Avoiding Bias in Big Data

Article By:

Vito Petretti

Katherine B. O'Keefe

As the use of big data in almost every area of business continues to grow, companies need to carefully consider their own use. We've previously discussed data analytics and use restrictions in a three-part series (Part 1, Part 2, and Part 3), and this post will highlight another issue that can arise when using big data: discriminatory outcomes. The Federal Trade Commission (FTC) recently issued the report Big Data: A Tool for Inclusion or Exclusion? Understanding the Issues, which focuses on avoiding harmful bias in big data and certain consumer protection laws applicable to big data. The report gathered information and commentary from participants in a workshop that the FTC held last year and supplemented that information with additional research.

The report notes that companies' use of big data has led to some major benefits, including increased educational attainment and access to employment for certain individuals, better access to credit using nontraditional methods, and healthcare tailored to individual patients' characteristics.

The report, however, lists three main concerns with big data. First, as with all data, the quality of the data used in big data analytics, including its accuracy and completeness, is important. Second, big data is a great tool for showing correlations, but the FTC is quick to point out a well-known truth that correlation is not causation. Last, big data may be used to categorize consumers in a way that would lead to exclusion of certain populations from certain benefits. In particular, big data use may result in

- companies mistakenly denying individuals opportunities based on the actions of others,

- the creation or reinforcement of existing disparities,

- the exposure of sensitive information,

- the inadvertent targeting of vulnerable consumers for fraud,

- the creation of new justifications for exclusion, and

- higher-priced goods and services for lower-income communities.

The report outlines several laws that may apply to big data practices. These laws include the Fair Credit Reporting Act, federal equal opportunities laws (the report specifically lists the Equal Credit Opportunity Act, Title VII of the Civil Rights Act of 1964, the Americans with Disabilities Act, the Age Discrimination in Employment Act, the Fair Housing Act, and the Genetic Information Nondiscrimination Act), and Section 5 of the Federal Trade Commission Act.

As to how one of these laws might apply to big data, consider the following example: A company asks a consumer for his or her zip code and social media behavior on an application, and then sends the application to a data analytics firm after removing the consumer's identifying information. The firm analyzes the creditworthiness of people in the same zip code with the same social media behavior and provides an analysis of such information to the company, knowing that the analysis will be used for a credit decision. If the company uses the consumer's information to create an analysis of a group that shares certain characteristics with the consumer and then ultimately uses that analysis to make a decision about the consumer, the FTC may very well view such analysis as a consumer report, and the Federal Credit Reporting Act would apply. However, if the company just used the same analysis to inform its general policies, it is unlikely that the FTC would argue that the Federal Credit Reporting Act would apply, because the consumer report was not used for a particular consumer. This example illustrates the distinctions that companies should consider when analyzing and using the data that they collect.

This list of applicable laws is nonexhaustive, and finding whether a practice violates any of these laws is a highly fact-specific inquiry. To assist businesses, the report provides two sets of questions that companies can use when engaging in data analytics—one for legal compliance and one for policy-related issues—and concludes that companies should be mindful of applicable laws when using big data analytics to ensure that their practices do not result in what could be unlawful bias.

Source URL:https://natlawreview.com/article/avoiding-bias-big-data