FTC | Aiming for truth, fairness, and equity in your company's use of AI

Advances in artificial intelligence (AI) technology promise to revolutionize our approach to medicine, finance, business operations, media, and more. But research has highlighted how apparently "neutral" technology can produce troubling outcomes – including discrimination by race or other legally protected classes. For example, COVID-19 prediction models can help health systems combat the virus through efficient allocation of ICU beds, ventilators, and other resources. But as a recent study(link is external) (https://academic.oup.com/jamia/article/28/1/190/5893483) in the Journal of the American Medical Informatics Association suggests, if those models use data that reflect existing racial bias in healthcare delivery, AI that was meant to benefit all patients may worsen healthcare disparities for people of color.

The question, then, is how can we harness the benefits of AI without inadvertently introducing bias or other unfair outcomes? Fortunately, while the sophisticated technology may be new, the FTC's attention to automated decision making is not. The FTC has decades of experience enforcing three laws important to developers and users of AI:

- **Section 5 of the FTC Act.** The FTC Act prohibits unfair or deceptive practices. That would include the sale or use of – for example – racially biased algorithms.
- **Fair Credit Reporting Act.** The FCRA comes into play in certain circumstances where an algorithm is used to deny people employment, housing, credit, insurance, or other benefits.
- **Equal Credit Opportunity Act.** The ECOA makes it illegal for a company to use a biased algorithm that results in credit discrimination on the basis of race, color, religion, national origin, sex, marital status, age, or because a person receives public assistance.

Among other things, the FTC has used its expertise with these laws to report on big data analytics and machine learning (https://www.ftc.gov/reports/big-data-tool-inclusion-or-exclusion-understanding-issues-ftc-report); to conduct a hearing on algorithms, AI and predictive analytics (https://www.ftc.gov/news-events/events-calendar/ftc-hearing-7-competition-consumer-protection-21st-century); and to issue business guidance on AI and algorithms (https://www.ftc.gov/news-events/blogs/business-blog/2020/04/using-artificial-intelligence-algorithms). This work – coupled with FTC enforcement actions – offers important lessons on using AI truthfully, fairly, and equitably.

Start with the right foundation. With its mysterious jargon (think: "machine learning," "neural networks," and "deep learning") and enormous data-crunching power, AI can seem almost magical. But there's nothing mystical about the right starting point for AI: a solid foundation. If a data set is missing information from particular populations, using that data to build an AI model may yield results that are unfair or inequitable to legally protected groups. From the start, think about ways to improve your data set, design your model to account for data gaps, and – in light of any

shortcomings – limit where or how you use the model.

Watch out for discriminatory outcomes. Every year, the FTC holds PrivacyCon, a showcase for cutting-edge developments in privacy, data security, and artificial intelligence. During PrivacyCon 2020 (https://www.ftc.gov/news-events/events-calendar/privacycon-2020), researchers presented work showing that algorithms developed for benign purposes like healthcare resource allocation and advertising actually resulted in racial bias. How can you reduce the risk of your company becoming the example of a business whose well-intentioned algorithm perpetuates racial inequity? It's essential to test your algorithm – both before you use it and periodically after that – to make sure that it doesn't discriminate on the basis of race, gender, or other protected class.

Embrace transparency and independence. Who discovered the racial bias in the healthcare algorithm described (https://www.ftc.gov/system/files/documents/public_events/1548288/privacycon-2020-ziad_obermeyer.pdf) at PrivacyCon 2020 and later published in Science? Independent researchers spotted it by examining data provided by a large academic hospital. In other words, it was due to the transparency of that hospital and the independence of the researchers that the bias came to light. As your company develops and uses AI, think about ways to embrace transparency and independence – for example, by using transparency frameworks and independent standards, by conducting and publishing the results of independent audits, and by opening your data or source code to outside inspection.

Don't exaggerate what your algorithm can do or whether it can deliver fair or unbiased results. Under the FTC Act, your statements to business customers and consumers alike must be truthful, non-deceptive, and backed up by evidence. In a rush to embrace new technology, be careful not to overpromise what your algorithm can deliver. For example, let's say an AI developer tells clients that its product will provide "100% unbiased hiring decisions," but the algorithm was built with data that lacked racial or gender diversity. The result may be deception, discrimination – and an FTC law enforcement action.

Tell the truth about how you use data. In our guidance on AI (https://www.ftc.gov/news-events/blogs/business-blog/2020/04/using-artificial-intelligence-algorithms) last year, we advised businesses to be careful about how they get the data that powers their model. We noted the FTC's complaint against Facebook (https://www.ftc.gov/system/files/documents/cases/182_3109_facebook_complaint_filed_7-24-19.pdf), which alleged that the social media giant misled consumers by telling them they could opt in to the company's facial recognition algorithm, when in fact Facebook was using their photos by default. The FTC's recent action against app developer Everalbum (https://www.ftc.gov/news-events/press-releases/2021/01/california-company-settles-ftc-allegations-it-deceived-consumers) reinforces that point. According to the complaint, Everalbum used photos uploaded by app users to train its facial recognition algorithm. The FTC alleged that the company deceived users about their ability to control the app's facial recognition feature and made misrepresentations about users' ability delete their photos and videos upon account deactivation. To deter future violations, the proposed order (https://www.ftc.gov/enforcement/cases-proceedings/1923172/everalbum-inc-matter) requires the company to delete not only the ill-gotten data, but also

the facial recognition models or algorithms developed with users' photos or videos.

Do more good than harm. To put it in the simplest terms, under the FTC Act, a practice is unfair (https://www.ftc.gov/about-ftc/what-we-do/enforcement-authority) if it causes more harm than good. Let's say your algorithm will allow a company to target consumers most interested in buying their product. Seems like a straightforward benefit, right? But let's say the model pinpoints those consumers by considering race, color, religion, and sex – and the result is digital redlining (similar to the Department of Housing and Urban Development's case against Facebook (https://www.hud.gov/sites/dfiles/Main/documents/HUD_v_Facebook.pdf) in 2019). If your model causes more harm than good – that is, in Section 5 parlance, if it causes or is likely to cause substantial injury to consumers that is not reasonably avoidable by consumers and not outweighed by countervailing benefits to consumers or to competition – the FTC can challenge the use of that model as unfair.

Hold yourself accountable – or be ready for the FTC to do it for you. As we've noted, it's important to hold yourself accountable for your algorithm's performance. Our recommendations for transparency and independence can help you do just that. But keep in mind that if you don't hold yourself accountable, the FTC may do it for you. For example, if your algorithm results in credit discrimination against a protected class, you could find yourself facing a complaint alleging violations of the FTC Act and ECOA. Whether caused by a biased algorithm or by human misconduct of the more prosaic variety, the FTC takes allegations of credit discrimination very seriously, as its recent action against Bronx Honda (https://www.ftc.gov/news-events/press-releases/2020/05/bronx-honda-to-pay-over-1-million-to-settle-charges) demonstrates.

As your company launches into the new world of artificial intelligence, keep your practices grounded in established FTC consumer protection principles.