

Federal Trade Commission

Emerging Issues in Big Data and Analytics

February 4, 2016

Tiffany George, Attorney

Federal Trade Commission

Division of Privacy and Identity Protection

The views expressed are my own and not those of the Commission or any individual Commissioner.

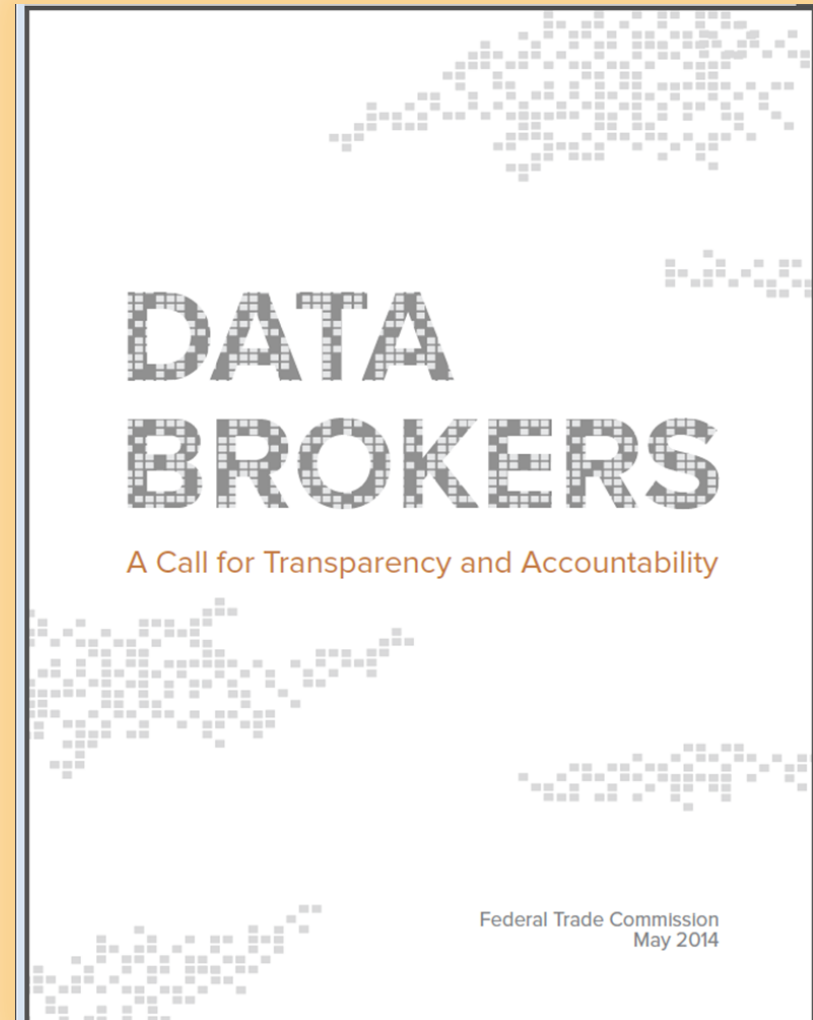


What is Big Data?

- The ability to collect consumer data from a variety of sources and use algorithms to:
 - extract hidden information
 - identify correlations
 - make predictions
 - draw inferences
 - glean new insights
- Three Vs:
 - volume,
 - velocity, and
 - variety.

Data Broker Report

- Sent information requests to nine data brokers:
 - Nature and sources of data?
 - Use, maintenance, and dissemination of data?
 - Give consumers access, and the ability to correct and/or opt out?
- The report:
 - summarizes findings,
 - proposes legislation, and
 - recommends best practices.



Data Broker Report Findings: Characteristics of the Industry

- Data brokers collect consumer data from numerous sources, largely without consumers' knowledge.
- The data broker industry is complex, with multiple layers of data brokers providing data to each other.
- Data brokers collect and store billions of data elements covering nearly every U.S. consumer.
- Data brokers combine and analyze data about consumers to make inferences about them, including potentially sensitive information.
- Data brokers combine online and offline data to market to consumers online.

Data Broker Report Findings: Consumer Choices

- To the extent data brokers offer consumers choices about their data, the choices are largely invisible and incomplete.
 - Marketing: Limited Access and not all allow correction.
 - Risk Mitigation: Not all provide access and only one allows for correction.
 - People Search Products: Not all allow consumers to opt-out.

Data Broker Report Findings: Benefits and Risks

- Consumers benefit from many of the purposes for which data brokers collect and use data:
 - prevent fraud
 - improve product offerings, and
 - deliver tailored advertisements to consumers.
- Many of the purposes for which data brokers collect and use data pose risks to consumers. Examples:
 - inability to conclude a transaction based on an error in a risk mitigation product,
 - “Biker Enthusiast” data segment could help a motorcycle dealership offer a consumer coupons, but it could also be used by an insurance company to assume the consumer engages in risky behavior.
- Storing data about consumers indefinitely may create security risks.

Recommendations

- Legislative
 - Notice and disclosure
 - Access and correction
 - Opt out and suppression
- Best Practice
 - Privacy by Design
 - Collect only data needed
 - Dispose of data as it becomes less useful
 - Refrain from collecting from children and teens
 - Ensure downstream users don't use information for FCRA or discriminatory purposes

Big Data Report

- Big Data: A Tool for Inclusion or Exclusion
 - September 2014 Workshop
 - Spring 2014 Seminar on Alternative Scoring Products
- The report:
 - Life cycle
 - Benefits and risks
 - Potentially applicable laws
 - Recent research



Life Cycle of Big Data

- Collection
- Compilation and consolidation
- Data mining and analytics
- Use

Benefits

- Increase educational attainment for individual students.
- Provide access to credit using non-traditional methods.
- Provide healthcare tailored to individual patients' characteristics.
- Provide specialized healthcare to underserved communities.
- Increase equal access to employment.

Risks

- Result in more individuals mistakenly being denied opportunities based on the actions of others.
- Create or reinforce existing disparities.
- Expose sensitive information.
- Assist in the targeting of vulnerable consumers for fraud.
- Create new justifications for exclusion.
- Result in higher-priced goods and services for lower income communities.
- Weaken the effectiveness of consumer choice.

Applicable Laws

- Fair Credit Reporting Act
 - Eligibility determinations
- Equal Credit Opportunity Act
 - Disparate treatment
 - Disparate impact
- Section 5 of FTC Act
 - Deceptive or unfair practices

Questions for Legal Compliance

- If you compile big data for others who will use it for eligibility decisions (such as credit, employment, insurance, housing, government benefits, and the like), are you complying with the accuracy and privacy provisions of the FCRA?
- If you receive big data products from another entity that you will use for eligibility decisions, are you complying with the provisions applicable to users of consumer reports?

Questions for Legal Compliance

- If you are a creditor using big data analytics in a credit transaction, are you complying with the requirement to provide statements of specific reasons for adverse action under ECOA?
- Are you complying with ECOA requirements related to requests for information and record retention?
- If you use big data analytics in a way that might adversely affect people in their ability to obtain credit, housing, or employment:
 - Are you treating people differently based on a prohibited basis, such as race or national origin?
 - Do your policies, practices, or decisions have an adverse effect or impact on a member of a protected class, and if they do, are they justified by a legitimate business need that cannot reasonably be achieved by means that are less disparate in their impact?

Questions for Legal Compliance

- Are you honoring promises you make to consumers and providing consumers material information about your data practices?
- Are you maintaining reasonable security over consumer data?
- Are you undertaking reasonable measures to know the purposes for which your customers are using your data?
 - If you know that your customer will use your big data products to commit fraud or for discriminatory purposes, do not sell your products to that customer. If you have reason to believe that your data will be used for these purposes, ask more specific questions about how your data will be used.

Summary of Research Considerations

- Consider whether your data sets are missing information from particular populations and, if they are, take appropriate steps to address this problem.
- Review your data sets and algorithms to ensure that hidden biases are not having an unintended impact on certain populations.
- Remember that just because big data found a correlation, it does not necessarily mean that the correlation is meaningful. As such, you should balance the risks of using those results, especially where your policies could negatively affect certain populations. It may be worthwhile to have human oversight of data and algorithms when big data tools are used to make important decisions, such as those implicating health, credit, and employment.
- Consider whether fairness and ethical considerations advise against using big data in certain circumstances. Consider further whether you can use big data in ways that advance opportunities for previously underrepresented populations.

Additional Resources

- Internet of Things Report
- Cross Device Tracking Workshop
- PrivacyCon Workshop
- FTC.gov

Federal Trade Commission

Questions?
tgeorge@ftc.gov

