THE ISSUES OF DISABILITY INFERENCE IN THE PERSONAL AND CONSUMER DATA ECOSYSTEM

Christian J. Rozolis[†]

ABSTRACT

This Article argues for a new approach to consumer data protection that includes specialized protection of individual disability information. The consumer data ecosystem consists of multiple actors that use data across multiple layers of transfers, from collection, to aggregation, to applied data-driven insights. Within this data system, markers of disability exist implicitly and explicitly in the data, raising the risks of data-driven discrimination. The Article first explores the three layers of the consumer data ecosystem, before situating disability status within it as a unique feature that is unlike other forms of personally identifiable information. Because disability status is a mutable trait, and because the Americans with Disabilities Act recognizes associational disability discrimination through its "regarded as" prong, there are unique data regulatory approaches that should be employed to better prevent the risks of data-driven discrimination. The Article next examines existing interlocking data protection laws and how they intercede at different points of the commercial data ecosystem, in contexts like venue ticket sales, housing, credit decisions, and criminal sentencing. Finally, this Article concludes by arguing for new regulatory approaches to protect an individual's disability data. Basic adjustments, such as recognizing disability as a protected trait in credit decision-making or limiting distribution of any necessarily collected disability identifications, are simple ways for existing regulatory frameworks to be used to enhance data protections for individuals with disabilities.

TABLE OF CONTENTS

DOI: https://doi.org/10.15779/Z38HQ3S09C

† J.D. Candidate, 2025, University of California, Berkeley, School of Law. Prior to attending law school, Christian was a data scientist with experience in credit modelling, supply chain forecasting, product recommendation systems, and large language model building.

I'd like to extend my gratitude to Silvia Yee, Elizabeth Zirker, and Arlene Mayerson who encouraged me to bring this paper to life, to David Koeller who listened to my ideas as I drafted this piece, to my family who has supported me through the years, and to my wonderful partner Vicky Ho whose love and support lifts me up every day.

I. INTRODUCTION

Data analytics is an industry most people do not think twice about. With the campaign to rebrand data analytics into data science, machine learning, AI, and countless other buzzwords, consumers have grown accustomed to their data being collected by companies for use in marketing, personalization, and product recommendations. From Instagram likes to website clicks, there is a massive amount of data floating between companies every day in a virtual panopticon. Big Data, historically defined by the "Three V's" (Volume, Velocity, and Variety), has exploded in recent years, with data scientists enumerating as many as 39 other Vs to characterize their datasets.[1] Plainly, the consumer data ecosystem is thriving.

Of course, in a thriving ecosystem of symbiotic data relationships between companies, data has become a huge source of profits for companies. For example, even though tech giants like Meta or Google offer free use of their platforms to consumers, they turn profits by selling consumer data to advertisers trying to gain access to those consumers.[2] Identifying the sources

1. Posting of Tom Shafer to KDnuggets, *The 42 V's of Big Data and Data Science* (2017), https://www.kdnuggets.com/2017/04/42-vs-big-data-data-science.html.

2. Scott Goodson, *If You're Not Paying For It, You Become The Product*, FORBES (Apr. 14, 2022), https://www.forbes.com/sites/marketshare/2012/03/05/if-youre-not-paying-for-it-you-become-the-product/ (noting Google's "main source of income" in 2022 was its

of a data problem is a challenge when there are so many responsible actors in the data-trading space. As more and more data are compiled and combined, the original data sources and measurements lose clarity and transparency. How exactly does a social media company know a user is a dog-lover? Does this knowledge come from the user's screen engagement, advertisers' reports, or perhaps from third-party purchasing patterns data? Most internet users will receive personalized content that could be derived from many sources. But what happens when a person is being implicitly discriminated against by their own data?[3] Some argue that algorithms, which are fed by data, are not discriminatory; rather, their underlying datasets reflect the bias and patterns of the real-world they measure. Others argue that because of this underlying real-world bias, more data safeguards should be built into data systems so as to not perpetuate discrimination in a self-fulfilling prophecy.

This Note will argue that disability status, whether it's an explicit or implicit data point, warrants specialized protections in the consumer data ecosystem because of the risks of disability discrimination inherently present in consumer data datasets. Part II will define and illustrate the consumer data ecosystem and then explain how data moves and evolves within it, before exploring the reasons that disability status is uniquely vulnerable to data-driven discrimination. Part III will explore how current law, state and federal, responds to the consumer data ecosystem, and how disability fits within it. Finally, Part IV will synthesize the two Sections to make modest legal proposals to begin effectively protecting disability status as a data point.

II. THE CONSUMER DATA ECOSYSTEM

At the outset, it is helpful to orient oneself within the consumer data "ecosystem." In general, there are three interacting layers of consumer data use: (i) collection and initial storage, (ii) harvesting and brokerage, and (iii)

Advertising revenue stream which yielded $38 Billion annually); Emily Sherman, *The Real Cost of Free Apps and Services*, U.S. NEWS (Jan. 25, 2024), https://money.usnews.com/money/personal-finance/family-finance/articles/real-cost-of-free-apps-and-services (identifying "Data Collection" and sales to marketers are one of three "costs" associated with free services).

3. *See* Noa Yachot, *Your Favorite Website Might Be Discriminating Against You*, ACLU (June 29, 2016), https://www.aclu.org/news/privacy-technology/your-favorite-website-might-be-discriminating-against-you (identifying examples of algorithmic discrimination where big data contributed to racially discriminatory car insurance policies, advertisements for payday loans, and search engine results); *see also* Press Release, NTIA, Off. of Pub. Affs., NTIA Launches Inquiry on how Data Practices Affect Civil Rights (Jan. 18, 2023), https://www.ntia.gov/press-release/2023/ntia-launches-inquiry-how-data-practices-affect-civil-rights (announcing a Request for Comment on how data practices can lead to discriminatory acts using job advertising, dating profiles, and other online user activity).

insight and application. Personal and commercial data is the resource that flows between the layers. Importantly, a single party can control all three layers, or a party may hand off data to another party between the layers in a sale or mutually beneficial business arrangement.[4]

A. LAYERS OF THE CONSUMER DATA ECOSYSTEM

Understanding these layers is important because personal data is not only utilized within them, but also easily moves between them. The first layer ("Layer One") consists of initial data collection. Data collection can be thought of as one of two forms: active or passive.[5] Active data collection occurs when data is specifically and explicitly requested from a consumer. For example, a hotel requests a customer's phone number when the customer signs up for a rewards program or a hotel reservation system presents a checkbox for voluntary disclosure of disability when a customer books a room. Passive data collection, on the other hand, is completed by observation with no real voluntary or interactive input.[6] Examples include a manufacturer collecting a car's terabytes of daily video and velocity data or a department store tracking a phone's movement via Wi-Fi to discern walking patterns.[7] Passive data collection has exploded in recent years on the internet as tech companies have started selling website tracking codes that can produce advertising insights.[8]

4. For example, Company A sells data to Marketing Consultancy B, who, in exchange for the data and rights to keep it, will reciprocally share valuable marketing insights back to Company A.

5. *See Differences in Human Data Capture: Active & Passive*, TRIASSIC (June 13, 2024), https://trinsicdata.com/blog/differences-in-human-data-capture-active-passive; Valeh Nazemoff, *Leveraging Passive and Active Data Using AI in Your Digital Business*, LINKEDIN (Feb. 17, 2022), https://www.linkedin.com/pulse/leveraging-passive-active-data-using-ai-your-digital-valeh-nazemoff/.

6. Some may argue that "Terms of Use" and "Consent" agreements make this collection voluntary, but a reasonable person is unlikely to read the fine print assuming it contains provisions on data sharing. *See* Geoffrey A. Fowler, *I Tried to Read All My App Privacy Policies. It Was 1 Million Words.*, WASH. POST (Mar. 31, 2022), https://www.washingtonpost.com/technology/2022/05/31/abolish-privacy-policies/ (noting that nobody feels in control of nor reads terms of use policies).

7. *See* Posting of Ashkan Soltani to Fed. Trade Comm'n: Tech. Blog, *Privacy Trade-Offs in Retail Tracking* (Apr. 30, 2015), https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2015/04/privacy-trade-offs-retail-tracking (describing privacy concerns of companies like Nordstrom or Philz Coffee when they discovered consumer tracking in their stores); *see also* Michele Bertoncello, Christopher Martens, Timo Möller & Tobias Schneiderbauer, *Unlocking the Full Life-Cycle Value from Connected Car Data*, MCKINSEY & CO. (Feb. 11, 2021), https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/unlocking-the-full-life-cycle-value-from-connected-car-data (detailing the successful business practices of car manufacturers that can access "1 to 2 terabytes of raw data per car" daily).

8. META PIXEL, https://www.facebook.com/business/tools/meta-pixel (last visited Nov. 27, 2023); GOOGLE ANALYTICS, https://marketingplatform.google.com/about/

This explosion in part led to the popularization of the term "surveillance capitalism,"[9] a crude description for a company's profit-motivated practice of mass collection of consumer and behavioral data.

As more data is collected from disparate sources, the second layer ("Layer Two") takes shape: harvesting and brokerage. Data brokers are data hoarders: they collect and collate as much information as they can from a wide variety of sources.[10] The data they collect could be years old or it could be a day old, but what is most important to the brokers is that the data is identifiable and voluminous. Data collection in Layer One is often structured so it can be effectively shared with data brokers in Layer Two. That structure is dependent on unique consumer identifiers. For most people, a social security number, a cell phone number, or an email address will function as a unique identifier. Other identifiers, like a landline phone number or desktop IP address, are semi-unique, but may suffice for data collectors depending on how they intend to use them.[11] Layer Two data brokers buy up disparate data sources and sew them together with these unique or semi-unique identifiers. The value of the complete, collated dataset is largely dependent on identification of unique consumers. While an anonymous dataset may be helpful to define macro-trends (i.e., "Converse sales are rising"), individualized data can help refine and elucidate the trends (i.e., "for 18- to 25-year-olds that retail shop at Urban Outfitters, Nike, and Journeys").

Finally, once enough data is collated, it is moved to the insights and action layer ("Layer Three"). Data brokers may work in both Layer Two and Three themselves or they may sell their datasets to a variety of buyers, like a corporate in-house marketing department, who will extract their own insights from the datasets. The technologies used to identify these insights have flashy names like "machine learning" or "AI," and they are frequently described as "black-

---

analytics/ (last visited Nov. 27, 2023); LINKEDIN INSIGHT TAG, https://business.linkedin.com/marketing-solutions/insight-tag (last visited Sept. 28, 2024).

9. John Laidler, *High Tech is Watching You*, HARV. GAZETTE: BUS. & ECON. (Mar. 4, 2019), https://news.harvard.edu/gazette/story/2019/03/harvard-professor-says-surveillance-capitalism-is-undermining-democracy/.

10. Thorin Klosowski, *Big Companies Harvest Our Data. This is Who They Think I Am*, N.Y. TIMES: WIRECUTTER (May 28, 2020), https://www.nytimes.com/wirecutter/blog/data-harvesting-by-companies/; Yael Grauer, *What Are 'Data Brokers,' and Why Are They Scooping Up Information About You?*, VICE: MOTHERBOARD (Mar. 27, 2018), https://www.vice.com/en/article/bjpx3w/what-are-data-brokers-and-how-to-stop-my-private-data-collection.

11. Phone numbers present a unique challenge in data matching. Many people today rely on a single, non-shared cell-phone number. However, landlines often represent a shared household of people. If customer rewards data is linked by phone number, it may be limited to representing a group of people (a household) rather than an individual. People with a shared IP address on a computer, likewise, may be targeted for advertisements at a household level.

boxes" when their results seem confounding to developers.[12] But data scientists almost always have a solid understanding of the inner mathematical functions of most advanced machine learning models, even if the results are confounding.[13] Large datasets, coupled with machine learning, enable companies to learn many things about their customers: affinity for fiber-based snacks, likelihood of pregnancy, or reading burnout rates.[14] When an insight becomes valuable, a company will act on it; this may be a targeted coupon campaign that encourages spending, personalized product recommendations, or a full redesign of an ecommerce site to facilitate customer retention.[15]

B. HOW DATA EVOLVES THROUGH THE LAYERS

As data flows across the layers from consumer to company, data profiles are born. If a company cannot identify its customers individually, the most valuable insights it can generate will likely only describe general customer profiles or the characteristics of an average or median customer. But the days of a "median customer" are gone, and today, personalization has taken center

12. *See* Lou Blouin, *AI's Mysterious 'Black Box' Problem, Explained*, UNIV. OF MICH. DEARBORN NEWS (Mar. 6, 2023), https://umdearborn.edu/news/ais-mysterious-black-box-problem-explained (outlining the practical problems of decoding confounding results, for example when "an autonomous vehicle strikes a pedestrian" when it should have avoided the pedestrian); *see also* Ian Moura, *Addressing Disability and Ableist Bias in Autonomous Vehicles*, DREDF (Nov. 7, 2022), https://dredf.org/addressing-disability-and-ableist-bias-in-autonomous-vehicles-ensuring-safety-equity-and-accessibility-in-detection-collision-algorithms-and-data-collection (describing the issues of disability recognition for autonomous vehicles, including the inability to correctly process a "wheelchair user who moved in a 'non-standard' way.").

13. *See* Cynthia Rudin & Joanna Radin, *Why Are We Using Black Box Models in AI When We Don't Need To?*, HARV. DATA SCI. REV. https:// hdsr.mitpress.mit.edu/pub/f9kuryi8/release/8 (Nov. 22, 2019) (arguing that explainable model functions should be used over "complicated functions of the variables that no human can understand how the variables are jointly related to each other to reach a final prediction.").

14. Klosowski, *supra* note 10; Kashmir Hill, *How Target Figured Out a Teen Girl Was Pregnant Before Her Father Did*, FORBES (Feb. 16, 2012), https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/?sh=59c74d376668; Kari Paul, *'They Know Us Better Than We Know Ourselves': How Amazon Tracked My Last Two Years of Reading*, GUARDIAN (Feb. 3, 2020), https://www.theguardian.com/technology/2020/feb/03/amazon-kindle-data-reading-tracking-privacy.

15. Hill, *supra* note 14; Edin Sabanovic, *How to Use Google Analytics to Improve Your Web Design Projects*, SHOPIFY PARTNERS (June 13, 2017), https://www.shopify.com/partners/blog/google-analytics-to-improve-web-design-projects; *see generally* Pushpendra Kumar & Ramjeevan S. Thakur, *Recommendation System Techniques and Related Issues: A Survey*, 10 INT'L. J. INFO. TECH. 495 (2018).

stage.[16] With individually identifiable customer data, companies have begun building what are known as affinity scores and groups.[17] These scores infer somebody's probable affinity for almost anything—dog-lover, "sad teen," or afternoon indie-music enthusiast[18]—and can be used to target advertising or sell consumer insights en masse.

While all these algorithmic affinity groups are derived as insights in Layer Three, they ultimately reflect the data collected and combined in Layers One and Two. Take the affinity group dog-lover, for example. Layer One collections of data for that group could come from Instagram, Chewy.com, and a local park district. Suppose the park district data is five years old, possibly due to a municipal policy restricting the sale of more recent data to marketers, and that the park data provides the number of registered dog licenses per home address. When a Layer Two broker collects the park district data, they know it carries a heavy implication of pet ownership. However, due to the age of the data, some dogs may no longer be alive, or households may have moved. Thus, the broker may combine Chewy delivery data and find regular buying patterns of food and grooming supplies for a select group of people. Assuming there are thousands of records, and other variables of value, the broker can validate residents that are most likely to currently own a dog. This becomes a Layer Three insight that the broker can then leverage with the Instagram data—they can determine which of these users engage the most with advertising and dog content. As a result, the broker can instruct a Chewy competitor like PetSmart to efficiently buy advertising space on the social media feeds of specific individuals.

But what if a data broker is also in the business of selling their insights to landlord companies? It is not hard to imagine a landlord that wants to exclude a dog owner from their buildings given the noise and physical nuisance dogs can create. They could use this dataset to limit who they advertise apartments to or even crosscheck the emails tied to dog-lovers with the emails of

16. *See generally* Nidhi Arora, Wei Wei Liu, Kelsey Robinson, Eli Stein, Daniel Ensslen, Lars Fiedler & Gustavo Schüler, *The Value of Getting Personalization Right—or Wrong—is Multiplying*, MCKINSEY & CO. (Nov. 12, 2021), https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/the-value-of-getting-personalization-right-or-wrong-is-multiplying.

17. Sandra Wachter, *Affinity Profiling and Discrimination by Association in Online Behavioral Advertising*, 35 BERKELEY TECH. L.J. 367, 369 (2020).

18. *See* Michael Reilly, *Is Facebook Targeting Ads at Sad Teens?*, MIT TECH. REV. (May 1, 2017), https://www.technologyreview.com/2017/05/01/105987/is-facebook-targeting-ads-at-sad-teens/; *Get Fresh Music Sunup to Sundown With daylist, Your Ever-Changing Spotify Playlist*, SPOTIFY (Sept. 12, 2023), https://newsroom.spotify.com/2023-09-12/ever-changing-playlist-daylist-music-for-all-day/ (Spotify's new time-of-day playlist is built algorithmically from a listener's personal listening history).

prospective tenants and filter them out of selection. Unbeknownst to the landlord, the "dog-lover" dataset had no controls for people with service animals. In one long evolutionary chain, a five-year old park district registration could turn into an unknowing inference of disability, namely service animal ownership. Even when a landlord has no intent to discriminate against people with disabilities, they could unknowingly and unintentionally become collateral in the otherwise legally permissible dog-discrimination. Congress considered such "benign neglect" when it passed the Americans with Disabilities Act (ADA).[19]

Just as "benign neglect" permeates society with "thoughtlessness and indifference," it also permeates the consumer data ecosystem.[20] This type of behavior is particularly hard to identify, especially since disability status can enter the data ecosystem without detection. For example, at Layer One, "dog owner registered with the park district," is not an explicit disability marker considering there are likely no distinctions for service animals versus pets in the park district's data. Yet the implicit marker for disability, owning a service animal, managed to pass between Layer Two and Three before ultimately being used for discriminatory purposes—denying housing because of the animal.[21] If an individual was unaware of this violation of their Fair Housing Act (FHA) rights, they may never think to challenge rental rejections as discriminatory. In such circumstances, where a landlord practice is facially permissible, individual instances of discrimination may go unnoticed until somebody alleges a broader pattern of discrimination and makes a case for disparate impact.[22]

Because data moves across layers and often comes from disparate sources, benign attempts to derive insights from it tends to tread on sensitive, personal

19. Alexander v. Choate, 469 U.S. 287, 289 (1985).

20. *See id.*

21. *See infra* Section II.C for discussion on Fair Housing Act obligations; *see also Assistance Animals*, U.S. DEPT. HOUS. & URB. DEV., https://www.hud.gov/program_offices/fair_housing_equal_opp/assistance_animals (last visited Nov. 27, 2023) (providing a public-facing explanation of Fair Housing Act obligations regarding assistance animals).

22. For example, a disparate impact analysis may show that the landlord's dog-rejection policy, applied across multiple properties, led to a disproportionate rejection of prospective tenants with disabilities. *See* Nick Adjami, *Disparate Impact: A Crucial Fair Housing Protection Under Attack*, EQUAL RTS. CTR. (Oct. 14, 2019), https://equalrightscenter.org/disparate-impact-under-attack/ (describing how "disparate impact" analysis has been applied in the housing context to demonstrate how neutral housing policies led to illegal discrimination); *see also Title VI Manual, Section VII — Proving Discrimination — Disparate Impact*, U.S. DEPT. OF JUST. CIV. RTS DIV. (outlining how disparate impact claims are litigated and referencing Tsombanidis v. W. Haven Fire Dep't, 352 F.3d 565 (2d Cir. 2003), where a city's fire code was struck down for violating the FHA).

features. Moreover, these types of harms are extremely difficult to detect, especially in the absence of any clear data markers.

C. WHAT MAKES DISABILITY DATA DIFFERENT?

Disability status is uniquely situated as a data feature in the consumer data ecosystem. This is because consumer products and interactions contain implicit disability-coded features in ways that other personal and protected characteristics do not. That is not to say that race, sex, gender, sexual orientation, age, and other protected characteristics do not have their own associated challenges with discriminatory data practices. However, disability status is unique compared to these other characteristics in that it is more apparent in a physical consumer world. Disability status is often explicitly coded in a way that is highly visible and identifiable. When Congress passed the ADA, they included legislative findings that explicitly called out the persistent structural barriers and exclusion of people with disabilities in day-to-day life.[23] The purpose of the ADA was "to provide a clear and comprehensive national mandate" to end disability discrimination.[24] Title III of the ADA was an essential component of that mandate, functionally applying these anti-discrimination principles to private businesses who previously avoided the obligations of Section 504 of the Rehabilitation Act of 1973.[25] Title III brought equal access to public accommodations and the removal of architectural barriers that impact how a person with a disability navigates the world.[26]

Although this mandate commanded equal enjoyment, some practical physical differences necessarily exist for people with disabilities navigating public life. For example, accessible seating options at venues have movable armrests and companion seats, alternative AV options in theaters include closed captioning devices, and some accessible rooms in hotels have light-based rather than sound-based doorbells. Because accessible options for equal enjoyment contain these physical distinctions, Congress recognized a need to balance the requirement for accessibility with the costs of building the accommodations.[27] Not every seat in a stadium is designed accessibly, not every showing in a movie theater has subtitles, and not every hotel room has

23. *See* 42 U.S.C. § 12101(a)(1)–(6).

24. *Id.* § 12101(b)(1).

25. *See* 29 U.S.C. § 794. Private businesses rarely received federal funds that would require Section 504 obligations.

26. 42 U.S.C. §§ 12182–83.

27. *See id.* § 12183 (distinguishing new construction from existing infrastructure, allowing consideration of "cost and scope" in alterations, and clarifying the requirements for larger changes such as elevator installation).

light-based doorbells. However, if somebody regularly engages with a business by using its accessible options, and the business keeps track of its consumer data, the physical distinctions may become trackable data points. Thus, the practical differences and underlying policy judgments of the ADA have an access point into the consumer data ecosystem in ways distinct from other protected classes like race, sex, and gender.

1. Volume and Identification Opportunity

To establish how disability-status is unique in data ecosystems, one needs to consider how Layer One data collection has broad access to affirmatively coded disability data. Congress expansively regulated architectural barriers to disability, with a broad, statutory definition of "public accommodation" that enables disability to be expressly accounted for in a multitude of physical contexts.[28] Layer One data collection has access to this affirmatively coded disability data. The most obvious examples of this are public venue ticketing sales. For example, the Department of Justice (DOJ) has promulgated Title III guidelines for concert halls, movie theaters, sports arenas, and other venues. The guidelines require that these venues offer accessible seating, auxiliary aids, and accessible ticketing options.[29] But a close look at the DOJ guidelines involving ticketing reveals no considerations of data protection.[30] That is, it appears a sports arena could collect and sell data that a fan routinely purchases tickets in their accessible seats. Beyond ticketing, consider hotel room or restaurant reservations, both of which could include affirmative requests for accessible table heights, roll-in showers, or specific parking spots.[31] Consider transit options too: Uber and Lyft can easily identify a user that utilizes their

28. *See id.* § 12181(7) (codifying disability discrimination protection in places like hotels, grocery stores, auditoriums, zoos, museums, and even bowling alleys).

29. *See* 28 C.F.R. § 36.308 (addressing seating in assembly areas); 28 C.F.R. § 36.302(f) (modifications in policies, practices, or procedures for ticketing); 28 C.F.R. § 36.303 (auxiliary aids).

30. Consider, for example, 28 C.F.R. § 302(f)(2), which requires venues to identify available accessible seating upon inquiry when distributing tickets but leaves unclear whether such information can be used in other contexts. And, while the regulations at 28 C.F.R. §§ 302(f)(6)–(7) do address the secondary ticket market, they generally only require that individuals with disabilities be allowed to participate "under the same terms and conditions" as those without disabilities, leaving one wondering how exactly information on accessible seat labels, purchasing patterns, and the associated data may be used.

31. *See* Disability Rts. Educ. & Def. Fund, Comment Letter on Advance Notice of Proposed Rulemaking, Titles II and III of the Americans with Disabilities Act (ADA) CRT Docket No. 113, RIN 1190-AA64, at 55–59 (Feb. 3, 2011), https://www.regulations.gov/comment/DOJ-CRT-2010-0008-0157 (noting that bed renovations and upgrades in lodging and medical settings led to higher, less accessible bed heights, and proposing that the DOJ update its bed height guidelines under the ADA.).

wheelchair accessible vehicle services and public transportation systems like Bay Area Rapid Transit (BART) may be able to identify by a user's transit card that they only utilize station entrances closest to, or only accessible via, the elevator. Even further, consider the use of screen readers on websites that use tracking cookies and proprietary code to collect user activity. The code could recognize routine "keyboard only" activity from a user (rather than mouse usage) which is a strong proxy for screen reader use.[32]

Of course, there are also examples of other non-disability characteristics that can be affirmatively or implicitly collected for consumer categorization purposes. For example, the frequent purchase of tampons or the routine purchase of discounted senior and veteran tickets generally indicate some trait about the consumer. But there are two critical distinctions for disability data: volume and identification. The first distinction is the sheer volume of scenarios in which disability-coding can arise. A savvy Layer Two data harvester could piece together a request for airline boarding assistance, coupled with an accessible rental car, an accessible hotel room, and accessible tickets to a Broadway show all from one weekend. It is far less likely that another personal characteristic such as race or gender would be so routinely and explicitly identifiable in the same chain of consumer interaction. For example, if somebody forgot to pack tampons for their trip, they would purchase them at the hotel or a nearby convenience store. However, that sex-coded[33] data point is a single occurrence in comparison to the four other disability-coded data points across the weekend trip.

The other distinction is the intrinsic identification spectrum that matches a data transaction to a person. Consumer transaction data tied to a specific person impliedly belongs to that person, and disability data is no different. However, disability accommodations often involve an interactive process when a person with a disability must voluntarily disclose disability status in seeking accommodations.[34] Moreover, disclosing one's disability is a socially sensitive and vulnerable experience. Between regulatory requirements for participation and the social stigmas of disclosure, it seems more likely that disability status compared to other data disclosures, is coming directly from the person with a disability, making disability data directly identifiable. Put another way, it seems likely that a person with a disability is the one purchasing

32. These are data points reflecting the world today and do not account for continued inaccessibility. As accessibility grows, it is inevitable that more data points that implicitly represent disability will appear.

33. To be specific, as used here, the "sex-coded" characteristic means a "person with a vagina."

34. *See infra* note 45.

their own tickets for venues accessible under the ADA or making accessible reservations.[35] Other protected classes, while having facially apparent data features, are not always directly identifiable through data habits; they are more susceptible to false positives. For example, a single father could routinely be purchasing tampons for his daughters, or a heterosexual brother could purchase pride merchandise to be an ally for their queer-identifying sibling. In these cases, the purchaser is not the person with the identifiable status or identity inherent in the transaction. To be sure, marketers have gotten their predictions right in these categories before,[36] and nothing is stopping them from refining their datasets. The point is, however, that disability status is much closer to direct identification in consumer data, and thus is likely to have a much lower false positive rate.

2. Staleness and Mutability

Second, disability status can be impermanent, but data collection practices may not account for such nuance.[37] Layer Two data brokerage can amplify disability data even as the explicit labeling of disability falls to the background. Of major concern in Layer Two is the risk of old, stale data continuing to perpetuate discrimination. Ordinarily, personal traits, unlike consumer interests, are thought to be stagnant, but staleness is an industry-wide issue for consumer data because everyone's general tastes change from time to time. Thus, so too will the consumer insights personally tailored to those tastes. The more current a dataset is, the more marketable its insights. The dog-lover example presented above provides an example of staleness—namely the park district data being five years old. To validate which records are not stale, a data harvester must combine more recent data before they can make effective use of the older records.

Disability status has a unique relationship with the concept of "staleness." Notably, enshrined in the definition of disability under the ADA is the "regarded as" prong.[38] This definition protects individuals who have an "actual or perceived" disability, and thus prohibits discrimination against somebody

35. While in some contexts, it is possible that booking or data disclosure is done by a family member (i.e., hotel rooms, restaurants, theaters), other concerns arise about the potential to *associate* the transaction with disability status. *See infra* Section I.C.3.

36. *See* Hill, *supra* note 15.

37. For example, consider an airline that offers preboarding for those who need extra time to walk down the jet bridge. A person with a chronic intermittent physical disability may use this option on one trip, but not for others.

38. 42 U.S.C. § 12102(1)(C).

with a disease in remission or somebody who had a prior disability.[39] Effectively, *prior* disability cannot be used to discriminate in the *present*. Likewise, when a prior data point can be used to discriminate in the future, a stale data point can facilitate continued discrimination. For example, if somebody at one point had a functional limitation requiring accessible seating and then no longer needed accessible seating, their disability-coded ticket purchases could persist within a data harvesters' collection for years. The immutability of historic consumer data patterns stands in tension with how much some disabilities, chronic conditions, and symptoms can fluctuate over time.[40]

This lack of permanence does not guarantee that stigmatization and discrimination cease. People with disabilities can still be subject to harsh discrimination in the absence of actual functional limitations. Few other categories of protected characteristics exhibit the same spectrum of mutability or lack of facially evident indicators.[41] The lack of a consistent public presentation of oneself as disabled or queer often leads to stigmatizing social suspicions of the truth of somebody's identity claims. For people with disabilities, this exacerbates the challenges of self-advocacy, accommodation seeking, healthcare seeking, and more, across a wide swath of daily contexts. Stale data risks perpetuating the same stigmas.

3. Replications of Patterns and the Physical World

Finally, Layer Three interacts with unique disability-coded features generated from a wide variety of sources, including those that project real-world architecture into virtual datasets. Such features may include explicit data indicators like data points noting whenever somebody purchased accessible seat tickets or requested audio assistance devices with their tickets. However, Layer Three data may also contain implicit indicators about real-world structures and relationships that manifest as data *patterns* rather than as clear

39. *Id.* § 12102(3)(A); *see* Sch. Bd. of Nassau Cnty. v. Arline, 480 U.S. 273, 279 (1987) (holding that the "regarded as" definition in Section 504 of the Rehabilitation Act extended discrimination protections to a perceived record of disability for a schoolteacher who was discharged for her record of tuberculosis).

40. Ironically, in determining that disability status is reviewed under the rational basis standard, the Supreme Court came to a generalized conclusion that disability is "immutable." *See* City of Cleburne v. Cleburne Living Ctr. 473 U.S. 432, 438, 442 (1985) (discussing the "immutability" of particular disabilities and how people with disabilities lack political power).

41. Sexual orientation may be the only other category in which there has been a long history of harsh discrimination for a characteristic that can be mutable in an individual and often lacks facially evident indicators. The same stigma surrounding disclosure and veracity (i.e., proving non-visible disability) persists along the sexual orientation dimension as well (i.e., "coming out" or "they are just confused").

data *points*. Imagine a complex interweaving of dozens of data sources coming together in Layer Three: grocery store purchasing, rental car histories, restaurant reservations, Instagram likes, and so forth. The quantity of data points from disparate sources makes identifying the Layer One source of a data insight much harder than for a single dataset. For example, is disability-status identifiable alone from a rental car history of paratransit van rentals? What if the data is then combined with a restaurant reservation website's data showing a history of disability disclosures for an accessible table? Or consider a video streaming provider has data on closed caption usage—would that alone suffice to indicate a hearing disability, or does combination of that data with a movie theater dataset showing assistive device usage confirm disability status? As data scales, it compounds and creates a mosaic picture of a person. Within that picture, any individual piece of data may not obviously be discriminatory or expressly identify disability, but collectively the pieces can fit together into a broader pattern identifying personal traits. Two further examples help to illustrate why this mosaic combination in Layer Three is especially concerning for disability status: home Wi-Fi association and arena ticket purchasing.

First, home Wi-Fi networks can create associational data for consumers. A home's Wi-Fi IP address is a weak data identifier because many people can be on a single network, so data attached to an IP address is not guaranteed to come from one unique person. However, advertisers can leverage IP addresses to place advertisements to other people in the home who are using the same Wi-Fi.[42] Most people think their Alexa is listening to their conversations; and while Amazon has confirmed they sell specific voice interactions with Alexa to marketers,[43] it is far less clear that Alexa is listening in the background. Rather, a simple data trick accounts for that creepy feeling when a roommate or sibling gets an advertisement for the product somebody was just discussing. The conversation was audible, but advertisers were really listening to the data that was coming from the two devices connected to the same Wi-Fi network.[44]

42. WCCO NEWS, *Targeted Ads: What They Are & How They End up on Your Devices*, CBS MINN. (Dec. 19, 2018), https://www.cbsnews.com/minnesota/news/targeted-ads-internet-mobile-devices-websites-shopping/.

43. Jennifer Pattison Tuohy, *Researchers Find Amazon Uses Alexa Voice Data to Target You with Ads*, VERGE (Apr. 28, 2022), https://www.theverge.com/2022/4/28/23047026/amazon-alexa-voice-data-targeted-ads-research-report.

44. One roommate browses Amazon for a widget. The other browses Instagram. Roommate A talks about the widget as they browse. In the background, that browsing data is glued into a massive Layer Two dataset associated with that household's Wi-Fi IP address. Then, that data is transferred to Layer Three, where advertisers learn that the other identifiable devices on the network have not browsed for the widget. The insight is remarkably simple: show the other devices (the other roommate) widgets on Instagram in hopes that they see the products and mention them back to Roommate A as an additional product consideration.

By household association alone, one person's profile can be used as leverage against the others. Any given data feature can be traded around in the background to facilitate these insights. This inferred relational association for product placements is not unlike disability association. For example, the spouse of a person with a disability could routinely purchase companion seats or request accessible rental cars even when they themselves do not have a disability. In this way, the very same data features that identify a person with a disability can also identify a person's family or close contacts—expanding risks of discrimination by association.

Aside from associational risks, Layer Three also presents another prominent risk: inferring physical barriers into a dataset. This is best illustrated by way of hypothetical example:

Jamie is a Golden State Warriors superfan who uses a wheelchair. Suppose the Warriors have ten years of ticketing data that show Jamie attended games in Section 210 for a few years before making a switch to Section 114.[45] Their ticket forecasting algorithms flag longtime fans who moved from the 200 level to the 100 level. Through years of testing, the Warriors know that once a fan that moved from the 200 to 100 level buys a courtside seat, the fan is likely to engage with a marketing campaign for premiere season tickets after their courtside experience. Thus, if the Warriors target that fan with promotional materials, they could convert them from a 100 level purchaser to a higher value season ticket holder. Naturally, the box office sends their premiere marketing material to fans the day after a courtside experience. Will the box office ever offer Jamie premiere season tickets?

Under the above marketing scheme, Jamie may never receive an email for premiere season tickets. The Chase Center does not list any accessible courtside seats on their arena map, which likely means that Jamie may never buy a courtside seat and consequently never trigger the marketing algorithm for premiere season tickets.[46] Alternatively, the algorithm could supplement Jamie's data by flagging the "type of seat purchased" to include the inherently disability-coded feature (Sections 210 and 114 contain accessible seats). But what if the Warriors embargo the use of "type of seat" in their algorithms so as not to discriminate against disabled people? The good-faith effort to prevent disability-coded data from playing a role in marketing algorithms unfortunately still allows disability status to leak into the picture because of the underlying

45. *See* Map of the Golden State Warriors Arena, CHASE CTR., https://www.chasecenter.com/maps (last visited Nov. 27, 2023) (published online as required by guidelines established under the ADA).

46. *Id.*

real-world stadium features preventing Jamie from ever buying ADA-accessible tickets below the 100 level.

Undoubtedly, many non-disability data features present risks for covert and overt discrimination. It's entirely possible in an anachronistic, data-less world, an individual sales representative reproduces the same discriminatory result as an algorithm, operating from their own internalized biases and stigmas. But the world is full of data now, and bias can be replicated across datasets rapidly at scale. Disability status is situated uniquely among protected characteristics in the consumer data ecosystem, due to its sheer volume of occurrence, lack of status permanence, opportunity for misidentification (like with people living in the same home), and replication of real-world barriers (like in Jamie's case). For these reasons, disability data warrants special legal protections.

III. HOW THE LAW INTERSECTS WITH THE LAYERS OF THE CONSUMER DATA ECOSYSTEM

Within the layers of the data ecosystem, there are various points at which legal intervention could address disability data discrimination. Importantly, the law can intervene within a layer (i.e., prohibiting the collection or inference of "disability"), or it can intervene between the layers (i.e., prohibiting sales of disability data or inferences between layers). While the following Section will examine some existing laws and cases, grouped by layer, it is important to consider that laws need not be restricted to one layer. In fact, robust laws that address multiple layers of the data process could incentivize better non-discriminatory data practices by creating liability for multiple actors.

A. INSULATING LAYER ONE

The most rigid legal solution in the data ecosystem is to stop disability data at the point of intake. For example, a total embargo on *retaining* disability markers would mean a restaurant reservation system could *collect* a disability disclosure for an individual reservation, but then must delete it once the reserved meal occurs. But an embargo of this type raises autonomy concerns and equal enjoyment concerns regarding how a person with a disability is treated as a repeat customer.[47] Moreover, some disclosures will persist for a

47. For example, a preference for outdoor seating could be kept for a person without a disability (assuming there are no accessible outdoor seats), streamlining their future booking experience when the preference is automatically populated during future bookings. But if a company cannot retain disability disclosures for accessible tables, the person with a disability will always have to repeatedly disclose their disability for each future booking. This solution

much longer time in a dataset until "expiration."[48] A more practical Layer One intervention would still allow data to be collected and retained for benign purposes. Those purposes may be statutorily permissible, and often, they are necessary to ensure a business can accommodate a person with a disability. For example, the DOJ's promulgated regulations for ticketing under the ADA allow venues to solicit a general affirmation of disability to prevent purchasing fraud.[49] This, however, creates a chicken and egg problem, which appears in many places of disability law: to accommodate a person, a public accommodation or employer generally must engage in an interactive process that implicitly involves being notified of a disability.[50] Without some collection of disability data, the process of accommodation breaks down.

Although stopping collection may be damaging to the ability to provide accommodations, stopping the transfer of collected disability data presents almost no risks to accommodation. Put plainly: data never has to leave Layer One for the goals of the ADA to be fulfilled. Disability characteristics need never be sold for commercial purposes. Whether that characteristic is an accessible hotel room reservation or a customer note indicating a sign language interpreter should be present at check-in, the data can be used for the purposes of business operations without it ever needing to leave that business.

The Driver's Privacy Protection Act of 1994 (DPPA) is a successful example of protecting data at Layer One.[51] The Act prohibits states from disclosing "personal information" obtained through state department of motor vehicle (DMV) records.[52] The statute specifically defines personal information so as to include "medical or disability information," and the enumerated permissible uses of such data include only the "normal course of business."[53] Under the DPPA, personal information can only be shared for "marketing" or "bulk distribution" (i.e., to Layer Two harvesters) when the

not only creates a different booking process, but also intrudes on a person with a disability's autonomy if they wish to disclose their disability and allow its re-use.

48. Consider the dog license registration example outlined above. If the park district required a disability disclosure alongside a dog license application, it would retain that disclosure for the life of the license. If the license were valid for five years, should disability status be used to embargo all of a citizen's data, or only the data particular to the dog license? Could the park district disclose the information after five years?

49. 28 C.F.R. § 36.302(f)(8) (allowing disability disclosure solicitation to combat ticketing fraud).

50. 29 C.F.R. § 1630.2(o)(3) (describing the "interactive process" for reasonable accommodations under labor guidelines).

51. 18 U.S.C. § § 2721–25.

52. *Id.* § 2721(a)(1).

53. *Id.* §§ 2721(b)(3), 2725(3).

State obtains "express consent" from the identified person.[54] Moreover, the DPPA also applies to private recipients, enforcing resale and disclosure requirements on anybody who receives data from a state DMV, and includes a five year record-keeping requirement to identify any redisclosures.[55] In effect, the DPPA enforces a Layer One restriction on data disclosures and includes accountability measures for any Layer Two recipient to ensure they too do not misuse the data.

In 2000, the DPPA passed constitutional muster in *Reno v. Condon*,[56] when the Supreme Court found the Act to be a valid exercise of the congressional commerce power. There, South Carolina passed a state law that enabled DMV data sales under the more lenient condition that "information will not be used for *telephone solicitation*."[57] The Court recognized that Congress had focused on regulating disclosures in the DPPA because of the states "significant revenue[]" interests in selling their DMV data.[58] Importantly, Chief Justice Rehnquist specifically noted "[t]he DPPA's provisions do not apply solely to States," rather, they "regulate[] the universe of entities that participate as suppliers to the market for [DMV] information."[59] Ultimately, the Court upheld the DPPA and found Congress could regulate the DMV data as an "article in interstate commerce."[60]

Recently, the DPPA has been used to challenge a disability data transfer that occurred between Layer One and Layer Two. The ongoing case, *Gershzon v. Meta Platforms, Inc.*, involves Meta's pernicious use of its proprietary software tracking code, the Meta Pixel ("Pixel").[61] Meta uses Pixel to "acquire[] personal information . . . for its advertising business" by installing tracking code on a given website.[62] A website owner typically will install the Pixel, or similar software, to help "keep track of user activity on their websites," presumably to figure out what features can be improved or changed.[63] Mikhail Gershzon is a California resident who visited the California DMV website, which is alleged to have Pixel installed.[64] The Pixel tracks outbound link requests for a given user, and it was able to discern Gershzon's first name, an email address, and

54. *Id.* § 2721(b)(12).
55. *Id.* § 2721(c).
56. Reno v. Condon, 528 U.S. 141 (2000).
57. *Id.* at 147 (emphasis added).
58. *Id.* at 143–44.
59. *Id.* at 146, 151.
60. *Id.* at 149–50.
61. *See* Gershzon v. Meta Platforms, Inc., No. 23-cv-00083-SI, 2023 WL 5420234 (N.D. Cal. Aug. 22, 2023); META PIXEL, *supra* note 8.
62. *Gershzon*, 2023 WL 5420234, at *1 (quoting Compl. ¶ 16).
63. *Id.*
64. *Id.*

the URLs he visited.[65] Because Gershzon visited California DMV links regarding his specific "application for a disabled parking placard" and the related links to monitor his application status, Pixel was able to capture a strong proxy, if not actual, indicator for his disability status.[66] As it stands, Mikhail Gerhszon survived a motion to dismiss, and sufficiently pled a case alleging that his disability personal information was shared in violation of the DPPA.[67]

Yet, the DPPA may have limits as it applies to tracking software installed on a DMV website information. In the last few months, the Northern District of California declined to extend the logic from *Gershzon* to a similar case.[68] There, the plaintiff Jacqueline Jackson alleged that LinkedIn's version of the Pixel, the "Insight Tag" was installed on the California DMV website and thus obtained her personal information from DMV URLs when she renewed her disability placard.[69] In what can only be described as procedural hair-splitting, the court granted LinkedIn's motion to dismiss because Jackson insufficiently alleged how the "disability information embedded within URLs qualifie[d] as a motor vehicle record."[70] Citing out-of-circuit cases that held that personal information on public DMV websites were not "motor vehicle record[s]" within the meaning of the DPPA, the court concluded Jackson failed to state a claim under the DPPA.[71] The court found it meaningful that Mikhail Gershzon had alleged the URLs he accessed were in his "MyDMV" account, ostensibly a "motor vehicle record," whereas Jackson failed to allege the URLs she visited were part of a "MyDMV" account application.[72] The court's ruling may only seem to erect a trivial barrier to Jackson, who can amend her complaint, but the ruling is a warning sign about the limits of disability data protection. If an individual logged into their MyDMV account and visited the "how to" page numerous times without initiating an actual disabled placard renewal application, their tracked data would never be part of a "motor vehicle record," and thus free for LinkedIn's taking. Of course, the disability status is less concretely identified, but surely given the discussion above about data harvesting and insights, LinkedIn could infer disability status using the data.

State DMVs are not the only websites involved in litigation for their use of Meta's Pixel. Another case, initially dismissed, *Salazar v. National Basketball*

65. *Id.* at *5, *13.
66. *Id.* at *5.
67. *Id.* at *13.
68. *See* Jackson v. LinkedIn Corp., No. 24-cv-00812-PCP, 2024 WL 3823806, at *4–5 (N.D. Cal. Aug. 13, 2024).
69. *Id.* at *1.
70. *See id.* at *4.
71. *See id.* at *4–5.
72. *Id.*

Association, involved Meta's Pixel receiving viewer data from the National Basketball Association's (NBA) website.[73] The plaintiff brought suit under a statute fairly similar to the DPPA, the Video Privacy Protection Act of 1988 (VPPA), 18 U.S.C. § 2710.[74] Michael Salazar alleged that Pixel monitored his viewing habits and also received data from the NBA about his "[website] usage, . . . items [he] purchase[d], and the . . . frequency . . . of [his] interactions" with the website along with his name, email, and more.[75] The district court ruled Salazar failed to state a claim under the VPPA because of his lack of "subscriber" status and because of his purported consent to the NBA's privacy policy. However, the Second Circuit vacated and remanded the case, holding that Salazar qualified as a "subscriber" under the VPPA even though he did not pay for his subscription.[76] Regardless, the case provides valuable insight into the type of consumer data being collected by technology companies like Meta.[77]

In re Nickelodeon Consumer Privacy Litigation was a similar case involving Google, Meta's marketing competitor, where the court also considered the VPPA and the website users' viewing habits.[78] It was another failed attempt to use the VPPA to hold software tracking code liable for personal information collection.[79] The plaintiffs alleged that Viacom (Nickelodeon and its website's owner) disclosed data to Google on a "child's gender, . . . birthdate, . . . and web communications, including . . . detailed URL requests and video materials requested and obtained from Viacom's children's websites."[80] In assessing the VPPA, the Third Circuit agreed with the Seventh Circuit's reasoning in another VPPA case that the statute was "not well drafted."[81] Ultimately, Google's motion to dismiss was granted because the court, like its sister circuit, reasoned

73. *See* Salazar v. Nat'l Basketball Ass'n, 685 F. Supp. 3d 232 (S.D.N.Y. 2023), *vacated and remanded to* 118 F.4th 533 (2d Cir. 2024).

74. *Id.* at 235. The VPPA aims to prohibit the wrongful disclosures of video tape rental and sale records from a video service provider. *See* 18 U.S.C. § 2710. Like the DPPA, the VPPA includes provisions on disclosures with consent or disclosures in the service provider's "ordinary course of business." *Id.* §§ 2710(b)(2)(B), (E); *cf. id.* §§ 2721(b)(3)(A)–(B) (discussing the limited circumstances where state DMVs are permitted to disclose personal information for "use in the normal course of business by a legitimate business or its agents").

75. *Salazar*, 685 F. Supp 3d at 236–37.

76. *See Salazar*, 118 F.4th at 552–53.

77. *See Salazar*, 685 F. Supp 3d at 244–45 (summarizing a laundry list of personal data that the NBA shares as part of its user agreements).

78. 827 F.3d 262 (3d Cir. 2016).

79. *Id.*

80. *Id.* at 269.

81. *Id.* at 278 (citing Sterk v. Redbox Automated Retail, LLC, 672 F.3d 535, 538 (7th Cir. 2012)).

and interpreted the VPPA to limit disclosure liability to video tape providers only, not the recipients of such data.[82]

Neither case litigated under the VPPA involved any claim of disability information disclosure. However, both cases, brought as putative class actions, could easily have represented web users with disabilities. Viewership data might include requests for closed captions and both websites are likely to have URLs that explain accessibility features or contain disability-coded markers akin to the California DMV. Finally, outside the purview of the VPPA, *Salazar* left unclear whether Pixel's data collection on "items purchased" on NBA.com could include ticket purchasing data; data that, per the discussion above, could implicitly tag a customer's suite of web habits with a disability marker if the tickets purchased were ADA compliant seats.

While other laws can be conceptualized as Layer One interventions too,[83] the DPPA and VPPA provide effective, simple examples of what can be done to protect data at its initial collection point. Coincidentally, Congress passed both acts in response to major pop culture events: the DPPA to the stalking and homicide of actress Rebecca Schaeffer, and the VPPA to the publication of Supreme Court nominee Robert Bork's household video rental history.[84] Both Acts were a response to improper uses of commercialized data; however, the two laws are different in one key respect: extended accountability. Where the DPPA extends liability to *any* recipient of protected data through redisclosure standards, the VPPA limits liability only to the original disclosing video service provider.[85] Put another way, the DPPA transcends Layer One and advances into Layer Two, limiting how a harvester can use the data, whereas the VPPA reaches its bounds at Layer One, and imposes no requirements on how someone who receives data can subsequently use it.

B. DISCLOSURES AND PRIVACY IN LAYER TWO

The lack of uniform regulations in the data privacy space is well-documented, and few states have filled the federal void.[86] Bloomberg Law

82. *In re Nickelodeon*, 827 F.3d at 281.

83. *See, e.g.*, Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191 (1996) (a privacy law with strict limitations on how health information can be shared outside of certain medical relationships).

84. *The Drivers Privacy Protection Act (DPPA) and the Privacy of Your State Motor Vehicle Record*, ELEC. PRIV. INFO. CTR., https://epic.org/dppa/ (last visited Nov. 27, 2023); *In re Nickelodeon*, 827 F.3d at 278; *see also Salazar*, 118 F.4th at 544–45 ("The Bork Tapes [were] a catalyst for the VPPA").

85. *In re Nickelodeon*, 827 F.3d at 281.

86. Thorin Klosowski, *The State of Consumer Data Privacy Laws in the US (and Why It Matters)*, N.Y. TIMES (Sept. 6, 2021), https://www.nytimes.com/wirecutter/blog/state-of-privacy-laws-in-us/.

recently reported that only twenty states have "comprehensive data privacy laws."[87] Only three states have biometric information privacy laws, while five others are on track to capture biometric privacy in their more comprehensive frameworks.[88] Understanding the federal and state ecosystem as a fractured patchwork helps explain why Layer Two data protection is particularly challenging to implement. Data harvesters collect their information from disparate and fractured sources. When those disparate sources are all subject to different state and international disclosure laws, information easily slips between the cracks.[89] Thus, individual states have begun responding to these concerns in a variety of ways, as has the federal government.

Illinois was the first to pass a biometric privacy law, the Biometric Information Privacy Act (BIPA).[90] Aimed at protecting biometric information like scans of retinas or facial geometry, BIPA has already proven successful at penalizing a notorious Layer Two data harvester. Meta (at the time doing business as Facebook) settled for $650 million in the *In re Facebook Biometric Information Privacy Litigation*, after it became apparent they would be liable to a full class of users whose facial data was collected from photos.[91] Although disability was not spotlighted in the suit, it is entirely possible a subclass of users with disabilities were disproportionately harmed by facial recognition algorithms that inferred disability from physical appearance.[92]

Notably, BIPA was legislatively centered on an important data insight: permanence. The Illinois state legislature included in their codified legislative findings the rationale for protecting biometric information: unlike a social security number that can be replaced when breached, biometric data, cannot

87. *Which States Have Consumer Data Privacy Laws?*, BLOOMBERG L., https://pro.bloomberglaw.com/brief/state-privacy-legislation-tracker/ (last updated Sept. 10, 2024).

88. *Is Biometric Information Protected by Privacy Laws?*, BLOOMBERG L., https://pro.bloomberglaw.com/brief/biometric-data-privacy-laws/ (last visited Nov. 27, 2023).

89. *See* Klosowski, *supra* note 86.

90. 740 ILL. COMP. STAT. 14 (2008).

91. 326 F.R.D. 535 (class action derived from Patel v. Facebook, Inc., 290 F. Supp. 3d 948 (2018)); *see Judge Approves $650M Facebook Privacy Lawsuit Settlement*, ASSOCIATED PRESS (Feb. 26, 2021), https://apnews.com/article/technology-business-san-francisco-chicago-lawsuits-af6b42212e43be1b63b5c290eb5bfd85 (describing the privacy lawsuit based on Facebook's failure to obtain consent before "using facial-recognition technology" on user photos).

92. Liability could arguably fall under the "regarded as" disabled prong of the ADA when a facial scan is perceived as someone with a disability; *see also* Ian Moura, *Addressing Disability and Ableist Bias in Autonomous Vehicles*, DISABILITY RTS ED. & DEF. FUND (Nov. 7, 2022), https://dredf.org/wp-content/uploads/2024/08/DREDF-Moura-AV-AI-Brief-Nov-2022-UPDATE.pdf (discussing algorithmic vision model issues in self-driving cars and the inability to recognize persons in wheelchairs).

be rewritten or changed.[93] By focusing on data permanence, Illinois has shown it is possible to effectively regulate data based on the real-world features associated with its content. Data permanence is a feature of biometric data that is inextricably intertwined with the identifying nature of the data itself. For example, a facial scan produces geometric content that identifies a person, and once collected, can always be accurately used to describe, or identify somebody absent extraordinary medical circumstances. Data permanence is sometimes accurate as to disability status, but other times the inverse is true. Because it is possible that an individual's disability is in fact *not* permanent, their disability coded data could be ineffective, and it could fail to accurately describe their present disability status. Yet as discussed above, the ADA protects this impermanent status because the perception and stigma of disability may persist even after somebody no longer satisfies the definition of disability.[94] Both the ADA's "regarded as" prong and BIPA's focus on data immutability contain strong, judicially tested rationales for intervention in Layer Two harvesting. BIPA effectively prohibits Layer Two harvesters from retaining certain data because it never goes stale and poses serious concerns if breached.[95] Put another way, the logic of BIPA is to create a policy barrier in Layer Two that allows permissible collections at Layer One. Similarly, the ADA seems to suggest that harvesters should not retain disability data because of the risk of continued stigma, even when the disability is impermanent, and the data has gone stale.

Other states have focused on Layer Two interventions that highlight transparency and deletion. The goals of these laws are to empower individuals to demand harvesters stop using their data. In an ideal world, enough people would opt-out of data collection such that the harvesters' small sample datasets cannot produce broad discriminatory results. California has been somewhat of a stalwart when it comes to these kinds of consumer data laws. The California Consumer Privacy Act of 2018 (CCPA)[96] is an all-purpose steroid pack for consumer data transparency and deletion. It contains provisions creating a right to know what is collected or sold, a right to opt out of sales, and a right to delete personal information.[97] Of note, the definition of personal information includes any identification at a household level and extends to

93. 740 ILL. COMP. STAT. 14 / 5(c) (2008).

94. *See supra* Section I.C.2.

95. *See* 740 ILL. COMP. STAT. 14 / 15(a) (requiring entities in possession of biometric data to destroy it once its "initial purpose for collecti[on]" is satisfied or after three years of initial collection).

96. CAL. CIV. CODE §§ 1798.100–1798.199.100.

97. *Id.* §§ 1798.105, 1798.110, 1798.115, 1798.120.

physical characteristics and medical information.[98] Thus, disability is likely covered both individually and by household identifications. The CCPA works in tandem with California's requirement that all data brokers register with the state.[99] These rules, recently amended by the Delete Act,[100] now provide the ability for somebody to request deletion of their personal information from data brokers in one place.[101] The amendments closed a loophole in the CCPA that required individuals to request deletion from every business individually (something nearly impossible when the average consumer does not know the name of most data brokers nor what personal data is being traded).[102]

The effectiveness of transparency regulations to date has been questionable. A chief example involves Clearview AI, a facial recognition technology known for its partnership with law enforcement agencies.[103] Clearview has a well-documented history of algorithmic bias and unethical data harvesting conduct that has resulted in foreign penalties.[104] Yet, as *New York*

98. *Id.* § 1798.140(v)(1). Presumably the household level protections apply to the scenarios contemplated *infra* Section I.C.3.

99. *Id.* §§ 1798.99.80–89.

100. S.B. 362, 2023 Leg. (Cal. 2023) (Data broker registration: accessible deletion mechanism).

101. CAL. CIV. CODE § 1798.99.86.

102. *See* Jedidiah Bracy, *California Governor Signs Delete Act into Law*, IAPP: THE PRIVACY ADVISOR (Oct. 11, 2023), https://iapp.org/news/a/california-governor-signs-ca-delete-act-into-law/.

103. CLEARVIEW AI, https://www.clearview.ai/law-enforcement (last visited Nov. 27, 2023).

104. *See, e.g.*, Thaddeus L. Johnson & Natasha N. Johnson, *Police Facial Recognition Technology Can't Tell Black People Apart*, SCI. AM. (May 18, 2023), https://www.scientificamerican.com/article/police-facial-recognition-technology-cant-tell-black-people-apart/ (discussing research that shows facial recognition technology "can worsen racial inequities in policing," in part due to programmer and dataset biases); Katherine Tangalakis-Lippert, *Clearview AI Scraped 30 Billion Images From Facebook and Other Social Media Sites and Gave Them to Cops: It Puts Everyone Into A 'Perpetual Police Line-Up'*, BUS. INSIDER (Apr. 2, 2023), https://www.businessinsider.com/clearview-scraped-30-billion-images-facebook-police-facial-recogntion-database-2023-4 (detailing criticisms to Clearview's acquisition of images and the subsequent use of its technologies by multiple law enforcement agencies); James Vincent, *Clearview AI Ordered to Delete Facial Recognition Data Belonging to UK Residents*, VERGE (May 23, 2022), https://www.theverge.com/2022/5/23/23137603/clearview-ai-ordered-delete-data-uk-residents-ico-fine (identifying the U.K. as the fourth country to order Clearview to purge its residents' data from its system); Natasha Lomas, *Italy Fines Clearview AI €20M AND Orders Data Deleted*, TECHCRUNCH (Mar. 9, 2022), https://techcrunch.com/2022/03/09/clearview-italy-gdpr/ (reporting that Italy's data protection agency announced a €20 million fine for violations of the European Union's General Data Protection Regulation rules).

Times reporter Kashmir Hill has pointed out,[105] less than one thousand people in California have used their CCPA rights to request access and deletion over the last two years—a miniscule number given California's population. Further, Hill is unable to exercise any CCPA protections herself to request deletion of her facial data because she is a resident of New York.[106] The House Energy and Commerce Committee offered one proposal, the American Data Privacy and Protection Act (ADPPA) (H.R. 8152), to fill the federal void. Although the Committee referred the bill out of committee, it has since stalled. Importantly, along with regulations like those listed above, the bill has a section focused on civil rights. Although "disability" is undefined, the Senate's version explicitly lists it as one of the classes of data that cannot be "collect[ed], process[ed], or transfer[red]" in a discriminatory way.[107] However, in the absence of any uniform federal law, consumers will continue to have challenges with the fractured state-by-state approaches to regulating Layer Two data harvesters.

Although BIPA has been around for more than a decade, laws like the CCPA and the Delete Act are still in their infancy. The effectiveness of Layer Two data harvester regulations has not been fully proven. Whether or not a federal law can fill the void, disability status remains a data feature that is floating around largely unchecked in the data harvesting ecosystem.

C. INFERENCE INTERVENTIONS IN LAYER THREE

Finally, Layer Three data and algorithmic inferences that can result in discrimination are the hardest to identify but often create the most impact in the data ecosystem. Many examples of Layer Three discrimination are not facially apparent in the data but can permeate through two different inference models: forecasting and affinity grouping. Forecasting is predictive in nature and takes historical consumer data to predict future demand, sales, or some other business metric. Affinities, on the other hand, are descriptive in nature and are used to estimate general qualities about someone or something.

---

105. Kashmir Hill (@kashhill), Twitter/X, https://twitter.com/kashhill/status/1714748117042577672; *see also Privacy Policy*, CLEARVIEW AI, https://www.clearview.ai/privacy-policy (last visited Nov. 8, 2024).

106. *Fresh Air: Exposing the Secretive Company at the Forefront of Facial Recognition Technology*, NAT'L PUB. RADIO (Sept. 28, 2023), https://www.npr.org/2023/09/28/1202310781/exposing-the-secretive-company-at-the-forefront-of-facial-recognition-technology (last visited Nov. 27, 2023) (describing her inability to "get Clearview AI to delete" photos of her because she lacks privacy protections as a resident of New York).

107. Press Release, U.S. Senate Comm. on Com., Sci., & Transp., House and Senate Leaders Release Bipartisan Discussion Draft of Comprehensive Data Priv. Bill (June 3, 2022), https://www.commerce.senate.gov/2022/6/house-and-senate-leaders-release-bipartisan-discussion-draft-of-comprehensive-data-privacy-bill.

1. Forecasting Perpetuates Prior Discriminatory Practices

The line between data driven discrimination and real-world discrimination is often thin. Jamie, the Warriors superfan in search of season tickets, is not a farfetched example of how physical accommodations can distort the data forecasting ecosystem.

The recently filed *Yaniz v. Chicago White Sox, Ltd.* is illustrative. The case is barely months old but contains the same hallmarks of disability data discrimination that the Warriors hypothetical above illustrates.[108] The plaintiffs allege the Chicago White Sox did not provide an equal opportunity to purchase ADA compliant season tickets through their online portal and instead only offered fans with disabilities less convenient alternatives to purchase season tickets.[109] Further, the complaint alleges the White Sox withheld accessible seats closer to the field from ticket packages until the team was out of playoff contention.[110] Thus, any accessible ticket patterns that appear in the White Sox ticketing datasets are likely to be distorted by the organization's own discriminatory practices. This particular suit was brought under Title III of the ADA and the "equal enjoyment" requirements of Section 12182(a),[111] yet it could easily become a larger disability data discrimination problem when sales are forecast in the future. An algorithm using the White Sox data may incidentally recommend that fewer accessible seats be available to the public early in future seasons to meet an incorrectly inferred demand forecast. The discriminatory pattern could take years to identify or may never become apparent absent repeat violations.

Undoubtedly, generalized trends can hide individualized factors of discrimination in a dataset. One such example is the COMPAS recidivism algorithm that some state courts use to assess risks of individual recidivism when sentencing a defendant.[112] The algorithm gained national attention in 2016 when it was the center of a due process controversy in *State v. Loomis*.[113] There, the Wisconsin Supreme Court ruled that it was permissible for Wisconsin's state trial courts to use the black-box algorithm to help "predict recidivism" when sentencing defendants.[114] To assist with sentencing

108. *See supra* Section I.C.3.

109. Complaint at ¶ 2, Yaniz v. Chi. White Sox, Ltd., No. 1:23-cv-10714, E.C.F. 1 (N.D. Ill. 2023) (Mich. Civ. Rts. Litig. Clearinghouse).

110. *Id.* ¶¶ 3, 71.

111. *Id.* ¶ 12.

112. Ed Yong, *A Popular Algorithm Is No Better at Predicting Crimes Than Random People*, ATLANTIC (Jan. 17, 2018), https://www.theatlantic.com/technology/archive/2018/01/equivant-compas-algorithm/550646/.

113. *See* State v. Loomis, 881 N.W.2d 749 (Wis. 2016).

114. *Id.* at 772.

decisions, a judge used the COMPAS tool, which takes inputs from a defendant's "criminal file," combined with an "interview with the defendant" to predict the defendant's recidivism risk.[115] The court emphasized that the COMPAS tool was only "one of many factors" to be weighed at sentencing and ultimately affirmed that such use comported with due process.[116] However, the court offered a 15-paragraph caution for the use of COMPAS going forward, including the warning that use of COMPAS should be flexible as more data became available.[117] Just five months later, such caution was exercised and reviewed in *State v. Jones*, when COMPAS had a run-in with disability.[118] In an unpublished opinion, a three-judge panel affirmed the trial court's skepticism of a defendant's COMPAS scores.[119] The sentencing judge reasoned that the scores could reflect a "potential[] . . . learning disability" rather than recidivism risks and took that into account as a mitigating factor when reviewing the algorithm's results.[120]

Judicial caution *is* warranted when using tools like COMPAS, as is apparent from the discussion in *Jones*. The court noted that the trial court's brief discussion of the defendant's vocation and education category scores had indicated a potential disability.[121] Yet, absent from such a "brief discussion," was the implicit policy choice the Supreme Court of Wisconsin made in *Loomis*: a single judge must venture into the results and weed out possible disability bias in the algorithmic scores. This is no small task. Countless journalists and academics have dedicated an intense amount of time to doing such work.[122]

The 8-page, 137-item questionnaire used in generating an individual's COMPAS scores can easily introduce disability bias into the proprietary algorithm from multiple categories of questions.[123] The trial judge on review in *Jones* rightly focused on the education and vocation sections of the

115. *Id.* at 754.
116. *Id.* at 769, 772.
117. *Id.* at 767–68.
118. State v. Jones, No. 2015AP2211-CRNM, 2016 WL 8650489 (Wis. Ct. App. 2016).
119. *Id.* at *5.
120. *Id.*
121. *See id.*
122. *See* Julia Angwin, Jeff Larson, Surya Mattu & Lauren Kichner, *Machine Bias*, PROPUBLICA (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing; Jeff Larson, Julia Angwin, Surya Mattu & Lauren Kirchner, *How We Analyzed the COMPAS Recidivism Algorithm*, PROPUBLICA (May 23, 2016), https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm; Andrew L. Park, *Injustice Ex Machina: Predictive Algorithms in Criminal Sentencing*, UCLA L. REV.: LAW MEETS WORLD (Feb. 19, 2019), https://www.uclalawreview.org/injustice-ex-machina-predictive-algorithms-in-criminal-sentencing/.
123. *See* Yong, *supra* note 112.

questionnaire, where questions about behavior in school, job retention, and general conflicts could unfairly punish a person with a learning disability. The court however failed to discuss that the questions about housing stability, number of recent addresses, drug use, and general behavior, could be proxy indicators for disability or cyclical institutionalization. Moreover, even with good-faith efforts, judicial caution alone is unlikely to weed out dangerous bias. After all, COMPAS is "immune from third-party scrutiny" as a private algorithm, so judges cannot see into the opaque model to fully understand how such categories are ultimately reflected in final scores.[124]

But not all algorithms are immune from scrutiny. The Equal Credit Opportunity Act (ECOA), 15 U.S.C. § 1691, is an example of a robust law that encourages scrutinizing algorithm explainability to prevent bias from creeping into credit decision model results. The law requires that a creditor who takes adverse actions against a party (e.g., a creditor who terminates credit or denies a loan) must provide that party a statement of reasons.[125] On its face, the adverse action notice requirement does not appear algorithmically oriented, but the Consumer Financial Protection Bureau has promulgated regulations that put serious guardrails on algorithmically forecasted credit decisions.[126] Creditors must disclose when a factor, such as "length of residence" or "age of automobile," played a role in a credit scoring decision.[127] This disclosure requirement works in tandem with Section 1691(a)(1), which defines the prohibited data attributes that cannot be used in credit decisions. The product of this regulatory scheme keeps credit algorithms in check, presumably because if a party received a statement that race or age played a role in an adverse decision, they would resort to complaint or litigation. Thus, credit companies have strong incentives to prevent an algorithm from ever weighing such factors in the first place so as not to risk algorithmic identification of implicit biases in the data. ECOA therefore provides a clear example of a statutory scheme that incentivizes companies to scrutinize their algorithmic decisions for bias and prevent harms before they occur.

However, the unfortunate catch with ECOA is that disability is absent from the protected categories listed in Section 1691(a)(1). Disability

124. *See* Park, *supra* note 122 (noting the trade secret law protections COMPAS enjoys).

125. 15 U.S.C. § 1691(d)(2).

126. 12 C.F.R. § 1002.9(b)(2) (2023); *see also* FED. TRADE COMM'N, CREDIT DISCRIMINATION, https://consumer.ftc.gov/articles/credit-discrimination (last visited Nov. 28, 2023) (explaining consumer Equal Credit Opportunity rights); *see also* U.S. DEP'T OF JUSTICE, EQUAL CREDIT OPPORTUNITY ACT, https://www.justice.gov/crt/equal-credit-opportunity-act-3 (last visited Nov. 28, 2023) (detailing consumer credit rights under the Equal Credit Opportunity Act alongside other consumer credit rights).

127. 12 C.F.R. § 1002 (Supp. I 2023).

discrimination in credit lending does not enjoy a unified scheme of credit discrimination protections. In cases involving "residential real estate-related transaction[s]," the Fair Housing Act prohibits discrimination on the basis of disability.[128] Other FHA conditions also protect against disability discrimination when advertising and setting the terms of a sale or rental of property.[129] But the remainder of other types of credit transactions do not enjoy this same protection, and they seem to fall only under Title III of the ADA through the public accommodation framework.

At least one state supreme court has made this fractured credit framework even more challenging for people with disabilities. In *Webster Bank v. Oakley*, a bank exercised its right to accelerate the defendant's mortgage loan.[130] The defendant had defaulted on her mortgage payments, after she took unpaid medical leave when her disabilities "rendered her unable to perform her work duties."[131] Because of the default in payments, the bank accelerated the defendant's loan and did not give the plaintiff reasonable accommodations for "foreclosure policies, practices and procedures" that she sought given her disability.[132] After affirming that various sections of the FHA did not require a foreclosing lender to make reasonable accommodations for disability, the court concluded that only Title III of the ADA covered disability in mortgage lending.[133] The court then elaborated that Title III of the ADA only regulates "access" to and not the "content" of goods and services. Effectively, an individual is entitled to access to foreclosure policies to accommodate a disability, but they are not entitled to a change in the policy content.[134] Put another way, the defendant may have been entitled to reasonable accommodations as to *how* the actual foreclosure unfolded, but they were not entitled to accommodations as to *whether* foreclosure was permissible, since the bank used the same general policy for all customers, therefore making "access" to the general "content" of the foreclosure policies equal.[135] Finally, after reviewing precedent from six other circuits with analogous holdings involving the access/content distinction in insurance policies, the court affirmed the denial of any relief under the ADA.[136]

128. 42 U.S.C. §§ 3605(a) (2018).
129. 42 U.S.C. §§ 3604(b)–(f) (2018).
130. Webster Bank v. Oakley, 830 A.2d 139, 144 (Conn. 2003).
131. *Id.* at 143.
132. *Id.* at 143–44, 160.
133. *Id.*
134. *Id.* at 161.
135. *See id.* at 161–62.
136. *Id.* at 161–63.

Although *Webster Bank* did not involve any algorithms or data usage, the practical result was such that credit algorithms could potentially fall outside the purview of the ADA because a person with a disability has equal "access" to the same algorithms as people without disabilities. The irony of course being equal access to an otherwise discriminatory algorithm creates anything but equal enjoyment of nondiscrimination in lending. Consider any number of forecasting algorithms in mortgage lending that a bank may use to determine whether a loan should be extended to somebody, whether a loan should be proactively restructured, or whether to accelerate a default judgment. A forecasting algorithm that considers a history of timeliness of medical payments could disproportionately penalize someone with a disability when they are seeking a mortgage in a number of ways.[137] It seems that legal protections for people with disabilities in the field of mortgage lending are at best weak and at worst a disparate system that allows discrimination in the gaps between statutory coverage.

2. How Affinity Profiling Identifies Disability

The other common model of inference in Layer Three is affinity profiling. Affinity groups are a type of inference that probabilistically classify individuals into certain affinity categories. For example, an individual could be classified as a "dog owner," "video gamer," or "sad teen." One scholar, Sandra Wachter, has gone as far as to claim that these groups should enjoy legal protections as a class in their own right.[138] Wachter observes there is a distinct category of affinities defined by "human-comprehensible characteristics," that do not enjoy legal protections but could still be used in decision-making and exacerbate inequality.[139] On their face, these groups do not invoke any traditionally protected characteristics, however individuals' inclusion in these groups can point to disability. People identified as "dog owners" could be individuals with service animals, for example, and people identified as "sad teens" could be individuals who are seeking therapy or developing mental health disorders for the first time in their lives. Additionally, "sad teens" could be students experiencing bullying about their learning or other disabilities, a reflection of the social reaction to disability rather than a product of a disability

137. Perhaps too many payments alone for medical services could be too high a risk for one lender—somebody who regularly works with physical therapists may be disqualified. Perhaps another lender highly disfavors any sporadic payment histories in somebody's financial profile and prefers clients with consistent, routine payment schedules—somebody with an intermittent chronic condition could be disqualified.

138. Sandra Wachter, *The Theory of Artificial Immutability: Protecting Algorithmic Groups under Anti-Discrimination Law*, 97 TUL. L. REV. 149 (2022).

139. *Id.* at 158 (highlighting groupings like "video gamers" and "income potential").

itself. This exact kind of affinity grouping needs express legal protections, due to the high potential of the groups including people with protected characteristics.

Because there are so many possible groupings, affinity groups are not easy to identify or enumerate statutorily. However, the CCPA has made an exceptional first attempt at identifying such inferred groups. Nested within the definition of "personal information" is the inclusion of "inferences drawn" that appears to cover any inference of "psychological trends, . . . behavior, . . . [or] abilities" along with traditional "consumer[] preferences."[140] The language is sufficiently broad, such that it could include disability and disability-like inferences. Thus, if a plaintiff can identify an affinity group being used by a company in some discriminatory way, liability would attach under the CCPA.

Moreover, affinity groups have already been the subject of recurrent controversy when reviewed under the Fair Housing Act. Three such cases provide examples of how provisions in the FHA interact with affinity groups. The first case, *Fair Housing Council of San Fernando Valley v. Roommates.com LLC* ("*Roommates*"), demonstrates how affinity attributes can be misused to discriminate in roommate listings.[141] The website at issue, appropriately named, was designed to match people looking to fill empty rooms with possible roommates.[142] Users of the site created profiles where they had an ability to disclose various personal attributes and provide additional comments.[143] One such attribute related to family status protected by the FHA: whether somebody would bring children to a roommate match.[144] The website's search engine allowed users to filter and exclude possible matches based on family status and other protected attributes.[145] After an extended discussion as to platform immunity under the Communications Decency Act, the court held that the website could be liable for discriminatory practices under the FHA.[146] Even though users themselves supplied information regarding their membership in affinity groups (like "parent with child"), the court attached discrimination liability to the website. Although the ruling did not explore disability, because disability is expressly protected under the FHA, like family status, the practical result is that disability as an affinity group (either

140. CAL. CIV. CODE § 1798.140(v)(1)(K).

141. Fair Hous. Council of San Fernando Valley v. Roommates.com LLC, 521 F.3d 1157 (9th Cir. 2008).

142. *Id.* at 1161.

143. *Id.*

144. *Id.* at 1161, 1166.

145. *Id.* at 1169.

146. *Id.* at 1176 (discussing 47 U.S.C. § 230 and the website's role in developing content used for discriminatory purposes).

individually or by familial relationship) is likewise protected from online discrimination.

The second case played out much more prominently in the public eye. After two years of investigative reporting from ProPublica, and after Facebook's initial denials, it became clear that Facebook was using "ethnic affinities" on its housing advertising platform.[147] The National Fair Housing Alliance (NFHA) subsequently filed a suit against Facebook in the Southern District of New York in March of 2018 over the company's use of "affinity" groups. In their complaint, NFHA identified some of those groups as expressly disability oriented: "Disabled American Veteran, Interest in Disabled Parking Permit, Interest in Disability.gov."[148] Almost a year after filing, the parties settled, with Facebook agreeing to spin off a dedicated housing portal that severely curtailed the available affinity categories to those permissible under the FHA.[149] This of course leaves one wondering: what happened to those affinity groups when they were not used in the housing context?

Finally, the most recent dispute also involves Facebook. In *Vargas v. Facebook, Inc.*, an ongoing case, plaintiffs have sufficiently pled their allegations that Facebook enables their advertising system to place discriminatory advertisements.[150] Rosemarie Vargas is one of the plaintiffs and is "a disabled female of Hispanic descent and a single parent living in New York City."[151] When she looked for miscellaneous housing ads on Facebook Marketplace, she saw no available housing, but when a "Caucasian friend" conducted a search with the "same search criteria," he received more ads than Vargas did in her preferred areas.[152] Vargas alleges that Facebook knew she was a single parent with a disability and that this information factored into the

147. *See* Chandler N. Spinks, Comment, *Contemporary Housing Discrimination: Facebook, Targeted Advertising, and The Fair Housing Act*, 57 HOUS. L. REV. 925, 935 (2020) for a synopsis of the public dispute; Julia Angwin & Terry Parris Jr., *Facebook Lets Advertisers Exclude Users by Race*, PROPUBLICA (Oct. 28, 2016), https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race; Julia Angwin, Ariana Tobin & Madeleine Varner, *Facebook (Still) Letting Housing Advertisers Exclude Users by Race*, PROPUBLICA (Nov. 21, 2017), https://www.propublica.org/article/facebook-advertising-discrimination-housing-race-sex-national-origin.

148. Complaint at 86, Nat'l Fair Hous. All. v. Facebook, Inc., No. 1:18-cv-02689, 2018 WL 8343918 (S.D.N.Y. 2018).

149. Press Release, Nat'l Fair Hous. All., Fair Hous. Groups Settle Lawsuit with Facebook: Transforms Facebook's Ad Platform Impacting Millions of Users (Mar. 19, 2019), https://nationalfairhousing.org/national-fair-housing-alliance-settles-lawsuit-with-facebook-transforms-facebooks-ad-platform-impacting-millions-of-users/.

150. No. 21-16499, 2023 WL 6784359 (9th Cir. 2023)

151. *Id.* at *1.

152. *Id.*

advertisements she received.[153] The Ninth Circuit briskly dismantled Facebook's immunity defenses and analogized the case to *Roommates* before determining the social media giant could be held liable under the FHA (at least at the pleading stage) for their "patently discriminatory tool."[154]

At the very least, this line of cases demonstrates that existing discrimination law frameworks can stymy discriminatory practices by affinity grouping in the third layer of the consumer data ecosystem. While these cases arose under the FHA, and its specifically enumerated protections, there is no reason to think other regulatory and anti-discrimination frameworks could not operate similarly to prevent disability discrimination.

IV. CLOSING THE GAPS

Given the frameworks that currently apply to various layers of the consumer data ecosystem, there is much that can be done to advocate for and protect the data of persons with disabilities. The same rationales behind the ADA, motivated by architectural and societal barriers to people with disabilities enjoying public life, directly extend into the consumer data ecosystem, and they are applicable to data that describes or alludes to disability status. Legal interventions like those described above can be tailored and applied in ways that secure and protect disability data.

A. REFINING GUIDELINES BETWEEN LAYER ONE AND LAYER TWO

Legal intervention at the point of collection in Layer One is the strongest tool for restricting disability data's movement in the consumer data ecosystem. These types of interventions would prohibit any explicitly marked disability data from transferring to a Layer Two data harvester and making its way to discriminatory inference models in Layer Three.

Practically, the DOJ could first experiment with providing guidelines for Layer One data in the event ticketing market. The DOJ is responsible for guidelines for ticketing as a public accommodation under Title III. Venue ticketing represents a small subset of accommodation types that directly connect to the consumer data ecosystem (through online resellers, for example). The DOJ could experiment in this subset by administering regulations that restrict the sharing of any non-pertinent accessibility information. For example, venues could still sell data that a consumer purchased particular seats, but the venues may not sell information regarding those seats' characteristics, such as if the seats are ADA compliant. This type

153. *Id.*
154. *Id.* at *3.

of regulation could be modeled after the DPPA, which, as discussed earlier, limits any information about "accessible seating" to uses in the "normal course of business."[155] Although the VPPA failed to extend liability to recipients of information, like data harvesters,[156] the DPPA extended liability and record-keeping requirements on redisclosure to recipients. Thus, any potential federal ticketing guidelines should focus on redisclosure liabilities as well, including oversight on data-sharing with subsidiaries or entities with shared ownership interests. For example, how would redisclosure work for a company like Live Nation that owns both venues and a ticketing service, or an entity like the NBA that facilitates basketball game ticketing, but also has an expansive online profile as demonstrated in *Salazaar*?[157] Redisclosure liability can also be easily built into the existing regulations, as the guidelines today recognize and advise on the secondary ticket market.[158]

Aside from the ticketing context, the DOJ should explore how similar regulations could work with other public accommodations and government programs. Logically, one could argue that an "equal enjoyment" of services would mean that a disabled consumer does not have to disclose any data that others do not have to. But sometimes the same piece of information being disclosed is a product of having a disability for some but not for others. For example, park district pet license data about dog ownership could never have explicit disability markers (i.e., there is no data field that says the animal is a service animal).[159] But one owner could appear in the dataset due to a disability, while another appears in that same dataset just by owning a pet. Thus, that data disclosure puts them on unequal grounds.

A practical solution may be to solicit general markers for disability (like in ticketing), so that the data can be withheld when the dataset is sold or shared. For example, one could indicate they are registering with a park district for their service animal, and thus any information regarding that registered animal cannot be sold. An even better solution would let the individual maintain control of their data and consent to its disclosure, much like what is allowed

155. 18 U.S.C. § 2721(b)(3).

156. *See supra* Section II.A, at 121.

157. *See* Press Release, Rob Bonta, Att'y Gen., Att'y Gen. Bonta Files Lawsuit Against Live Nation, Ticketmaster (May 23, 2024), https://oag.ca.gov/news/press-releases/attorney-general-bonta-files-lawsuit-against-live-nation-ticketmaster (alleging the use of "exclusive contracts with promoters and venues" that helps Live Nation maintain a ticketing/venue monopoly); *see supra* note 73.

158. 28 C.F.R. § 36.302(f)(7) (facilitating exchanges for accessible seating for individuals with disabilities who acquired tickets in the secondary ticket market).

159. Assuming arguendo that the park district would not be covered under Title II of the ADA.

under the DPPA. There are dozens of other contexts in the consumer marketplace in which disability data can be generated voluntarily and to which a DPPA-type framework could be applied: renting hotel rooms, renting cars, requesting rideshares, making restaurants reservations, and more. Preventing explicitly coded disability information from being shared within the consumer data ecosystem would also seriously limit the amount of data that can be used to generate affinity profiles by Layer Three actors. Preventing those who collect disability data from sharing it at the earliest stage possible would have beneficial effects further downstream in the data ecosystem when it comes to protecting disability status.

B. PRESUMPTIONS AGAINST OPAQUE INFERENCES IN LAYER THREE

As for potential interventions applied to Layer Three consumer data, increased judicial caution should be the bare minimum. Many algorithms like COMPAS are wholly opaque methods of inference, and courts should be skeptical that these models are innocent of perpetuating issues of benign neglect. Judicial skepticism of algorithmic models should be the default in every context, not just criminal sentencing. Future consumer data protection regulations should consider ways in which such skepticism can be codified. For instance, regulations could require particularized evidentiary standards for algorithms or specialized requirements of expert testimony.

Another way to prevent harmful discrimination could be to institute a presumption of discrimination against an algorithm or dataset when a company cannot or will not transparently elaborate its sourcing. California is a stalwart for transparency disclosures: AB-2013 was just signed into law requiring artificial intelligence training data transparency, and the CCPA effectively mandates transparency in other data sourcing contexts, requiring a business to inform customers what it is collecting from them and why.[160] Coupling a policy of transparency in data with transparency in algorithms would incentivize companies to proactively audit their algorithms and applications for risks of disability bias and discrimination akin to ECOA's adverse action notice model. For ECOA, the prohibition on discrimination, coupled with the transparency requirements in adverse action notices, keep creditors in check so that their models do not inadvertently leverage protected characteristics. Applying the ECOA model to the context of disability data, if a plaintiff can demonstrate discrimination, which can be difficult to do given the facts of cases like those in *Vargas v. Facebook, Inc.*, a burden could be placed on any companies that use affinity profiles to explain exactly how such profiles or algorithms are created and used. This would encourage Layer Three actors

160. CAL. CIV. CODE §§ 3111(a), 1798.100(a)(1).

to proactively remove risky variables and datasets that may lead to discriminatory results.

Finally, despite the difficulties of implementation, Congress should close the gap in the Equal Credit Opportunity Act and expressly enumerate disability status as a class protected from discrimination in credit lending. Today, even if a person with a disability identifies discriminatory loaning algorithms used against them, they cannot do much to stop the discriminatory practices. There is no clear reason why disability status should not have the same protections as other enumerated characteristics in ECOA.

C. ADDITIONAL CONSIDERATIONS

Designing data interventions and proactive protections requires nuance and considerations beyond simple legal questions. In a world where personalization has become the norm, it is possible that well-intentioned legal solutions could still adversely impact people with disabilities. Some data collection could actually help people with disabilities. For example, while detecting screen reader usage on a website implicitly identifies users with disabilities, it could also be used to identify parts of a website that are not seeing engagement due to screen reader inaccessibility. Companies could then use this information to make their websites more accessible.[161] Of course the counterargument would be that people should design with disability in mind at the outset, but when so many digital and physical spaces already exist, many of the entrenched design decisions of the past still need to be made more accessible today.

Data deletion or restrictions could also prevent people with disabilities from receiving beneficial services. For example, predictive notifications to purchase more dog food are just as relevant for service animal owners as they are for nonservice animal owners. The best solution in these cases might be to allow individuals to voluntarily consent to data disclosures so as to preserve their autonomy, even if they remain at risk of discrimination. As discussed, disability data insights can both positively and negatively impact people with disabilities, but the autonomy of people with disabilities must remain at the forefront of any future regulation so that data laws do not become paternalistic exercises of power.

161. *See* Heather Burns, *Detecting Screen Readers in Analytics: Pros and Cons*, POWERMAPPER (Aug. 17, 2016), https://www.powermapper.com/blog/accessibility-analytics/ (surveying the developer debate on whether screen reader usage data should be used in web design).

V. CONCLUSION

People with disabilities continue to face structural barriers in the physical world today. Unfortunately, many of those structural barriers have been carried over into the digital marketplace. The consumer data marketplace is a wildly unregulated space for all Americans, but the unique features and information encoded in the consumer data of people with disabilities means they are susceptible to continued discrimination through their data. Legal intervention can take place at different points in the personal and consumer data ecosystem, and there are existing models showing how Congress can respond to this pervasive, continued problem. Although the days of the median consumer are gone, it seems the problem of benign data neglect has persisted. Further action is desperately needed to combat this neglect.