Algorithmic Accountability in the Area of Consumer Privacy: Is it Make Believe or for Real?

Anna Aurora Wennakoski*
Ph.D. student in privacy & data protection law
Benjamin N. Cardozo School of Law
New York City, NY
anna.wennakoski@gmail.com

Artificial intelligence (AI) is increasingly everywhere – and its promoters certainly do not shy away from its promises. We find ourselves in a world awash in big data, analytics, the Internet of Things (IoT). It is a blossoming issue amongst marketers, and corporations are increasingly adopting novel AI techniques to capture new clients and to better serve their needs.

This paper's primary focus is on consumer protection as one aspect of the algorithmic society we find ourselves in, in which we ask the questions: how accurately does artificial intelligence succeed in delivering customer expectations and in capturing their preferences in fields such as targeted marketing; is AI able to accurately classify, filter, or segregate consumer data containing personal or otherwise sensitive data when using various types of tracking software? And from a legal standpoint, what issues are raised with regard to the use of AI in targeted marketing, especially with regard to the enforcement of privacy policies?

Emerging AI Issues in the Consumer Privacy Space

Described as the "science and engineering of making intelligent machines,"[1] AI is not even seven decades old since it made its first appearance in the literature. First confined to software deployed in mainframe computers, the field of AI now spans a gamut of applications, from being embedded in web-based Google and Amazon searches, to self-driving cars, to all manner of IoT speaking and recording devices and wearables. Today, "computers read[] and understand[] human language whether it's text or voice and then tak[e] defined actions based on that

* The author is the recipient of the Global Privacy Student Award of the International Association of Privacy Professionals (IAPP) in both 2016 and 2017. She wishes to thank Professor Felix Wu at Benjamin N. Cardozo School of Law in New York for his inspiring lectures, as well as Jason R. Baron, one of the coordinators of the DESI VII Workshop, for his comments and editorial assistance.

[1] Shannon, Claude E. and John McCarthy, eds., Automata Studies, Annals of Mathematics Studies, Number 34, 1956.

interpretation of human language"[2] – sometimes "designed to remedy or supplement the shortcomings of human reasoners."[3]

One approach to algorithms is to cast a look at contracts. For consumers, it is important to be clear on when a contract has been formed. However, with the use of algorithms, there is a question as to who (as the legal subject) is then the concluder of an agreement. The device, the system itself, the coder, or the user? And what if the machine or device that participates in making an agreement makes an error?

Recently, some experts have begun to critically ask whether the data emitted by various IoT devices actually is reliable. One example is Fitbit data, and there are many others.

Essentially, the new forms of software technologies extrapolate more exact results from a known subset or pattern of data. For example, to find relevant evidentiary needles in corporate digital haystacks, technology assisted review techniques have emerged which hold out the promise of largely automating e-discovery search. However, in order to have valid and reliable results, one should aim to create data sets which enable the validation and re-iteration of the results. Such methods can include various classification and training sets. Yet, the end results can still vary, given the immaturity of some of these technologies. For sure, it is not that all algorithms will end up being fallacious; rather, the ways in which some of them operate will still leave room for improvements.

Importantly for us here, *the method used to ask for relevant data affects what kind of data is received as a result.* This is because the algorithms which are at the very foundation of these technical tools all need data as an input; and if that data is not valid, or the software has not been adequately trained to focus on the relevant questions, neither will we obtain accurate or valid answers. So too, there may be a gray area between what is considered relevant and non-relevant data. One therefore needs to have in place some kind of quality control feedback loop, as part of any kind of active, supervised machine learning process.

A second point to note is that algorithms can be biased and thus, lead to wrong end-conclusions, notwithstanding the fancy mathematics at work in the "black box." To the extent we care about decisionmaking processes, we should be asking whether decisions based on algorithms are capable of being fully justified after the fact, or in which ways can they be said to be limited. This becomes a more urgent question in a number of domains, as we are experiencing a shift from less human-centric to more machine-centric systems and processes – where AI is taking over.

---

[2] Dera J. Nevin, E-discovery counsel at Proskauer, at the NYSBA Annual Meeting on January 25 2017, in New York City.

[3] Daniel Martin Katz: Quantitative Legal Prediction - or how I Learned to stop Worrying and start Preparing for the data-driven future of the Legal Services Industry, p. 928.

These observations lead us to consider how algorithmic accountability is obtained in a world of consumer personal data. This is of interest not only in the US, but also in jurisdictions such as the EU, which find themselves within a legal regime (e.g., soon with that of the General Data Protection Regulation or GDPR), which contains a very broad definition of what constitutes "personal data."[4]

Policing Consumer Privacy Policies in a World of AI

US privacy policies are enforced largely by the Federal Trade Commission (FTC), which monitors statements made by companies about their practices with regard to personal information. Importantly, however, privacy policies are not contracts in the traditional sense, as they typically lack the other party's express consent. Courts view such policies as rather like unilateral promises, or general statements of policy.[5] Contract claims based on privacy policies typically fail due to the inability to show damages. Some authors have noted how the terms used in privacy policies are standardized, akin to mere boilerplate—casting serious doubt on the individual's role in the process. Indeed, as the court noted in *Google*[6], plaintiffs who are not Gmail or Google apps users are not subject to any of Google's express agreements. The same can be said to hold true with respect to other companies, vis à vis their users as well as third parties.

From a practical standpoint, in many privacy policies, the definition of "personal information" is also relatively broad and open ended, e.g., phrased in terms of "including but not limited to. . . . " Furthermore, sometimes the policies provide no explanation as to the "aggregated or non-personally identifying information" which "may be shared with third parties for advertising or other purposes." Yet consumers typically expect that the relationship they engage in with a company is a straightforward, one to one, i.e., a two-party relation – not one involving third parties such as data brokers.

In addition, in some cases what is important is not what is written in the policy but what is missing therefrom, including (i) descriptions of the use of data mining and other automated techniques of varying kinds, (ii) whether the corporation employs the use of encrypted connections; (iii) what is corporate policy regarding the recording of conversations with consumers of the product or service, and (iv) what expectations a consumer has about the security of his or her personal data in terms of it being "sold" to third parties.

---

[4] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) , OJ 119, 4.5.2016, p. 1–88, article 4(1).

[5] See *In re Google Inc, Gmail Litigation* (N.D. Cal 2013) (holding that a user's acceptance of these statements does not establish explicit consent, and that a reasonable user who read the privacy policies would not have necessarily understood that her emails were being intercepted).

[6] In re Google Inc. Gmail Litigation, N.D. Cal. Devc. 28 2012.

All these types of deficiencies arguably may form the basis of a claim of unfairness or deception by omission under section 5 of the FTC Act.[7] Note, however, that the FTC act does not provide for a private cause of action; instead, the FTC itself has authority to initiate enforcement actions where it sees fit.

In the EU, the Convention 108 and Directive 95/46 have had provisions on appropriate security measures for organizations to employ.[8] The Directive 95/46 obliged organizations to implement appropriate technical and organizational measures. Under the Directive, this was considered a duty of the controllers.[9] The requirement of data security has been confirmed in EU case law.[10]

Under the GDPR, each "controller and the processor shall implement appropriate technical and organizational measures to ensure a level of security appropriate to the risk."[11] This is one of the areas in which the risk-based approach the GDPR purports to take becomes most visible. Measures can include providing for encryption of data, as well as other technical tools. In practice, organizations can choose between a whole host of technical tools to meet security requirements. It remains to be seen, however which of these will be deemed sufficient to meet organizational compliance obligations.

Specifically regarding data security related to the transmission of data when fulfilling the new data portability requests, the WP29 has stated that it belongs to the data controller -- apparently referring to the initial controller, to whom the data subject made the request – to ensure appropriate data security.[12] By its wording, the GDPR seems to address controllers, despite the fact that processors might have a more considerable role in the processing itself through their assistance. [13] Given how today so many platforms are highly interlinked, arguably it can be difficult to assess when exactly does one organization's liability end and other processor's liability begin.[14]

So far as the undersigned is aware, neither the FTC nor the Court of Justice of the European Union have yet grappled specifically with the AI issues here discussed. Nor is the author aware of US or EU litigation involving a claim of violation of consumer privacy due to the use of a biased algorithm – unless one views expressed in *Schrems* as touching on such themes

---

[7] See, e.g . *Sears Holdings Management Corp*. (FTC Aug, 31, 2009).
[8] See ETS No.108, article 7; Directive 95/46/EC, article 17(1).
[9] Directive 95/46/EC, article 17.
[10] Case C-342/12 Worten - Equipamentos para o Lar SA v Autoridade para as Condições de Trabalho (ACT). Reports of Cases published in the electronic Reports of Cases (Court Reports - general) ("Worten") para. 24-25 and 28-29 in particular.
[11] GDPR, article 32.
[12] WP29 Guidelines on data portability, p. 15.
[13] GDPR, article 28(3)e.
[14] In the EU, the Network and Information Security Directive imposes enhanced cybersecurity obligations on "essential service operators" and "digital service providers" (e.g., banks, energy companies, infrastructure companies, telecoms, cloud service providers, etc.). For these entities, the NIS Directive requires them to maintain "appropriate and proportionate" security measures to manage risk. NIS Directive, Art. 14(1).

where it recognized the "fundamental right to respect for private life."[15] On the other hand, the Court has held that Data Retention Directive 2006/24 "exceeded the limits imposed by compliance with the principle of proportionality in the light of Articles 7, 8 and 52(1) of the Charter, and thus was invalid."[16] It only seems a matter of time, however, before cases will arise where a principle of algorithmic accountability will be defined, leading to the development of a common law of AI.

AI and "Real" Consumer Privacy Protections

How to then evaluate artificial intelligence in a world of amorphous privacy protections for the consumer? Against what standard ought we measure AI performance? In many instances, "the quality of the resulting predictions and the underlying models supporting most forecasts is unclear."[17] And who will be liable in the case of AI making a mistake?

The notion of algorithmic accountability for consumers remains to be seen. Despite the fact that algorithm-based businesses are flourishing, no express laws have yet been crafted to address pertinent issues related to AI, including the ways in which AI creates data subsets when analyzing existing data (i.e. our black-box problem). Yet businesses do face a plethora of general legal obligations to comply with which arguably touch on the outcomes of AI (most notably, as handed down by the FTC and as embedded within the GDPR). These include taking reasonable and appropriate measures to safeguard data (especially personal data), including considering the use of encryption, but also general good corporate governance and duties related thereto.

From a user perspective, "becoming your own filter" is increasingly a burdensome challenge. One the one hand, these filter bubbles can accentuate erroneous results; and on the other, create also new concerns over critical thinking and analyses., Together, these features underscore the need for learning and understanding the ways in which new data is being generated as well as the ways in which algorithms function.

Regarding consumer privacy policies, there is oftentimes a disconnect between the wording in the policies and reasonable expectations of consumers with respect to protection of their personal data. It would therefore appear judicious to start a re-evaluation of the exact wording in corporate privacy policies – an excellent domain for lawyers in human form to still work on. To that, there is a need for more expert level knowledge on how to approximate the words and the deeds and to formulate the policies accordingly.

---

[15] Case C-362/14 Maximillian Schrems v Data Protection Commissioner Reports of Cases published in the electronic Reports of Cases (Court Reports - general) ("Schrems").

[16] Digital Rights Ireland Ltd v Minister for Communications, Marine and Natural Resources and Others, and Kärntner Landesregierung and Others (joined cases C-293/12 & C-594/12). Reports of Cases published in the electronic Reports of Cases (Court Reports – general, see ¶ 69).

[17] Daniel Martin Katz;, Michael J Bommarito; and Josh Blackman: A General Approach for Predicting the Behavior of the Supreme Court of the United States, 11 Dec 2016, 1.