CYBERLAW CLINIC
HARVARD LAW SCHOOL | BERKMAN KLEIN CENTER FOR INTERNET & SOCIETY

BEFORE THE COMMISSIONERS OF
THE FEDERAL TRADE COMMISSION

COMMENT REGARDING COMMERCIAL SURVEILLANCE
AND DATA SECURITY

COMMENT OF BERKMAN KLEIN CENTER
PROJECTS AND ASSOCIATES

Christopher Bavitz, Kasia Chmielinski, Sandra Cortesi, Sue Hendrickson, Adam Holland, Adam Nagy, Sarah Newman and Shreya Tewari (collectively, "Commenters") submit this comment in response to the Advance Notice of Proposed Rulemaking on Commercial Surveillance and Data Security ("Advance Notice") published by the Federal Trade Commission ("FTC").

1. INTRODUCTION

Commenters applaud the FTC's attention to the risks posed by commercial entities' collection, storage, use, and subsequent flow of data concerning an unprecedented range of human activity, from facial movements and information accessed in web browsing sessions to pinpointed physical location. Swift action to mitigate such risk is key to the FTC's mission of protecting the public from unfair and deceptive trade practices.

The range of potential policy solutions is wide, and this Comment will draw on the expertise of the Commenters to focus particularly on the regulation of commercial surveillance through transparency and disclosure mechanisms, including how such mechanisms impact vulnerable populations such as children and teenagers. We believe disclosure and transparency have a critical role in preventing harm to consumers as a result of commercial surveillance. However, transparency for its own sake, or implemented in isolation, is not necessarily helpful. Commenters have combined decades of experience in operating transparency projects, which

have yielded significant insight, offered here in hopes that it may contribute to the success of the FTC's efforts in this area.

1.1. Statement of interest of the Commenters

Sue Hendrickson is the Executive Director of the Berkman Klein Center for Internet & Society at Harvard University ("Berkman Klein"). Berkman Klein is an interdisciplinary research center that aims to explore and understand cyberspace to tackle the biggest challenges presented by the Internet.

Christopher Bavitz, Adam Holland, and Shreya Tewari lead Berkman Klein's Lumen database ("Lumen").[1] Lumen's goal is to illuminate the overall global climate of online expression, censorship, and rights enforcement by aggregating takedown requests made to online service providers ("OSPs") and facilitating research. Founded in 2002 as the Chilling Effects Clearinghouse, the Lumen database now includes over twenty million notices, and grows by approximately 40,000 every week. Participating OSPs include Google (products including Search, Drive, Blogger, and more), Twitter, YouTube, WordPress, and Medium. Lumen's database and associated research aim to educate the public; to facilitate research on the nature and sources of content takedowns; and to provide as much transparency as possible about the "ecology" of such takedown requests, in terms of who sends them, why, and to what effect.

Kasia Chmielinski and Sarah Newman are co-founders of the Data Nutrition Project ("DNP"),[2] an independent non-profit organization originally incubated at Berkman Klein as part of its Assembly program.[3] DNP was founded in 2018 through the Assembly Fellowship offered by Berkman Klein "to mitigate the harms that arise from the use of inappropriate data in developing artificial intelligence." DNP, now incorporated as an independent 501(c)(3) non-profit organization, will shortly launch the third version of a standardized label, akin to nutrition labels on food, to highlight "key ingredients" in the datasets "such as meta-data and populations, as

---

[1] LUMEN DATABASE, https://www.lumendatabase.org (last visited Oct. 18, 2022).

[2] DATA NUTRITION PROJECT, https://datanutrition.org (last visited Oct. 18, 2022).

[3] ASSEMBLY AT BKC, https://www.berkmankleinassembly.org (last visited Oct. 18, 2022).

well as unique or anomalous features regarding dataset distribution, and both intended and known uses of a particular dataset" that impact how the data can be responsibly used.

Adam Nagy and Christopher Bavitz lead Berkman Klein's Risk Assessment Tool Database project ("RiskDB").[4] RiskDB collects public information about the design and implementation of risk assessment instruments widely used across the U.S. criminal justice system and tracks associated judicial and legislative developments.

Dr. Sandra Cortesi is a psychologist who specializes in the impacts of digital technologies on young people's lives and directs Berkman Klein's Youth and Media project ("YaM").[5] YaM aims to gain detailed insights into youth practices and digital fluencies to help shape evolving regulatory and educational frameworks to advance the public interest. YaM seeks to amplify youth voices and agency while understanding genuine concerns associated with online activity. YaM covers research, advocacy, and development initiatives around youth ages 12-18 and their use of digital technologies.

1.2. Recommendations

Commenters believe that the FTC should use its influence and regulatory powers to push collectors and processors of consumer data towards a system of end-to-end data transparency. Such a system would entail that data is used fairly, traceably, and transparently and require ongoing independent oversight that tracks whether the rules as instantiated are working as intended and continue to be appropriate for the facts at hand. The long-term policy goal should be a data collection ecosystem in which collection occurs on an opt-in basis; consumers may readily identify and retrieve data about themselves possessed by any company; companies can readily delete all data about a consumer upon request, without negative impacts on the goods or services they offer; vulnerable and underrepresented communities are protected from further harm and

---

[4] RISK ASSESSMENT TOOL DATABASE, https://criminaljustice.tooltrack.org (last visited Oct. 18, 2022).

[5] YOUTH AND MEDIA, http://youthandmedia.org (last visited Oct. 18, 2022).

proactively engaged in product design processes; and researchers are able to study, on an ongoing basis, the societal effects of the data collection and processing as well as the effects of the transparency and disclosure policies themselves.

To work toward this goal, Commenters make the following key recommendations:

- **Tailored design of transparency mechanisms**

  The FTC should design and implement transparency mechanisms that are scoped and architected to yield actionable data for specified stakeholders, while also providing some flexibility for unanticipated insights. Transparency mechanisms should be regularly reevaluated to ensure they are achieving policy objectives and keeping pace with innovation.

- **Specific and flexible disclosure requirements**

  The FTC should advance specific disclosure requirements. There is an opportunity, parallel to requirements under the European Union's Digital Services Act, to require disclosure of data collection (including purposes thereof) and requests to delete stored data in a standardized and interoperable system, an independent database, or both. With respect to algorithms, disclosure requirements should include predictive goals, information about the data and techniques used, output explanations, reliability tests, and documentation of efforts to mitigate bias that emerges from model design or data inputs.

  The FTC should implement transparency and disclosure regimes in a way that is sensitive to the complexities of data collection, whether through exempting or reducing requirements for older and active data, or by providing additional time for compliance.

- **Consumer protection for vulnerable populations**

  The FTC should use any avenues available to it to shift the industry to a default where user data is not collected unless users have explicitly opted in. An effective opt-in requires prior notice about 1) what data will be collected, and 2) how that data will be used, as well as constraints on the use of the opt-in option for data collection

beyond what is necessary for the operation of that particular product or service.

The FTC should additionally support the agency of youth and other underrepresented populations by encouraging companies to more intentionally include their perspectives on privacy, personalization, and product design processes.

2. TRANSPARENCY

2.1. Contexts in which transparency is effective

Transparency advances key interests in the regulation of technologies and consumer protection, but the efficacy of any transparency initiative depends on its design and implementation. The Advance Notice asks a critical question concerning the contexts in which transparency or disclosure requirements are effective, as transparency for merely transparency's sake often fails to be impactful. In Commenters' experience, transparency mechanisms are effective when they are tailored to yield actionable data for specific stakeholders and when they afford flexibility in response to the ongoing, and sometimes unanticipated, insights that such data generates. We offer three examples below to highlight the importance of these contextual elements and design features.

2.1.1 Tailored scoping

Tailored scoping of a transparency mechanism enables entities to effectively leverage their often-limited resources to achieve transparency around the most impactful issues. RiskDB, for example, aims to advance accountability in the use of risk assessment tools in the U.S. criminal justice system.[6] The project's impact is predicated on its relatively narrow scope.

Risk assessment instruments are designed to predict an individual's future risk for misconduct using various factors. As many as 60 different risk

---

[6] *About*, RISK ASSESSMENT TOOL DATABASE, https://criminaljustice.tooltrack.org/about/ (last accessed Oct. 20, 2022).

assessment tools are currently in use.[7] At least 23 states use risk assessment tools in criminal proceedings, and some even require their use.[8] The proliferation of these tools may contribute to bias and unfairness in criminal proceedings.[9]

Prior to RiskDB and other similar efforts, the lack of transparency around risk assessment tools, which are often defended under trade secret laws, made attempts to investigate their accuracy and potential bias extremely challenging.[10] Thus, RiskDB is designed to catalogue open-source information on instruments used to assess risks of recidivism in adult populations in pretrial, probation, sentencing, prison, and parole proceedings. RiskDB's extensive and detailed database includes prespecified queries generated by the research team to highlight impactful facets of tool design and implementation, such as user training and certification procedures and information around the functioning of dynamic inputs, if any.[11]

The database's utility to researchers and the public is predicated on its focus on a set of risk assessment tools and its exclusion of other types of algorithmic tools used in criminal proceedings, such as tools that assess mental health issues or substance use disorders. Through its tailored focus, RiskDB is able to provide high-impact transparency to researchers and the

---

[7] Sarah Picard-Fritsche, Michael Rempel, Jennifer A. Tallon, Julian Adler, & Natalie Reyes, Demystifying Risk Assessment, CTR. FOR CT. INNOVATION, 2017, at 1, https://www.courtinnovation.org/sites/default/files/documents/Monograph_March2017_Demystifying%20Risk%20Assessment_1.pdf/.

[8] RISK ASSESSMENT TOOL DATABASE, https://criminaljustice.tooltrack.org/ (last visited Oct. 20, 2022).

[9] Jennifer L. Skeem & Christopher Lowenkamp, *Risk, race, and recidivism: Predictive bias and disparate impact*, 54 CRIMINOLOGY 680, 685 (2016).

[10] Alex Chohlas-Wood, *Understanding risk assessment instruments in criminal justice, in* BROOKINGS INST. AI & BIAS (2020) https://www.brookings.edu/research/understanding-risk-assessment-instruments-in-criminal-justice.

[11] *Comparative Search*, RISK ASSESSMENT TOOL DATABASE, https://criminaljustice.tooltrack.org/comparative_search (last visited Oct. 20, 2022).

public and in turn advance accountability in this field.[12] In its use of disclosure requirements to address commercial surveillance, the FTC should carefully consider what data to exclude, as well as what data to include, to ensure that the transparency mechanism is useful and impactful for stakeholders.

2.1.2 Actionable data

Any transparency mechanism should be purposefully designed to yield actionable data. For example, Lumen was founded to illuminate the climate of online expression, censorship, and rights enforcement. As such, the Lumen database is designed to enable users to perform large-scale queries of metadata to identify trends about content takedowns across time, geography, and different OSPs. At the same time, Lumen also allows users to see details on individual takedown requests, up to and including the full text of many entries, depending on their level of access (e.g., whether they are researchers or the public). The combination of broad metadata queries and granular detail yields actionable insights on key policy topics:

- Tracking organized efforts to suppress information. Lumen allows qualified users to query the domains of content alleged to be infringed. This allows for the discovery of organized efforts to suppress information through the abuse of DMCA, such as using the "back-dated article" technique.[13] In this technique, the complainant creates a copy of the original article and back-dates it, creating a fake "original" article that appears to have been published prior to the true original. Then, the complainant sends a DMCA notice to the relevant OSP, alleging infringement by the true original and requesting its removal. Once the true original is taken down, the copier removes the fake original. As demonstrated by Tewari's

[12] *See e.g.,* Christopher Bavitz et al., *Assessing the Assessments: Lessons from Early State Experiences In the Procurement and Implementation of Risk Assessment Tools*, BERKMAN KLEIN CTR. FOR INTERNET & SOC'Y, (Dec. 2018).

[13] Shreya Tawari, *Over thirty thousand DMCA notices reveal an organized attempt to abuse copyright law*, LUMEN PROJECT (Apr. 22, 2022), https://www.lumendatabase.org/blog_entries/over-thirty-thousand-dmca-notices-reveal-an-organized-attempt-to-abuse-copyright-law.

research, between June 2019 and January 2022, 33,988 notices used the back-dated article technique to request takedowns; all of them used the domain "today-news.press" for the fake original's URL. Most of the targeted domains were Lithuanian, Ukrainian and Russian online news. Such insights from the metadata are critical for understanding the suppression of legitimate speech.[14]

- Tracking growth in the volume and content of takedown notices over time. Lumen allows qualified users to query the number of notices uploaded during a particular day, month, or year. Lumen has seen an exponential increase in the number of notices it receives since its founding: it took over ten years for Lumen to receive its one-millionth notice, a little over a year to receive its two-millionth, and most recently, the fourth million took only eight months. In parallel, research into the content of individual notices reveals that early on each DMCA notice contained on average only one or two URLs, but in recent years, notices may contain up to several thousand URLs. While Lumen does not encapsulate the entire landscape of takedown notices (since participation in Lumen is voluntary on the part of OSPs), these two findings in combination document significant changes in the volume of takedown actions and provide important insights into the evolution of the environment for online speech.

These examples of research outputs from the Lumen database illustrate the necessity of careful attention to the architecture of any transparency mechanism that the FTC may consider.

2.1.3 Accessible outputs

To be effective, the outputs of the transparency mechanism need to be not only actionable, but also accessible. DNP, which works to advance the accuracy and fairness of algorithms by producing understandable

[14] *Id.*

"nutrition" labels to accompany training datasets, is an excellent example.[15] The DNP label is designed to provide "at-a-glance" information about the attributes of the dataset and its common use cases. For instance, each label provides standard metadata about the dataset, but also includes information about who funded the dataset, how it was/is intended to be used, known uses, how it should not be used, and an entire section about possible "inference risks" of the dataset across four categories.

Raw data is insufficient for effective transparency in contexts where stakeholders lack the tools or incentives to process it themselves. By making many relevant attributes of a given dataset readily accessible, DNP increases awareness on behalf of both companies and the public, which fuels responsible collection and use of datasets as well as accountability when they are misused. The FTC should consider whether, in addition to disclosure requirements, it may also wish to implement formatting requirements that increase accessibility of particularly salient data.

2.1.4 Flexibility for unanticipated insights

Transparency mechanisms should be purposefully designed to allow unanticipated insights. For instance, Lumen has included separate metadata categories for the sender and rightsholder. At the time of Lumen's founding, the sender and the rightsholder were most often the same person, but some DMCA notices were sent by lawyers on behalf of large media companies, the rightsholders.

In the two decades since, however, the disambiguation of these two metadata categories reveals an unanticipated insight: content owners are increasingly relying on third-party rights management vendors to send takedown notices. Out of the top ten senders in the eighteen million notices in Lumen's database, seven are third-party outside vendors. The most represented sender, Audiolock.net, has sent nearly one million notices. This insight is significant because content owners' reliance on outside vendors, like automation of the sending of notices, enables a

---

[15] *The Dataset Nutrition Label*, DATA NUTRITION PROJECT, https://datanutrition.org/labels/ (last visited Oct. 18, 2022).

significant increase in the volume of takedown notices and therefore the volume of content being taken down.

2.2. Incentives for transparency and the extent to which companies can reasonably comply

Disclosure and transparency are costly, complex, and in the absence of regulation or broadly accepted best practices, of uncertain benefit to the discloser. As a result, most companies that collect and use consumer data have failed to voluntarily pursue transparency initiatives. To incentivize transparency, the FTC must address concerns about cost, reputational risk, and operational hurdles.

2.2.1. Cost

Lumen has succeeded in recruiting companies, from tech giants like Google to smaller startups, to volunteer data on content takedown requests by adopting measures that lower the cost of participation, despite the absence of formal legal requirements. The Lumen Commenters observed that companies sometimes decline to participate in voluntary data-sharing due to budget constraints. To mitigate the financial disincentive, Lumen has built a plug-and-play extension tool, which automates data-sharing for participating companies under certain pre-determined criteria and thus significantly reduces costs for contributors.[16]

2.2.2. Reputational risk

Transparency and disclosure initiatives are sometimes perceived as posing a reputational risk. The DNP Commenters noticed that companies were more hesitant to adopt their label when the label included normative indicators (e.g., alerts that were perceived as negative) about datasets. DNP addressed this issue by changing the language of its schema to be more neutral in its provision of information relevant to considerations of data use.

Further, Commenters have also observed that some companies are receptive to framings of transparency initiatives that emphasize reputational benefits. Lumen has found that many of its participating

---

[16] Lumen Database, *Lumen Database API Client*, GitHub (2022), https://github.com/leeper/lumendb#lumen-database-api-client.

companies are motivated by the reputational benefits of data-sharing, as well as by the promise of socially impactful research that can be conducted on the data the companies make available. As scandals involving technology companies over the last several years have made clear, consumers are concerned about misuse of data, and participation in a transparency initiative may offer a form of credentialing that companies could use in their marketing efforts.

2.2.3. Operational complexity

Incentives for transparency in the current data collection environment are so slim that many companies have not even developed robust internal transparency mechanisms. Thus, any effective disclosure mechanism will force many companies to change their data collection and utilization processes as part of compliance.

To reduce the operational complexity of participating in a transparency and disclosure mechanism so that companies can reasonably comply, Commenters recommend thinking about data in two categories: collection status (previously collected versus currently being collected versus future collection) and usage in the system (inactive or "not being used" versus active or "being used").

The categorization generates five broad buckets of data, including: 1) *past and unused* datasets, 2) *active and unchanging* datasets, 3) *active and changing* datasets (being changed, being used), 4) *passive and changing* (being changed, not being used) datasets, and 5) *future* (new) datasets. Past and unused datasets are those that will have been fully collected before transparency rules are promulgated and are no longer in use. For example, past and unused datasets may include data from a product or service that is no longer offered. Active datasets are those that are no longer being added to but are still in use in the system. For example, active datasets may include a foundational look-up table filled with values that will not change over time, though the table is still being used as match criteria. Active and changing datasets describe those that will have begun to be collected before transparency rules are promulgated but that will continue to add new data after transparency rules are promulgated. For example, active and changing datasets may include data from a current product or service. Passive and changing datasets are those that will have begun to be collected before transparency rules are promulgated and will continue to change or have

information added to them after transparency rules, but which are seldom used. For example, passive and changing datasets may include data from products that are no longer in development but still used. Future datasets are those that will only begin collection and use after new transparency rules are promulgated.

Each of these dataset types implicates different privacy, administrability, and risk considerations. Datasets that are inactive pose less of a practical threat than their active counterparts. The barriers to transparency and disclosure are higher for old datasets and existing data in active datasets, particularly to the extent that compliance requires the adoption of a particular metadata schema. Transparency and disclosure regimes should be designed in a way that is sensitive to these complexities, whether through exempting or reducing requirements for older and active data, or by providing additional time for compliance. Distinguishing between existing datasets not in use and those still in use (e.g., active and unchanging datasets, active and changing datasets) can also help prioritize measures, as datasets still in use are more critical to address from a harm-reduction perspective than those that have been retired.

2.3. Disclosure schema based on the nature of services

Recommended disclosure schema should depend on the context and nature of services. For example, disclosure for companies may differ from disclosure for government actors. However, there are numerous types of information which we believe ought to be made available to the public by any actor. Scholars have highlighted, in the context of algorithmic systems, several categories of information to which consumers should be entitled, including: (1) general predictive goals and applications of algorithms, (2) relevant, available, and collectable data, (3) excluded data, (4) specific predictive criteria, (5) analytic and development techniques used, (6) principal policy choices, (7) nontransparent accountability, and (8) output explanations.[17] Based on the experience of Commenters, we propose several additional categories of information to be disclosed, including: (9) inter- and intra-rater reliability tests (10) showing that an algorithm effectively addresses bias against any protected class, and (11) data about overrides

[17] Robert Brauneis & Ellen P. Goodman, *Algorithmic Transparency for the Smart City*, 20 YALE J. L. & TECH. 103, 167–68 (2018).

(i.e., when decisions went against recommendations). Each of these eleven categories is important for algorithmic transparency and should be made available for scrutiny and discussion.

In the criminal justice context, disclosure is important both at the procurement stage (i.e., the stage at which the government is procuring algorithms for use in the criminal justice system) and at the implementation stage. Similarly, each of the above disclosure schemas may be relevant to ensuring consumers are properly informed at the procurement stage, the implementation stage, or both, in the context of commercial surveillance.

3. CONSUMER PROTECTION FOR VULNERABLE POPULATIONS

A number of the commissioners' statements regarding the Commercial Surveillance and Data Security Advance Notice of Proposed Rulemaking highlighted concerns about how commercial practices affect marginalized and vulnerable populations, especially children and teenagers.[18] The Commenters believe that these concerns are essential to the overall policy landscape, not least because transparency and disclosure mechanisms of the type contemplated in the previous section risk reduplicating harm done to vulnerable people. The YaM project's extensive research provides essential insight into the digital lives and perspectives of youth, and Commenters draw on that expertise to recommend that transparency mechanisms be designed to safeguard privacy and that the FTC use its influence to encourage companies that collect, store, and use data to adopt opt-in practices and include marginalized people in their product design processes.

3.1. Balancing transparency and privacy for vulnerable populations

The same transparency mechanisms that can support individual agency, effective remedy, and well-designed policy solutions can be simultaneously harmful to – even weaponized against – the people whose data is present

[18] Statement of Commissioner Rebecca Kelly Slaughter Regarding the Commercial Surveillance and Data Security Advance Notice of Proposed Rulemaking (Aug. 11, 2022), at 8-9; Statement of Commissioner Alvaro M. Bedoya Regarding the Commercial Surveillance and Data Security Advance Notice of Proposed Rulemaking (Aug. 11, 2022), at 3.

among the disclosures. Children on the Internet are especially vulnerable to targeted marketing due to "their low levels of online marketing literacy and specific developmental deficits which make them different from other vulnerable groups."[19] This in turn poses a risk that children will suffer detriments to their "developing self-concepts when using marketer-created content for online social exchanges and identity formation."[20] For other vulnerable populations, like victims of sexual harassment, the consequences can be even more acute. Victims whose non-consensual intimate images ("NCII") have been shared online often seek fast solutions and use any available mechanism that they can. Google has a separate process for NCII takedowns, in part to avoid sharing them with Lumen and immortalizing such sensitive materials in its database. However, victims often also request takedown under the Digital Millennium Copyright Act ("DMCA"). DMCA takedown notices are routinely shared with Lumen in an effort to have NCII removed as quickly as possible. These issues can severely impact the well-being of the affected parties and consequently merit careful consideration of strategies to mitigate these harms.

Thus, we must consider the strong interest in privacy and avoidance of further harm, balanced against the value in disclosure, analysis, and preserving of data. Disclosure requirements should be architected to minimize potential consumer harm. Mechanisms to minimize risk of harm from disclosure could include 1) tiered access, restricting more sensitive materials to accredited users; 2) timelocking or windowing, keeping data in a database only for a specified amount of time, or only making it accessible after a set period of time has elapsed; 3) periodic or ongoing review of the data being retained and relevant polices. These mechanisms are critical in particularly sensitive contexts such as children's data and NCII and help ensure that such data, if collected, is not publicly accessible.

---

[19] Ann-Marie Kennedy, Katharine Jones & Janine Williams, *Children as Vulnerable Consumers in Online Environments*, 53 J. CONSUMER AFFS. 1478, 1496 (2019).

[20] *Id*. at 1491.

3.2. Insufficiency of current protections and utility of opt-in data collection

At present, the sole legal protection specific to young people online is the Children's Online Privacy Protection Act of 1998 and related rulemakings ("COPPA").[21] We believe that COPPA in its current form does not adequately protect the interests of young people: neither those under 13 who it covers, nor teenagers who it exempts.

First, very few websites are actually subject to COPPA. COPPA prohibits operators of websites or online services from collecting personal information about children if either (1) the website is directed towards children or (2) the operator has actual knowledge a child is using the site (unless the operator has obtained verifiable parental consent). COPPA is underinclusive because few or no restrictions are imposed on operators of general audience or mixed audience websites, and further, many companies have chosen to avoid COPPA's regulatory framework through age-based bans.[22] Such bans, combined with minimal mechanisms to verify a user's age, allow companies to bypass COPPA's provisions to protect the privacy of underage users by avoiding actual knowledge. For example, companies like Facebook have chosen to ban users below the age of 13,[23] but research shows that millions of under–13 youth are on Facebook since underage users can simply lie about their birthdays when creating an account.[24]

Moreover, even where COPPA is applicable, the measures provided to mitigate risks to children are practically inadequate. The requirement of verifiable parental consent[25] assumes that parents are always available when

---

[21] 15 U.S.C. §6502.

[22] Danah Boyd, Eszter Hargittai, Jason Schultz & John Palfrey, *Why parents help their children lie to Facebook about age: Unintended consequences of the 'Children's Online Privacy Protection Act'*, FIRST MONDAY (Oct. 31, 2011), https://journals.uic.edu/ojs/index.php/fm/article/view/3850/3075.

[23] *Id.*

[24] *Id.*

[25] 15 U.S.C. §6502.

their children are on the Internet and that children will always be truthful when asked their age on a website. Both assumptions are of questionable validity. Underage children gain access by lying about their age during site registration, thereby allowing websites to avoid the "actual knowledge" required by COPPA to trigger its protections and prohibitions.[26] Even when parents are directly involved, they often help their children deceive websites to gain access.[27]

It is unlikely that changes to the COPPA rules would have a significant effect; developing a truly reliable system of age verification and parental consent and age attestation may be impossible. Therefore, we recommend that the FTC exert influence, through regulation or otherwise, to move online companies to a default setting of *not* collecting data about users unless they have explicitly opted in. Any data collection beyond what is necessary for the operation of the product or provision of the service at hand should be subject to a separate opt-in process, to prevent the policy from operating at a merely formalistic level, as many privacy policies do today. Such an opt-in regime would protect other marginalized and vulnerable populations as well as children and teenagers.

Default settings are often outcome determinative, such that they can "fundamentally influence social concerns" like privacy.[28] Default settings have a powerful effect on consumers, who often do not change default settings for reasons ranging from "cognitive and physical laziness; perceiving default as correct, perceiving endorsement from the provider; using the default as a justification."[29] In practice, default privacy settings, in combination with a requirement to opt out of data sharing, practically

---

[26] Boyd et al., *supra* note 22.

[27] *Id.*

[28] Michael J. Kasdan, *Is Facebook Killing Privacy Softly? The Impact of Facebook's Default Privacy Settings On Online Privacy*, 2 N.Y.U. INTELL. PROP. & ENT. LAW LEDGER 107, 114 (2011).

[29] Markus Tschersich & Reinhardt A. Botha. (2013). *Understanding the Impact of Default Privacy Settings on Self-Disclosure in Social Network Services - Building a Conceptual Model and Measurement Instrument*, 19 AMERICAS CONF. ON INFO. SYS. 5 (2013).

enable OSPs to "dictat[e] what kind of privacy users will or will not have."[30] Thus, the FTC can achieve a significant improvement to the state of consumer privacy online by pushing corporations to shift their default privacy settings to be more favorable to consumers.

3.3. Valuing youth and other underrepresented perspectives

Use of the Internet, and particularly social media platforms, is so essential to young people's social lives that many are very willing to allow data collection to maintain free access. As a result, it is unclear whether, particularly in the short term, an improved transparency and disclosure regime would change the online behavior of young users.[31]

Despite their willingness to disclose information, youth still care about, contemplate, and try to manage their privacy online.[32] Teenagers conceive of privacy differently from adults; they tend to have a much more individual and personal perception of privacy.[33] For example, teenagers often explain privacy in terms of their relationships: they care about whether information is available to their friends or family.[34] Teenagers had generally been less concerned about their data being available to institutional or commercial actors, but they are becoming more aware and skeptical of these practices. It's still not certain how their skepticism will affect actual usage and behavior. Nonetheless, these platforms are so essential to young people's social lives that they are willing to give up a lot to use these platforms, especially when they can use these platforms for free in exchange for allowing data collection.

---

[30] Kasdan, supra note 28, at 115 (internal citations omitted).

[31] Sandra Cortesi, Youth and the Participatory Promise 27 (Jan. 28, 2021) (Ph.D. thesis, University of Basel) (on file with University of Basel).

[32] JOHN PALFREY & URS GASSER, BORN DIGITAL: HOW CHILDREN GROW UP IN A DIGITAL AGE (2016).

[33] Cortesi, *supra* note 32.

[34] *Id.*

To respect youths' agency, we also recommend that the FTC encourage companies to incorporate youth engagement into their product design processes, and integrate their perspectives on privacy, personalization, and self-determination. As Berkman Klein affiliate Afsaneh Rigot has persuasively argued, basing product design on the most impacted populations, especially vulnerable and de-centered populations like youth, enhances the overall technology ecosystem.[35] As Commissioner Bedoya pointed out, social media can cause significant psychological harm (in addition to the many opportunities they offer) to young users who are particularly vulnerable because they are exploring their identities and facing feelings of insecurity. YaM has chosen to address this issue by researching overall well-being of young people, which is broader than and encompasses the mental health of young people. Such inclusion in product design helps young people feel empowered and will help companies better understand the needs of such users, such that the products offered can be more intentionally designed to match the expectations and perceptions of young people.

4. CONCLUSION

Commenters are heartened by the FTC's interest in promulgating rules related to commercial surveillance and data privacy. The rapid development of new technologies will only increase the need for effective regulation on this front. The FTC is right to consider transparency as essential to new trade regulation rules to protect consumers. To advance meaningful transparency, we recommend centering the following considerations: 1) intentional design of transparency mechanisms; 2) specific and flexible disclosure requirements; and 3) consumer protection for vulnerable populations. We intend for these recommendations to serve as "first steps" towards developing a culture of transparency with respect to data collection and use. Ultimately, Commenters believe the goal of these measures should be, over time, to create a traceable data pipeline through which any consumer's data can be easily tracked from collection to deployment to deletion. Commenters recommend that the FTC consider populations who will face a disproportionate portion of the negative effects of commercial

---

[35] Afsaneh Rigot, *Design From the Margins*, HARV. KENNEDY SCH. BELFER CTR. FOR SCI. & INT'L AFFS. (2022), https://www.belfercenter.org/publication/design-margins.

surveillance and implement rules that protect the interests of these groups. With the development of technologies like immersive platforms that can track virtually everything about a user, threats to data privacy become more salient. Therefore, Commenters recommend that the FTC consider the direction in which technology is developing when formulating new regulations to proactively address issues these technologies may pose to consumer privacy.

Respectfully submitted,[36]

Jessica Fjeld
Assistant Director, Cyberlaw Clinic
Lecturer on Law, Harvard Law School
Reginald F. Lewis Law Center
1557 Massachusetts Avenue
Cambridge, MA 02138
Tel: 617-384-9269
Email: jfjeld@law.harvard.edu

On behalf of Christopher Bavitz, Kasia Chmielinski, Sandra Cortesi, Sue Hendrickson, Adam Holland, Adam Nagy, Sarah Newman, and Shreya Tewari

[36] Commenters thank fall 2022 Cyberlaw Clinic student Alice Hu of Harvard Law School for their valuable and significant contributions to this comment.