Stanford University
Human-Centered
Artificial Intelligence

November 21, 2022

Response to the Request for Comments on Trade Regulation Rule on Commercial Surveillance and Data Security

Stanford Institute for Human-Centered Artificial Intelligence

We are pleased to offer comments in response to the Federal Trade Commission's advance notice of proposed rulemaking on commercial surveillance and data security. We are a group of Stanford-affiliated researchers, faculty, and graduate students, and we submit these comments on behalf of ourselves individually and provide our affiliation for informational purposes only.

As this is an advanced notice of rulemaking, we offer feedback on both the scope of the topics the Commission is considering as well as the substantive points raised by the individual questions.

In these comments, we focus on the intersection of automated decision-making and information privacy. We review the questions posed in Sections III, IV, and V, offer feedback on the framing of the questions in these sections, and provide some substantive answers to the questions posed by the Commission.

Section III: Collection, Use, Retention, and Transfer of Consumer Data

While Section IV specifically addresses automated decision-making systems (ADM), we believe the questions raised in Section III have implications for algorithmic systems generally, and those that rely on artificial intelligence (AI)[1], such as machine learning (ML), specifically.

[1] We use the term "AI" as shorthand to refer to a set of technologies that rely on modern machine learning approaches, particularly deep learning and neural networks.

Section III includes several questions about the possible introduction of two principles into the FTC regulatory framework: data minimization and purpose limitation. Data minimization and purpose limitation are two of the information privacy and data protection principles (FIPs) developed in late 1970s in the era before the existence of relational databases or the commercial internet and subsequently enshrined into policy and law around the globe.[2] The European Union's General Data Protection Regulation is one of the most significant examples of FIPs-based regulations; since 2018, data processors in the EU have had to put both data minimization (DM) and purpose limitation (PL) principles into practice.[3]

DM and PL may theoretically pose challenges to algorithmic decision-making systems that rely on large datasets of consumer data. To the extent that industry depends on using as many consumer data points as possible without a specified purpose in mind (e.g., data mining) when developing products and services, implementing these principles could curb these and similar practices.[4] With regard to AI specifically, massive datasets are enabling the development of increasingly robust and accurate models.[5] Accordingly, the normative view is that the more data, the better, and constraints on data use will hamper progress in AI, and constrain innovation in data-intensive applications.

However, the assumption that introducing DM and PL principles will inevitably lead to negative consequences for the ongoing development of ADM is apparent in Question 48: that DM and PL will "unduly hamper algorithmic decision-making or other algorithmic learning-based processes or techniques." This framing pits information privacy against ADM as zero-sum: You can have privacy at the expense of algorithmic decision-making, or algorithmic decision-making at the expense of privacy, but not both.

---

[2] For a general background, see Gellman, "Fair Information Practices: A Basic History – Version 2.22" (April 2022). Available at: https://doi.org/10.2139/ssrn.2415020.

[3] See Article 5(1)b-c and Article 6(1)a of the EU General Data Protection Regulation, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679.

[4] See generally, in support: Alon Halevy, Peter Norvig, and Fernando Pereira, "The Unreasonable Effectiveness of Data," in IEEE Intelligent Systems, 24, no. 2 (March/April 2009): 8-12; for a critique: danah boyd and Kate Crawford, "Critical Questions for Big Data," Information, Communication & Society (2012), 15:5, 662-79.

[5] Alec Radford et al., "Robust Speech Recognition via Large-Scale Weak Supervision" (2022). Available at: https://cdn.openai.com/papers/whisper.pdf; Dhruv Mahajan et al., "Exploring the Limits of Weakly Supervised Pretraining" (May 2018). Available at: https://doi.org/10.48550/arXiv.1805.00932.

We believe this is an unfair assumption and, if left unchallenged, supports a data ecosystem that today benefits the largest actors with the most access to data, constraining smaller actors and startups. In these comments, we discuss why we believe DM and PL are not incompatible with ADM. First, we suggest that DM and PL are important principles in a growing movement toward the responsible development of ADM, and that while they might pose challenges, they are not inherently incompatible with algorithmic innovation: Not only have EU-based companies been required to comply with DM and PL principles since the enactment of the GDPR in 2018, there is also a growing body of research demonstrating that building ADM systems within these constraints is feasible or even desirable. Second, we argue that companies must be held to high standards with regards to privacy and ethics when building products that rely on consumers' personal data. Third, DM and PL are integral to shifting the landscape of data acquisition and management away from one where companies are the primary stewards of consumers' personal data. Instead, we advocate for data intermediaries as a means by which consumers have more control over and receive direct benefit from their personal data, as well as investment in public data to create a data infrastructure for ADM that is more resilient and less dependent on private actors. Given the scope of these proposed rules and the long-range time frame they are intended to serve, the assumption that DM and PL are incompatible with ADM marries us to the existing status quo of platform dominance and assumes the broken data landscape of 10 to 20 years from now remains as it does today.

1. Data minimization and purpose limitation are not inherently barriers to innovation

With the introduction of the GDPR in 2018, companies operating within EU member states have been subject to DM and PL principles in practice for over four years. Critiques of the GDPR's embrace of these principles predominantly question the feasibility of DM in the context of technical advances in big data analytics, as well as highlight how such principles would "sacrifice considerable social benefit" by limiting the innovative potential of ML and AI techniques.[6] Other critiques don't so much question this approach to regulation as they criticize the post-GDPR implementation efforts to regulate data collection, with researchers pointing to a lack of

[6] Bert-Jaap Koops, , "The Trouble With European Data Protection Law" (August 2014). Available at: https://papers.ssrn.com/abstract=2505692; Mark MacCarthy, "In Defense of Big Data Analytics" (April 2018). Available at: https://doi.org/10.2139/ssrn.3154779.

consistency and specificity in the interpretation of DM and PL principles.[7] The absence of concrete definitions and guidelines—to some extent aided by a lack of attention from the computer science research community, which has focused more on fairness and transparency issues—may have led to only limited adoption of these principles and left implementation questions unresolved. For example, Biega and Finck, in their paper exploring the application of DM and PL to data-driven systems, question whether DM as practiced under the GDPR actually requires pseudonymization in order to practically comply with the regulation.[8]

However, a growing body of research has shown that researchers and practitioners are developing and implementing techniques that are adapting to the constraints posed by DM and PL. Many have had successes in implementing techniques that allow for the protection of data privacy without hindering AI innovation. A series of empirical studies have shown that the performance of data-driven systems can be replicated using comparatively less data.[9] Researchers have also proposed new, concrete frameworks to guide the technical implementation of DM and PL principles in order to bridge the gap between the legal principles and their practical realization.[10]

Various technical DM approaches and techniques are presently being tested and deployed at different stages of supervised ML models. Experts point toward feature selection methods and privacy-preserving methods (such as federated learning) that can be employed to minimize the amount of data used to develop models during the training phase.[11] At the inference phase, techniques that can effectively minimize data include converting personal data into less "human readable" formats, hosting ML models locally, and employing privacy-preserving query approaches

[7] Divya Shanmugam et al., "Learning to Limit Data Collection via Scaling Laws: A Computational Interpretation for the Legal Principle of Data Minimization" (June 2022). Available at: https://doi.org/10.48550/arXiv.2107.08096.
[8] Asia Biega and Michele Finck, "Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems" (December 2021). Available at: https://doi.org/10.26116/techreg.2021.004; Awanthika Senarath et al., "Understanding Software Developers' Approach Towards Implementing Data Minimization" (August 2018). Available at: https://doi.org/10.48550/arXiv.1808.01479.
[9] Biega and Finck (2021); Nicholas Vincent et al., "'Data Strikes': Evaluating the Effectiveness of a New Form of Collective Action Against Technology Companies" (2019). Available at: https://doi.org/10.1145/3308558.3313742; Hongyi Wen et al., "Exploring Recommendations Under User-Controlled Data Filtering" (2018). Available at: https://doi.org/10.1145/3240323.3240399.
[10] Shanmugam et al. (2022)
[11] Reuben Binns and Valeria Gallo, "Data Minimisation and Privacy-Preserving Techniques in AI Systems." Information Commissioner's Office (August 21, 2021). Available at: https://ico.org.uk/about-the-ico/media-centre/ai-blog-data-minimisation-and-privacy-preserving-techniques-in-ai-systems/.

(including anonymization). Other approaches across the phases include reducing the quality of data without reducing its overall quantity via techniques including data shuffling.[12]

Further, DM techniques have long been employed across the private sector for reasons that extend beyond data privacy considerations. Recognizing the additional security risks associated with collecting vast amounts of data, technology companies have developed tools and techniques to determine the minimal amount of data and set of features needed for ML models to make accurate predictions.[13] Common practices that can be adopted by companies to improve their ability to enact DM and PL requirements include clear data classification, review, and erasure processes, alongside data retention schedules.[14] At the same time, it is important to acknowledge the limitations in the existing literature: Many of the DM techniques have been developed with simplified assumptions in lab settings and have not been tested in real-world settings.[15] Addressing this limitation would, among other things, require developing DM techniques for a wider range of models.[16] More research is also needed to develop DM techniques that take into account the sensitivity of different data classes.[17] These limitations also point to the need for industry to devote resources toward developing this research, or to provide academic researchers the access to data that would allow them to do so.

There is still a need for more detailed research into the effects of DM and PL practices on other areas of concern such as algorithmic discrimination. For example, civil rights groups generally agree that DM is crucial in helping reduce harms caused by algorithmic bias.[18] Expansive data

[12] Biega and Finck (2021).
[13] Abigail Goldsteen, "Reduce Data Privacy Issues With Machine Learning Models," IBM Developer Blog (January 19, 2022). Available at: https://developer.ibm.com/blogs/data-minimization-for-machine-learning/; Goldsteen et al., "Data Minimization for GDPR Compliance in Machine Learning Models" (August 2022). Available at: https://doi.org/10.1007/s43681-021-00095-8.
[14] "Big Data, Artificial Intelligence, Machine Learning and Data Protection," Information Commissioner's Office. Accessed November 2022 via: https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf;
Malek, "Bigger Is Always Not Better; Less Is More, Sometimes: The Concept of Data Minimization in the Context of Big Data Section II: Focus Papers" (2021). Available at: https://heinonline-org.stanford.idm.oclc.org/HOL/Page?collection=journals&handle=hein.journals/ejplt2021&id=212&men_tab=srchresults.
[15] Wen et al. (2018).
[16] Goldsteen et al. (2022).
[17] Shanmugam et al. (2022).
[18] "Data Minimization: Key to Protecting Privacy and Reducing Harm," Access Now. Accessed November 2022 via: https://www.accessnow.org/cms/assets/uploads/2021/05/Data-Minimization-Report.pdf; Samantha Lai and Brooke Tanner, "Examining the Intersection of Data Privacy and Civil Rights," Brookings,

collection has historically proven particularly costly for minority groups and marginalized communities.[19] Not collecting certain data in the first place, including data on protected characteristics, can help prevent some discriminatory practices. However, experts have highlighted the potential unintended consequences of DM principles on algorithmic fairness efforts, amid concerns that they could jeopardize model fairness by impacting users differently.[20] Researchers have found that any small performance decrease caused by DM is likely to disparately impact marginalized groups, especially if such groups comprise a minority of the data.[21]

Others emphasize that DM may hinder fairness audits. The digital civil rights nonprofit Access Now has argued that DM requirements should include narrow exceptions for data that is collected for the purpose of addressing civil rights-related harms.[22] Recognizing that DM requirements may prevent or discourage organizations from collecting data on protected attributes, Access Now recommends that DM requirements include narrow carve-outs for organizations to collect data on protected classes for civil rights auditing and testing purposes. Scholars in the EU have also grappled with the question of whether the GDPR should grant exceptions for the collection and use of such data for narrow purposes such as preventing discrimination.[23] Yet others push back against collecting demographic data for fairness purposes, advocating instead for solutions including anonymization techniques that enable trusted third parties to collect and hold demographic data, as well as participatory data governance approaches.[24]

While we recommend that the various potential unintended consequences of DM and PL be taken into consideration in the FTC's regulatory framework, we do not believe that these present insurmountable barriers to the implementation of such principles. Ultimately, given the wide range

---

July 18, 2022. Available at: https://www.brookings.edu/blog/techtank/2022/07/18/examining-the-intersection-of-data-privacy-and-civil-rights/.

[19] Danielle Keats Citron and Daneil J. Solove, "Privacy Harms," 102 Boston University Law Review 793 (2022).

[20] Biega and Finck (2021); Gemma Galdon Clavell et al., "Auditing Algorithms: On Lessons Learned and the Risks of Data Minimization" (February 2020). Available at: https://doi.org/10.1145/3375627.3375852.

[21] Biega et al., "Operationalizing the Legal Principle of Data Minimization for Personalization" (July 2020). Available at: https://doi.org/10.1145/3397271.3401034.

[22] "Data Minimization: Key to Protecting Privacy and Reducing Harm" (2022).

[23] Marvin van Bekkum and Frederik Zuiderveen Borgesius, "Using Sensitive Data to Prevent Discrimination by Artificial Intelligence: Does the GDPR Need a New Exception?" (May 2022). Available at: https://doi.org/10.2139/ssrn.4104823.

[24] McKane Andrus and Sarah Villeneuve, "Demographic-Reliant Algorithmic Fairness: Characterizing the Risks of Demographic Data Collection in the Pursuit of Fairness" (May 2022). Available at: https://doi.org/10.48550/arXiv.2205.01038.

of feasible techniques that, in many cases, have already been implemented, it would be misleading at this stage in AI development specifically to assume that DM and PL inevitably harm the development and deployment of AI systems.

2. Companies Must Be Held to High Standards When Building ADM Systems Based on Personal Data

In addition to the range of potential and actual harms resulting from consumer data collection, the FTC also must consider the harm created during the collection process itself through the violation of individuals' right to privacy, autonomy, and freedom from commercial intrusion into all aspects of private life. Despite the potential effect of DM and PL principles on the development of ADM systems, unhampered "progress" by exploiting consumers' personal data at the expense of their privacy should not be a sanctioned goal. There are important ethical issues raised by commercial data collection and surveillance practices that DM addresses. For example, Access Now highlights that not only are most people unwilling to allow companies to collect data about them without restraint, but they also elaborate on the various harms that can arise from extensive data collection about individuals, including spillover effects from the consumer sector into government access to privately held data, as well as increased data security risks. They frame DM as a human rights issue, noting that "the most important impact of strong data minimization is harm reduction: data that is not collected cannot harm people."[25] Consumer Reports, in their whitepaper in support of DM, also points out that in addition to the harms that can result from secondary uses of data, "unwanted observation, through excessive data collection and use, is harmful in and of itself."[26]

We question both what "progress" entails with regard to ADM systems based on or utilizing personal data, as well as the assumption that the collection and use of greater amounts of data always leads to improvement.[27] Promoting the development of ADM systems without limitations

[25] "Data Minimization: Key to Protecting Privacy and Reducing Harm" (2022).
[26] "How the FTC Can Mandate Data Minimization Through a Section 5 Unfairness Rulemaking," Electronic Privacy Information Center. Accessed November 2022 via: https://epic.org/documents/how-the-ftc-can-mandate-data-minimization-through-a-section-5-unfairness-rulemaking/.
[27] For clarity, we refer to ADPPA's definition of personal information: "information that identifies or is linked or reasonably linkable, alone or in combination with other information, to an individual or a device that identifies or is linked or reasonably linkable to an individual, and may include derived data and unique identifiers." H.R.8152 - American Data Privacy and Protection Act. Accessed November 2022 via: https://www.congress.gov/bill/117th-congress/house-bill/8152/text#H46F71E2465BC4107A19A53B8A30F30D3.

on the type of data used or the scope of data collected, and without articulation of a clear purpose leaves a pathway open for developing systems that maximize the collection of data for data's sake: a race to quantify and capture as many forms of personal experience as possible without regard to the consequences. This isn't conjecture; tech companies are already demonstrating interest in building systems that attempt to anticipate all of an individual's wants and desires, across experiential contexts.[28] Today, a large suite of ADM-driven products, ranging from smart home devices and voice assistants to wearable health-monitoring gadgets, collect and leverage extraordinarily large amounts of personal data to anticipate users' needs and nudge their behavior in nearly all aspects of their everyday lives.[29]

The unconstrained development of ADMs using personal data could lead to ever more massive personal surveillance networks, across devices and contexts, in public and private, in an attempt to gather as much data from as many facets of people's lives as possible. Even if the results produced were highly predictive, one must question whether the costs of potentially unlimited surveillance are worth the benefits. Without PL regulation, there would be little assurance that that data collected in one context (e.g., real-time location data to provide directions) would not be used in another (to build a profile of an individual's movements across time for marketing purposes).

Twenty-five years after the emergence of the commercial internet, and nearly 10 years into the embrace of "big data," in our shift toward AI we face another technology that is dependent on yet more massive amalgamations of data. Accordingly, this shift comes with another round of assertions that innovation should be allowed to proceed unchecked, lest we quash the next new breakthrough.[30] However, critics are raising questions regarding the rapid speed of progress in some AI domains, and whether this progress represents an unalloyed good.[31] Even in domains that do not utilize personal data, the pace of progress in AI is causing some experts to voice concern

[28] Justin De Graaf, "How Consumer Needs Shape Search Behavior," Think with Google (May 2019). Available at: https://www.thinkwithgoogle.com/marketing-strategies/search/consumer-needs-and-behavior/.
[29] Chris Gilliard, "The Rise of 'Luxury Surveillance,'" The Atlantic (October 18, 2022). Available at: https://www.theatlantic.com/technology/archive/2022/10/amazon-tracking-devices-surveillance-state/671772/; Tanya Basu, "Amazon Has a New Plan for Its Home Robot Astro: To Guard Your Life," MIT Technology Review (September 28, 2022). Available at: https://www.technologyreview.com/2022/09/28/1060418/amazon-wants-astro-to-guard-your-life/.
[30] Kate Kaye, "Why an 'Us vs. Them' Approach to China Lets the US Avoid Hard AI Questions," Protocol (November 8, 2022). Available at: https://www.protocol.com/enterprise/us-china-ai-fear-military.
[31] Kevin Roose, "We Need to Talk About How Good A.I. Is Getting," The New York Times (August 24, 2022). Available at: https://www.nytimes.com/2022/08/24/technology/ai-technology-progress.html.

over this speed lest our haste cause us to barrel into computational accomplishments that bring with them inadvertent yet avoidable damage to society. The FTC should not embrace this hype, nor allow for the unrestricted exploitation of personal data, particularly when the gains for the data economy thus far appear to benefit large data holders to the enormous detriment of consumers. If we do allow personal data to be leveraged to build large-scale ADM systems, it should be within the constraints of DM and PL and more specifically to benefit individuals, as we discuss in the next section.

3. Shifting the Landscape of Data Acquisition and Governance

The questions raised by the FTC throughout the ANPR demonstrate an interest in putting a check on the practices that perpetuate surveillance capitalism. However, while introducing DM and PL would have an impact on these practices, these principles alone do not address underlying inequities and unfairness in the data ecosystem today. In order to address the existing market conditions that contribute to these problems, as well as the disproportionate advantages that have accrued to its winners, the FTC should support, through regulation, a data ecosystem that moves consumer control of personal data away from platforms and other first-party collectors toward one that supports the licensing of data through data intermediaries. We believe such a move would give consumers more direct benefits from commercial uses of their data. Additionally, the FTC, as well as the federal government more generally, must support the development of new public data infrastructures and resources to open the playing field for data use to a broader set of actors, and change incentive structures such that businesses are no longer incentivized to collect as much personal data as possible. We believe these measures are crucial not just to address the inequities caused by the present surveillance capitalist data ecosystem, but also to specifically support more ethical, fair, just, and privacy-preserving development of data-centric systems, including those utilizing ADM and AI.

Today, individual consumers have little control over the collection, use, and reuse of their personal data. Enacting DM and PL obligations on top of the existing data ecosystem would theoretically force companies to be more thoughtful and conservative in their data collection practices, and ideally translate to less personal data collected from consumers. However, these obligations will not curb companies' desire for massive amounts of data, especially for data-centric applications

of AI. The current lack of data regulations contributes to a condition where quantity (collecting free or low cost data at scale by any means necessary) supplants quality (reliable data, fit for purpose, collected ethically). To encourage a shift toward making higher quality, ethically sourced data the norm, we recommend the FTC consider both enacting regulations and supporting policy that explore several possible solutions to this conundrum: investing in public (rather than proprietary) datasets; creating and incentivizing data intermediaries; and developing best practices for filtering data for privacy violations.

DM and PL principles may appear to conflict with the demands of emerging AI technologies dependent on ever-increasing amounts of training data, but we believe that it is possible to use sources other than consumer products to create useful datasets, even in commercial contexts. Speech recognition technology is a classic example of a problem that was long believed to require data gathered from consumer products for other purposes. In 2007, Google publicly stated that they set up their 411 phone search service to collect consumer voice data in order to help develop other products.[32] This type of out-of-context data reuse still occurs in large companies, but recent advances by OpenAI,[33] a company without consumer product offerings, have demonstrated state of the art performance using large amounts of data gathered from across the web.[34] This, and similar advances in other problem domains,[35] is evidence that open data is already a better source for many if not most model training requirements, though we caution that we are not advocating for widespread data scraping as a solution, as we elaborate below. This conclusion follows logically from the demand for ever-larger datasets to improve the quality of ML models, since private data from any one product, or even a set of products under one corporate roof, is not as large as the aggregate of public data sources.

[32] Juan Carlos Perez, "Google Wants Your Phonemes," InfoWorld (October 23, 2007). Available at: https://www.infoworld.com/article/2642023/google-wants-your-phonemes.html.
[33] OpenAI does not have consumer products from which to harvest data. Their image generation model DALL_E, for example, was trained from images scraped from the web as well as from licensed data. However, this approach has demonstrably produced images that are biased. See: Khari Johnson, "DALL-E 2 Creates Incredible Images—And Biased Ones You Don't See," Wired (May 5, 2022). Available at: https://www.wired.com/story/dall-e-2-ai-text-image-bias-social-media/.
[34] See: https://cdn.openai.com/papers/whisper.pdf. Note that many of the most visible examples of large-scale AI systems are built upon "public" data (data scraped from public-facing websites), and while not immune from privacy concerns, they predominantly do not use private, personal data.
[35] Manohar Paluri et al., "Advancing State-of-the-Art Image Recognition With Deep Learning on Hashtags," Engineering at Meta (May 2, 2018). Available at: https://engineering.fb.com/2018/05/02/ml-applications/advancing-state-of-the-art-image-recognition-with-deep-learning-on-hashtags/

Our belief is that commercial and research groups can encourage this trend by creating options that shift data aggregation away from large platforms toward public or nonprofit solutions, such as building shared public datasets for common problem domains. For example, an industry and academia consortium like MLCommons,[36] or an open project like Mozilla's Common Voices,[37] create datasets for voice data, using public sources where participant consent is clearly obtained. This would allow any interested party to replicate OpenAI's speech training, with the expectation that added data or improved labeling would be contributed back to the public. Over time open resources like these will become increasingly more valuable, in turn creating incentives to move away from privacy-sensitive proprietary product data, especially for small to medium enterprises, governments, or non-profits that can't afford to compete with large platforms. Investment in other public data resources, such as those proposed by Stanford in response to the National Artificial Intelligence Research Resource (NAIRR) Task Force, is another means by which to support investment in public datasets.[38]

Another method to shift the data economy toward one that is more equitable and consent-oriented is to create regulatory frameworks and technical standards that support the development of data intermediaries that manage personal data on behalf of consumers. Data intermediaries, broadly construed, provide a way for consumers to maintain control over their data while sharing it in ways consonant with their desires and needs. Rather than relying on a model where access to "free" services is "paid" by giving up personal data directly to a service, a system that utilizes intermediaries would allow consumers to license their data as needed, and potentially reap direct benefits from doing so, rather than giving up control. A key benefit for businesses would be the ability to access data from willing consumers using a trusted infrastructure that meets compliance needs for consent and privacy.[39] A data infrastructure built upon data exchange and licensing rather than data extraction and exploitation would also dilute the current monopoly power over data that big platforms hold, and allow greater access, albeit potentially at a cost, to a broader range of

[36] https://mlcommons.org/en/
[37] https://commonvoice.mozilla.org/
[38] "National Research Cloud," Stanford Institute for Human-Centered Artificial Intelligence. Available at: https://hai.stanford.edu/policy/national-research-cloud.
[39] "Advancing Digital Agency: The Power of Data Intermediaries," World Economic Forum. Accessed November 2022 via: https://www.weforum.org/reports/advancing-digital-agency-the-power-of-data-intermediaries.

actors.[40] With regulatory support, the mitigated risks and compliance benefits of such a system could make "free" data a liability rather than an asset.

Finally, another practical measure the FTC can support are best practices for the technical processing of personal data, such as mandating the use of automated filters for sensitive content before any data is even stored in a dataset. Examples include running address, name, credit card, and email detectors like Presidio[41] on text to remove items that might leak identity information, and faces, license plates,[42] or sexual content analysis tools[43] on images to avoid such content from being captured. We know that unintended sensitive content occurs even in current well-known public datasets,[44] which is one reason we need to be more careful about ensuring meaningful consent from contributors. As a first step, we could ask that any datasets, either public or private, used commercially for AI training are run through these existing tools to generate a report on the likely amount of problematic data, along with an explanation of the failure that caused such data to be included. Once there is confidence that we have established best practices, it would seem natural to mandate that these tools must be run with a low threshold set for acceptable sensitive data rates. Including private data should be treated as a bug, with a requirement to remove it and ensure that processes are in place to ensure it does not get included in the first place.

Responses to specific questions in Section III:

Q38: In response to Q38, we recommend that the FTC limit commercial surveillance practices, and in particular apply the aforementioned purpose limitation principles to biometric data. Widespread repurposing of biometric data outside the initial context of collection will contribute to the proliferation of surveillance technologies, threatening privacy and individual autonomy.

Q43/Q44: Our responses above frame our overall answer to Q43: ADM should not be used as an excuse not to embrace data minimization and purpose limitation principles. But Q44 also brings up an important issue related to the length of data collection: Should privacy by design include data

[40] Jennifer King, "Advancing the Case for Data Intermediaries," Stanford Institute for Human-Centered Artificial Intelligence (February 16, 2022). Available at: https://hai.stanford.edu/news/advancing-case-data-intermediaries.
[41] https://github.com/microsoft/presidio
[42] https://github.com/understand-ai/anonymizer
[43] https://github.com/GantMan/nsfw_model
[44] https://twitter.com/dribnet/status/874389135577853952

deletion by default? We argue that yes, with some exceptions or at least with explicit, opt-in consent, most forms of consumer data should be deleted by default after a reasonable time period. For example, one of the authors of these comments discovered accidentally that Amazon still maintained detailed purchase information about orders she placed in 1999, including delivery and billing addresses, and even long-expired credit card information. Even granting some need for historical data to inform good data-security practices, there is no justifiable reason to keep detailed customer order data forever, and particularly with no option to expunge it if desired. One of the biggest threats to privacy from data collection results from aggregation over time; without clear limits on what constitutes disproportionate data collection, companies will continue to be able to produce dossiers of our entire lives, notably children from age 13 into adulthood.

Q45/Q46: Regarding Q45, we caution that determinations of purpose limitation should not be based on notice and consent. Companies must not be able to provide vague, indeterminate descriptions of data use in privacy policies as a means by which purpose limitation is retrospectively assessed. Q46 suggests a potential sectoral approach to this problem; one might also consider uses of sensitive data, or uses of data with high risk or legal impacts. These delimiters may help to avoid overregulating uses of personal data where there is little risk to repurposing it, within reason.

Q50/Q52: We are all already bearing the costs of the fallacy of "free" services. We would direct the FTC to our earlier comments regarding supporting data intermediaries, which would provide more direct access, control, and individual benefits to consumers rather than the current model of direct data collection by individual sites. We urge the FTC to proactively consider the consumer benefits of supporting a data ecosystem that does not put individuals at the mercy of data collectors, but instead gives individuals an actual stake in the use of their data and meaningful negotiation power, rather than merely the power to vote with their feet. As the status quo demonstrates, there are many examples where data monopolies make it difficult if not impossible for individuals to find reasonable alternatives to many online services. In response to Q52, moving to such a system would require going beyond mere data interoperability, and ideally obviate the need for tracking through persistent identifiers. Consumers' intermediaries would handle service-by-service negotiations, with the power to reject data collection entirely, or offer data exchange upon mutually agreed-upon conditions.

Section IV: Automated Decision-Making Systems

Most consumers already experience automated decision-making systems when they are shown personalized ads and content online. Users whose preferences are more "typical" are likely to have experiences where recommendations match their preferences because companies tend to have greater amounts of data to train their systems that match to similar users. Algorithms are usually beneficial for companies in one of three areas: They save time—faster than manual adjudication; they improve consistency as similar cases tend to be treated similarly by the algorithms; and they improve accuracy if the algorithms are trained on high-quality data.

While errors are inevitable in most settings as it is difficult to develop algorithms that never make mistakes, their failures have costs for both consumers and the companies that deploy them. We believe it is important to consider the cost-benefit tradeoffs of deploying algorithms as well as implementing rules to evaluate and mitigate algorithmic errors.

Three factors are important when evaluating cost-benefit tradeoffs. First, do algorithms, even if imperfect, improve over existing practice? For example, human feedback can be used to improve performance over time, weeding out bias.[45] Second, can algorithms be used together with human supervision to mitigate errors? Algorithms relying on human monitoring, a practice referred to as Human-in-the-Loop, is based on creating machine learning models that, in case of uncertainty or in contexts with a low level of confidence, yields the decision making to a human mind.[46] Human judgment can be valuable when it detects negative payoffs for a high-risk algorithmic decision, helping to avoid harmful outcomes.[47] Third, are algorithms being used for high- or low-risk decisions, where errors have significant legal or health and safety impacts?

The best way to measure algorithmic error is to have a well-annotated test dataset that is sufficiently large and diverse to represent likely scenarios that the algorithm could encounter in

[45] Nicole Turner Lee, et al., "Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms," Brookings (May 2019). Available at: https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/.
[46] Kerstin Dautenhahn, "The Art of Designing Socially Intelligence Agents: Science, Fiction, and the Human in the Loop," Applied Artificial Intelligence (June 2011). Available at: https://doi.org/10.1080/088395198117550.
[47] Ajay Agrawal et al., "Exploring the Impact of Artificial Intelligence: Prediction Versus Judgment," Information Economics and Policy (June 2019). Available at: https://doi.org/10.1016/j.infoecopol.2019.05.001.

practice.[48] It is also important to update test data over time so it stays fresh and relevant. Algorithmic "mistakes" may be more common in the consumer sector (e.g., recommending products) due to the high heterogeneity in the users, but depending upon the application these mistakes can be less critical. The FTC should consider working with the National Institute of Standards and Technology (NIST) as well as civil society organizations to create test datasets against which companies can evaluate their performance. This is particularly critical as the FTC attempts to evaluate claims by companies regarding effectiveness or accuracy; an AI-based service that makes claims to specific levels of effectiveness or accuracy should be able to defend those claims against established and vetted benchmarks. Ultimately, any company that makes claims about accuracy or effectiveness of an algorithmic system must be prepared to provide a means by which those claims can be evaluated by outside parties.

Accordingly, ideal conditions for evaluating algorithms are application-specific.[49] The FTC should consider establishing a clear guideline of low- versus high-risk applications of algorithms and target stricter regulations for high-risk settings (e.g. in healthcare, finance, education, etc.). It is difficult to broadly state what the acceptable level of accuracy algorithms need to meet to be used. For example, one could compare the performance of humans without algorithmic support and the performance of algorithms with potential human supervision. For example, if the latter demonstrates better performance and has additional benefits such as improved consistency, time savings, and lower bias relative to humans, then it may be acceptable.

Companies can mitigate algorithmic errors by having human experts provide the final supervision after algorithms make the initial triaging decisions. Another approach is to have standards of monitoring algorithmic performance over time, and to alert users if the algorithm starts to make more "wrong" decisions, for example, due to changes in the users' data.

Section V: Discrimination Based on Protected Categories

[48] Jiefeng Cjem et al., "Detecting Errors and Estimating Accuracy on Unlabeled Data With Self-Training Ensembles," Advances in Neural Information Processing Systems 34 (2021): 14980-92. Available at: https://proceedings.neurips.cc/paper/2021/hash/7dd3ed2e12d7967b656d156d50308263-Abstract.html.
[49] For example, see Daniel E. Ho et al., "Evaluating Facial Recognition Technology: A Protocol for Performance Assessment in New Domains," Stanford Institute for Human-Centered Artificial Intelligence (November 2020). Available at: https://hai.stanford.edu/sites/default/files/2020-11/HAI_FacialRecognitionWhitePaper_Nov20.pdf.

To measure and evaluate algorithmic discrimination, the FTC should establish clear metrics of consumer utility in a specific setting or application and then quantify the metrics if the utility of an underrepresented group suffer due to algorithmic decisions. For example, researchers have developed fairness metrics to determine that the AI-driven approach currently used in health insurance marketplaces significantly affects the pricing of healthcare insurance and harms undercompensated groups.[50] In addition, standards for evaluating algorithms relative to the current practice are also crucial to address algorithmic discrimination. When conducting such evaluations, it is useful to look at fine-grained groups beyond the broad demographic categories. The FTC should understand algorithms' performance and potential biases for underserved groups, such as rural communities or unhoused populations, while prioritizing sectors that Congress has explicitly legislated, especially since rules likely need to be modified for different sectors.

For more information on the effect of algorithmic discrimination and appropriate responses, please refer to the Stanford HAI letter submitted in January 2022 in response to the White House Office of Science and Technology proposal for an AI Bill of Rights that safeguards the American public against powerful technologies.

In closing, we appreciate the opportunity to provide comments to the Federal Trade Commission on these crucially important issues.

Sincerely,

Dr. Jennifer King, Privacy and Data Policy Fellow, Stanford HAI
Caroline Meinhardt, Graduate Student, Stanford International Cyber Policy Program
Abel Ribbink, Graduate Student, Department of Computer Science, Stanford University
Pete Warden, PhD Candidate, Department of Computer Science, Stanford University
Daniel Zhang, Policy Research Manager, Stanford HAI
James Zou, Assistant Professor, Biomedical Data Science and, by courtesy, of Computer Science and of Electrical Engineering, Stanford University

[50] Anna Zink, Thomas G. McGuire, and Sherri Rose, "Balancing Fairness and Efficiency in Health Plan Payments," Stanford Institute for Human-Centered Artificial Intelligence (November 2022). Available at: https://hai.stanford.edu/policy-brief-balancing-fairness-and-efficiency-health-plan-payments.