date :** December 22, 2021

Alicia Solow-Niederman, *Information Privacy and the Inference Economy* (Sept. 10, 2021), available at SSRN.

A decade ago, Charles Duhigg wrote a story for the New York Times that still resonates today, revealing that Target could predict its customers' pregnancies and delivery dates from changes in their shopping habits. This and similar revelations pose a difficult question: how do we protect vulnerable people from the power of inferences? At the time, I wondered aloud whether we ought to regulate harmful data-driven inferences and how we would do it, which sparked characteristically overheated responses from the libertarian punditry.

A decade on, the ceaseless progress of machine learning (ML) has exacerbated these problems, as advances in the state-of-the-art of prediction make Target's old algorithm seem like child's play. ML techniques have become more accessible and more powerful, fueled by advances in algorithms, improvements in hardware, and the collection and distribution of massive datasets chronicling aspects of people's lives we have never before been able to scrutinize or study. Today, obscure startups can build powerful ML models to predict the behavior and reveal the secrets of millions of people.

This important draft by Alicia Solow-Niederman argues that information privacy law is unequipped to deal with the increasing and sometimes-harmful power of ML-fueled inference. The laws and regulations on the books, with their focus on user control and notice-and-choice, say very little about the harmful inferences of companies like Clearview AI, which notoriously scraped millions of photos from Facebook, LinkedIn, and Venmo, using them as ML training data to build a powerful facial-recognition service it sells exclusively to law enforcement agencies. Unlike Target, which had a contractual relationship with its customers and gathered the data for its algorithm itself, Clearview AI had no connection to the individuals it identified, suggesting that protections cannot lie in laws focused primarily on user consent and control.

The first very useful contribution of this article is its important summary of recent advances in ML, how they raise the possibility of harmful inferences, and how they challenge outdated privacy laws built upon notice-and-choice. This makes Part II of the article an accessible primer on a decade's worth of ML advances for the non-technical privacy expert.

Solow-Niederman's most important move, in Part IV of the article, is to ask us to focus on actors beyond the dyad of provider and user. Like Salome Viljoen's magisterial work on Democratic Data (previously reviewed in these pages), Solow-Niederman deploys geometry. Where Viljoen added the horizontal dimension of people outside the vertical user/service relationship, Solow-Niederman asks us to move beyond the "linear" to the "triangular." She urges us to look outside the GDPR-style relationship between data subject and data controller, to consider the actions of so-called "information processors." These are companies like Clearview that amass massive data sets about millions of individuals to train machine learning models to infer the secrets and predict the habits *not just of those people but also of others*. We cannot protect privacy, Solow-Niederman argues, unless we develop new governance approaches for these actors.

This move — relational and geometric — leads her to focus on actors and relationships that get short shrift in other work. If we worry about the power of inference to harm groups and individuals, we need to scrutinize that which gives power to inference, she argues. Solow-Niederman focuses, for example, on how information processors amass "compute": the computer-processing infrastructure needed to harness massive data sets. She provocatively suggests that

regulators might cast extra scrutiny on mergers and acquisitions that lead companies to increase compute power, citing for inspiration the work of now-FTC-Chair Lina Khan, who has argued for similar shifts in antitrust law.

The triangular view also focuses attention on how companies like Clearview obtain data. Other commentators have been loath to focus on Clearview's scraping as the source of the problem, because many tend to be wary of aggressive anti-scraping restrictions, such as expansive interpretations of the Computer Fraud and Abuse Act (CFAA). Solow-Niederman suggests, contrary to the conventional wisdom, that the CFAA could have been useful in thwarting Clearview AI, had Facebook detected the massive scraping operation, asserted its Terms of Service, and sued under the CFAA. She even suggests FTC action against companies that purport to prohibit scraping yet fail to detect or stop scrapers.

These are two genuinely novel, even counter-intuitive, prescriptions that flow directly from Solow-Niederman's triangular intervention. They suggest the power of the approach, and we would be well-advised to see how it might lead us to other prescriptions we might be missing due to our linear mindsets.

To be clear, as I learned a decade ago, protecting people from the power of inference will raise difficult and important questions about the thin line between intellectual exploration and harm production. Inference can be harm, Solow-Niederman suggests, but she acknowledges that inference can also be science. Preventing the former while permitting the latter is a challenging undertaking, and this article defers to later work some of the difficult questions this differentiation will raise. But by focusing attention and energy on the ever-growing power of ML inference, by compellingly exploring how conventional information privacy law and scholarship cannot rise to the challenge of these questions, and by suggesting new means for considering and addressing inferential harm, Solow-Niederman makes an important and overdue contribution.

Cite as: Paul Ohm, *How to Regulate Harmful Inferences*, JOTWELL (December 22, 2021) (reviewing Alicia Solow-Niederman, *Information Privacy and the Inference Economy* (Sept. 10, 2021), available at SSRN), https://cyber.jotwell.com/how-to-regulate-harmful-inferences/.