Comments of the
Software & Information Industry Association (SIIA)
on the
Federal Trade Commission Hearings on Competition and
Consumer Protection in the 21st Century
August 20, 2018

Topic #9: The consumer welfare implications associated with the use of algorithmic decision tools, artificial intelligence, and predictive analytics

About SIIA

With nearly 700 member companies, SIIA is the principal trade association of the software and digital content industries. Our members are global industry leaders in the development and marketing of software and electronic content for business, education, government and consumer markets. They range from start-up firms to some of the largest and most recognizable corporations in the world. SIIA member companies are leading providers of, among other things:

- Data analytics and artificial intelligence,
- business, enterprise and networking software,
- publishing, graphics, and photo editing tools,
- corporate database and data processing software,
- financial trading and investing services, news, and commodities,
- online legal information and legal research tools,
- education software, digital content and online education services,
- specialized business media,
- open source software, and
- many other products and services in the digital content industries.

Introduction

Big data analytics, including machine learning (ML) and artificial intelligence (AI), are natural outgrowths of recent developments in computer technology, including the availability of massive data sets, vast increases in computing power, and breakthroughs in analytical techniques. These techniques promise transformative benefits for consumers, workers, and society at large. SIIA members are industry leaders using analytics to promote social and economic opportunity through, for example, risk and alternative credit scoring models, predictive analytics in education, detecting discrimination, and for other purposes.

Two MIT business professors claim that "the most important general-purpose technology of our era is artificial intelligence, particularly machine learning (ML) — that is, the machine's ability to keep improving its performance without humans having to explain exactly how to accomplish all the tasks it's given."[1] They reason is that it will dramatically decrease the cost of tasks that were previously thought to be the sole province of human workers.

One result of the use of machine learning predictive analytics is the transformation of many tasks into prediction tasks. For instance, driving is interpreted by machine learning algorithms as predicting what a human driver would do. Because AI makes prediction easier, it makes tasks involving prediction cheaper and encourages re-interpreting tasks involving prediction in order to make them cheaper.[2]

Companies big and small are also incorporating AI into their back-office operations, including those with extensive record keeping and assessment requirements. Insurance companies are finding that they can cut costs, do more work with the same number of workers and increasing revenues. These trends are just beginning and are expected to grow substantially over the next few years.[3]

The overall benefits of this transformation of work are bound to increase consumer welfare in the economy as a whole. Measuring productivity growth is always difficult, but there is no doubt that this growth has stagnated over the last two decades across various sectors of the economy. AI might be a fix for this problem. Former head of the Council of Economic Advisors, Jason Furman has said that his greatest worry about artificial intelligence is that "we do not have enough of it" to make a dent in our productivity slowdown. [4]

An extensive discussion of the issues associated with AI, labor productivity, and the effect on the workplace can be found in several SIIA issue briefs.[5]

Of course, these advanced analytic techniques can also pose challenges for fairness, lead to unintentional discrimination, or other undesirable outcomes, depending on how they are used. In 2014, the FTC held an informative workshop and comment process on this topic, during which SIIA's Mark MacCarthy participated on a panel for the workshop, "Big Data: A Tool for Inclusion or Exclusion?" The workshop usefully framed important developments in the use of data analytics for providing services to low income and underserved consumers and provided a forum for discussion of the possibility of unfair or discriminatory use of data analytics.

SIIA has also released several issue briefs on algorithmic fairness and the ethical use of advance algorithms.[6]

---

[1] Brynjolfsson, Eric and McAfee, Andrew, "The Business of Artificial Intelligence," Harvard Business Review, July 18, 2017, available at https://hbr.org/cover-story/2017/07/the-business-of-artificial-intelligence.

[2] Ajay Agrawal, Joshua Gans and Avi Goldfarb, "The Simple Economics of Machine Learning," Harvard Business Review, November 17, 2016 available at https://hbr.org/2016/11/the-simple-economics-of-machine-intelligence

[3] Steve Lohr, 'The Beginning of a Wave': A.I. Tiptoes into the Workplace, New York Times, August 6, 2018, available at https://www.nytimes.com/2018/08/05/technology/workplace-ai.html.

[4] Jason Furman, "Is This Time Different? The Opportunities and Challenges of Artificial Intelligence," Remarks at AI Now: The Social and Economic Implications of Artificial Intelligence Technologies in the Near Term, July 7, 2016, available at https://obamawhitehouse.archives.gov/sites/default/files/page/files/20160707_cea_ai_furman.pdf

[5] See New Economic and Policy Research on AI and the Future of Work (2017) and Artificial Intelligence and the Future of Work (2016)

[6] See Algorithmic Fairness (2016) and Ethical Principles for Artificial Intelligence and Data Analytics (2017)

SIIA remains confident about the opportunities for data analytics to ultimately promote inclusive economic growth. Examples of such initiatives include AI4ALL, a nonprofit organization formed by a group of Universities to increase diversity and inclusion in the field of artificial intelligence, with the goal for AI to be developed by a broad group of thinkers and doers advancing AI for humanity's benefit.[7] Additionally, Google's People + AI Research initiative (PAIR), which brings together researchers across Google to study and redesign the ways people interact with AI systems. The goal of PAIR is to focus on the "human side" of AI: the relationship between users and technology, the new applications it enables, and how to make it broadly inclusive.[8]

However, it is an important function of public policy to endorse and incentivize responsible application of data analytics, and to promote social and economic opportunity. The following are a series of considerations that SIIA raised during the 2014 FTC initiative, which should inform the upcoming hearings as well.

The use of analytics in regulated contexts covered by existing anti-discrimination and consumer protection laws.

Adequate and appropriate protections under existing anti-discrimination and consumer protection laws already apply to the use of analytics in regulated eligibility contexts such as lending, insurance, housing, and employment. These protections continue to function effectively with the increased use of AI/ML, and the ever-increasing amounts of data. Statutory constraints on discrimination in these areas include:

- Title VII of the Civil Rights Act of 1964 makes it unlawful for employers and employment agencies to discriminate against an applicant or employee because of such individual's "race, color, religion, sex, or national origin."[9] This is enforced by the Equal Employment Opportunity Commission and state fair employment practices agencies.
- The Equal Credit Opportunity Act makes it unlawful for any creditor to discriminate against any applicant for credit on the basis of "race, color, religion, national origin, sex or marital status, or age,"[10] which is enforced by the Consumer Financial Protection Bureau.[11]
- Title VIII of the Civil Rights Act of 1968, the Fair Housing Act**,** prohibits discrimination in the sale, rental, or financing of housing "because of race, color, religion, sex, familial status, or national origin."[12] The act also protects people with disabilities and families with children. It is enforced by the Department of Housing and Urban Development.

[7] AI4ALL was founded by leading AI technologists to increase diversity and inclusion in artificial intelligence. Since the program was first piloted at Stanford in 2015, AI4ALL has created transformative impact for more than 1000 underrepresented youth and is tripling its impact in 2018 in partnership with Universities including Boston University, Princeton, Carnegie Mellon, UC Berkeley and companies like NVIDIA, Autodesk and Google. Learn more about AI4ALL education programs here: http://ai-4-all.org/

[8] In July, 2017, Google launched the People + AI Research initiative (PAIR), which brings together researchers across Google to study and redesign the ways people interact with AI systems. The goal of PAIR is to focus on the "human side" of AI: the relationship between users and technology, the new applications it enables, and how to make it broadly inclusive. The goal isn't just to publish research; we're also releasing open source tools for researchers and other experts to use.

[9] 42 U.S.C. §2000e-2 available at http://www.law.cornell.edu/uscode/text/42/2000e-2

[10] 15 U.S.C. § 1691 available at http://www.law.cornell.edu/uscode/text/15/1691

[11] The Federal Reserve Board originally enforced the Equal Credit Opportunity Act, but the Dodd-Frank Act of 2011 transferred jurisdiction to CFPB. See Consumer Financial Protection Bureau, CFPB Consumer Protection Laws: ECOA, June 2013 p. 1 available at http://files.consumerfinance.gov/f/201306_cfpb_laws-and-regulations_ecoacombined-june-2013.pdf

[12] 42 U.S.C. 3604 available at http://www.law.cornell.edu/uscode/text/42/3604

- The Genetic Information Nondiscrimination Act of 2008 prohibits U.S. health insurance companies and employers from discriminating on the basis of information derived from genetic tests.[13] Enforcement is divided among a number of agencies including the Department of Health and Human Services (for health insurance) and the Equal Employment Opportunity Commission (for employment).

In addition, there are requirements under the Fair Credit Reporting Act when information is used for employment, insurance, or credit granting.[14] These requirements include notice of adverse action, disclosures, access and correction, use limitations, consent, and redress.

These extensive anti-discrimination and consumer protection laws apply to the use of advanced algorithms in these regulated contexts. For instance, The FTC ruled in the case of Spokeo that online data aggregators who use the latest machine learning techniques are covered under the FCRA when they make profiles available to third-parties for FCRA-regulated purposes.[15]

Due diligence and transparency mandates should not be applied to new contexts absent a showing that they are needed to prevent real harms.

The legal requirements appropriate for preventing unfairness and discrimination were imposed to cover specific cases where it was thought the dangers of harm were the greatest and where there was evidence that these harms were in fact occurring. Housing, employment, insurance, and credit granting are so important to the life prospects of individuals that unfair treatment in these areas could substantially reduce their chances of economic and social success.

This complex set of legal requirements was intentionally created narrowly to apply to specific protected classes and specific contexts. It should not be extended to new protected classes or to new contexts without a similar showing that the new requirements are need to prevent real harm.

Experts widely agree that merely targeting ads to different demographics can be done without creating disparate or undesirable outcomes, and frequently can be net beneficial for those businesses and specific communities. While the laws outlined above prohibit discriminatory ads for areas such as housing and employment, ads targeting a specific group are often necessary to expand markets and advertise specific services (e.g., Univision ads targeting native Spanish speakers). Targeted advertisements for products and services do not foreclose opportunities where any consumer could learn about those products and services through other means.

Trade secrets and business confidentiality should continue to protect algorithms from disclosure.

Complete algorithmic transparency would prevent the use of trade secrets in the area of data analytics. Such a proposal is harmful and it does not meaningfully advance any public policy purpose. Even the

---

[13] Pub. L. No. 110-233, 122 Stat. 881 available at http://www.gpo.gov/fdsys/pkg/PLAW-110publ233/pdf/PLAW110publ233.pdf

[14] http://www.consumer.ftc.gov/sites/default/files/articles/pdf/pdf-0096-fair-creditreporting- act.pdf

[15] FTC, Spokeo to Pay $800,000 to Settle FTC Charges Company Allegedly Marketed Information to Employers and Recruiters in Violation of FCRA, June 12, 2012, available at https://www.ftc.gov/news-events/press-releases/2012/06/spokeo-pay-800000-settle-ftc-charges-company-allegedly-marketed

industries regulated for anti-discrimination purposes and for fairness are permitted to maintain their algorithms as trade secrets.

The fundamental rationale for trade secrets is to provide those who invent something the ability to exploit it without others being able to take advantage of the fruits of their innovative efforts. Trade secrets promote innovation and competition.

Imagine how a proposal for algorithmic transparency would work in practice. Consider a company that invested billions of dollars developing software and analytical capabilities to provide new insights into healthcare. Under algorithm transparency, as soon these capacities are brought to market they must be revealed to the general public. This revelation would include the formulas used, the software code involved, and perhaps even the software engineering notes used to develop these analytical capabilities. The information would then be available to the company's competitors, who would immediately copy the successful techniques and rush to market with an alternative product that does much the same thing. But having spent no resources to develop the product, they can offer it at a fraction of the cost of the original developer. Why would analytics companies invest huge sums in developing, improving, updating their algorithms if they were required to disclose them immediately to competitors?

Data journalists interested in understanding the uses of algorithms for decision making recognize the difficulties of public transparency as a solution. One noted that companies are reluctant to make their statistical models public since "…exposing too many details of their proprietary systems (trade secrets) may undermine their competitive advantage, hurt their reputation and ability to do business, or leave the system open to gaming and manipulation."[16]

Gaming and manipulation are "real issues" he says, quoting Goodhart's rule that when a measure becomes a target it ceases to be a good measure. For example, measures for fraud prevention and identity authentication would become difficult if not impossible if the fraudsters and identity thieves knew the details of the algorithms used to detect them. Spammers and criminals already spend billions trying to game search results, a task that would be much easier if they know more about the internal workings of search algorithms.

As concluded by former FTC Commissioner, Julie Brill, full algorithmic transparency is affirmatively harmful to the industries involved and to the public interest in competition and innovation in the data analytics industry.[17]

The focus of policy concern should be on the uses of algorithms, not on the algorithms themselves.

The idea that an algorithm "makes a decision" needs to be rejected clearly and completely if there is to be any progress in understanding the issues raised by data analytics. In many cases, the notion of algorithmic decision-making is just shorthand for the idea that organizations use algorithms to aid them in making decisions. But like many such pieces of shorthand it can lead policymakers to focus in the

[16] Diakopoulosa, Nicholas, Algorithmic Accountability, Digital Journalism, 3:3, 398-415, 2015 available at http://www.nickdiakopoulos.com/wp-content/uploads/2011/07/algorithmic_accountability_final.pdf

[17] Brill, Julie, *Transparency, Trust, and Consumer Protection in a Complex World,* Keynote Address Before Coalition for Networked Information Fall 2015 Membership Meeting (December 15, 2015).

wrong area. In particular, it seems to suggest that the analytic tool is the issue rather than the use of the tool for particular purposes.

An audio file recognition program is a system that matches a new audio file against a library of such files and estimates the probability that the new file is the same as one of the files in its library. An organization determines what level of probability it will accept as an indication of a match, and that depends on the purposes for which it is attempting to match the files. The purposes could include letting people know what music they happen to be listening to or detecting copyright infringement. The algorithm determines neither the acceptable probability level nor the purpose for which a match will be used. If there is a policy issue, it relates to the use of the algorithm, not to the algorithm itself.

Some characteristics reveal race and ethnicity so clearly that they effectively function as proxies for them—two well-known examples are surname and census geography. There are advanced statistical techniques that can combine these two characteristics to generate a new variable that is an even better statistical proxy for race and ethnicity.

An approach that says to focus on the algorithm itself might say that the construction and development of these proxies is a matter of concern and should be discouraged. Of course, the use of proxies for race and ethnicity to make eligibility decisions in the areas of housing, lending, insurance and employment would be problematic under the existing anti-discrimination laws. Yet, these proxies are useful in assessing whether lending decisions have a disproportionate, adverse impact on protected classes, as the CFPB explained in 2014.[18] The issue for policymakers is not the existence of proxies for race and ethnicity but how these proxies are used. Some uses promote exclusion; some are essential in the fight against it.

In general, policymakers should be concerned with how algorithms are used, not with the construction or existence of an algorithm.

[18] Consumer Financial Protection Board, Using Publicly Available Information to Proxy for Unidentified Race and Ethnicity, Summer 2014 at http://files.consumerfinance.gov/f/201409_cfpb_report_proxy-methodology.pdf. p. 23