# DATA PRIVACY AND PLATFORM TRANSPARENCY: KEEPING TRACK OF THE MANY MOVING PIECES

There are few legal concepts that have evolved so fundamentally in such a short period of time as data privacy. The evolution of the rights of individuals to determine for themselves when, how and to what extent their personal information is shared with others is evident both in the meteoric increase in the datasets that are at risk absent protections, as well as the awareness of policymakers, regulators and, indeed, the public of the scope of what is now at risk and why, and the concomitant need for protection.

In short, that evolution has been driven by the explosive growth in the collection of personal data and the profitability of the business models that underpin that collection. The scope of online data at risk has grown exponentially as our online footprint has grown – ranging from what traditionally would have been deemed personal information (think applications for credit or credit cards, for example) to online and offline behavior (consider smartphones, connected cars, wearable fitness trackers, smart speakers and browsing history, among many others). Invariably, technology outpaces regulation – generative AI being another example, but we are fast approaching, if we have not already past, the point where regulation must catch up on a comprehensive, national basis, with technology to protect personal data.

A related theme is platform transparency, a catch-all for limits on targeted online advertising, increased transparency about how the platforms' algorithms operate in practice and modification of the blanket immunity platforms enjoy in respect of content moderation.

I provide below a snapshot of where we stand from regulatory standpoint on data privacy and platform transparency, and note that the most recent legislative salvo to address the latter was introduced this past Thursday.

## Background

As Robin Andruss, in her post, "A Brief History of Data Privacy, and What Lies Ahead," notes, early efforts in the United States to protect data grew out of the recognition that government services such as Social Security, Medicare and Medicaid were aggregating and storing ever-increasing amounts of personal data. In 1967, the Supreme Court (in *Katz v. United States*), in a case involving a wiretap of a public telephone, expanded the coverage of the Fourth Amendment beyond tangible property to that which a person "seeks to preserve as private" (what became known based on a concurring opinion as the "reasonable expectation of privacy" doctrine). While the zones of privacy would expand over time in US jurisprudence, personal data did not figure prominently. A more comprehensive way of thinking about personal data appeared in 1980, when the OECD set out its Privacy Principles. It would be another 20 years before data and security would collide.

The early days of the internet prompted concerns about security, and as passwords became ubiquitous (eventually followed, though not so broadly, by two-factor authentication), society largely conflated security and data privacy. Few gave much thought to who owns all the data that was being collected, and few had any inkling of the scope of data that was potentially harvestable (a concept, by the way, that would not burst into the public imagination until 2018). In short order, the platforms caught on to the idea that free services such as search engines would enable them to leverage an increasing amount of data, first to improve search capabilities and then to turbocharge online advertising. Then along came the social media

sites and data-driven algorithms to predict (and shift) behavior, with ever-increasing levels of consumer tracking and analysis to generate advertising revenue.

The wake-up call came with the revelations in March 2018 that Cambridge Analytical had harvested data, on an unprecedented scale, on up to 87 million unwitting users (according to Facebook), that had been shared on Facebook to build voter profiles and target political advertising in the 2016 election. (The data were collected via a personality test deployed on an app called thisisyourdigitallife, which provided access to millions more of the friends of the test-takers.)

Security concerns obviously never disappeared. In addition to the privacy issues surrounding ad-funded use of data, businesses found themselves at risk of data breaches (hacking) of sensitive personal data collected by them for their standard business operations. Businesses routinely have had to embrace cybersecurity defenses, prompted by a combination of the ease with which hackers are able to access IT systems, regulatory requirements and desires to avoid the costly financial, regulatory and reputational consequences of data breaches. Businesses also have to contend with so-called data residency requirements that define which forms of data need to be processed or stored locally and restrict the transfer of that data outside relevant jurisdictions.

## EU Data Privacy Protection

Those that operate in the regulatory space at the intersection of data privacy and online transparency look to the European Union as the undisputed leader. That leadership flows from two separate legislative initiatives, the General Data Protection Directive ("GDPR"), which regulates the processing of personal data, and the Digital Services Act ("DSA"), which is intended to harmonize across the European Union the rules regulating online intermediary services (*see* my previous briefing note, available here).[1]

The GDPR became effective in May 2018, and the first set of provisions under the DSA to become operative did so on February 17 of this year, with the balance largely applying across the European Union beginning February 17, 2024. The significance of these efforts is due in part to recognition of the need for a harmonized approach across the European Union and the fact that they both have extra-territorial effect.

In short, both regimes subject non-EU businesses to EU regulation, in the case of the GDPR, to the extent those businesses process personal data of subjects located in the European Union

---

[1] To understand the principles underpinning the GDPR, one needs to look back to concerns about state surveillance in Germany so clearly evident during the Nazi era. In response to state surveillance that continued following the partition of Germany in East Germany, the West German state of Hessen adopted data privacy protections (the *Datenschutzgesetz*) on September 30, 1970. A Federal Data Protection Act was passed in February 1977. In 1983, the German Federal Constitutional Court declared the right of "self-determination over personal data" to be a fundamental right.

The first national data protection law may have been the Swedish Data Act (the *Datalagen*) adopted in 1973, which became operative in 1974 following an amendment to the Swedish constitution. The European Commission adopted the Directive on Data Protection, which went into effect in 1998 and conditioned cross-border transfers of data outside the European Union on recipient countries meeting EU adequacy standards for privacy protection.

and, in the case of the DSA, to the extent they offer intermediary services to recipients located or established in the European Union (irrespective of from where the services are provided).[2] In both cases, EU policymakers chose not to wait for global standards to emerge, and instead proceeded based on the perceived urgency of protecting EU citizens. Not surprisingly, GDPR and the DSA are viewed by many as the gold standard, reflecting the status, as Columbia Law School professor Anu Bradford posited, of the European Union as an influential regulatory superpower, wielding its power in a manner she termed the "Brussels Effect."

The United States lacks a federal analogue to the GDPR and the DSA. There is a growing patchwork of state consumer privacy legislation (and a set of somewhat antiquated federal laws that target specific types of data, by functional area), and although comprehensive federal consumer privacy legislation failed in the last Congress, it is likely to be reintroduced imminently. Similarly, federal efforts to regulate social medial platforms failed in the last Congress, but have resurfaced again this year.

## The European Model

The GDPR regime distinguishes among "data controllers," "data processors" and "data subjects." Data controllers control the collection and use of "personal data," while data processors carry out the instructions of data controllers. The individuals whose data are controlled by data controllers are data subjects. Personal data covers any information that can identify an individual, including names, email addresses, ethnicity, banking details, IP addresses, biometric data and web cookies. GDPR sets out a list of data subject rights, including rights to:

- be informed that their data have been collected;
- request access to inspect their personal information (through "data subject access requests");
- request that errors in their personal information be corrected;
- request that their personal information be deleted ("right to be forgotten");
- request that their personal information be transferred to another entity;
- restrict the processing of data (though the data can still be stored);
- not be subject to automated individual decision-making; and
- object to certain processing of data, including profiling.

At the heart of the GDPR is the concept of consent. In order to collect, store, process or sell data, one of six legal bases must apply, the most fundamental of which is that the data subject gave its explicit consent to process the data. The other five bases are: to perform a contract to which the data subject is a party, to comply with a legal obligation, to protect vital interests, to carry out a task of public interest or for a legitimate interest (all as defined).

[2] On April 25, the European Commission designed 17 Very Large Online Platforms (VLOPs) and 2 Very Large Online Search Engines (VLOSEs) for purposes of the DSA.

## The Historical US Approach

The United States has no comprehensive federal legislation governing data privacy and we have no federal legislation governing online transparency. We do have a series of legislative regimes that target specific risks, such as:

- the US Privacy Act of 1974, which addresses data held by government agencies;
- the Health Insurance Portability and Accountability Act (HIPPA), which addresses healthcare and health insurance personal data;
- the Gramm Leach Bliley Act (GLBA), which addresses personal information held by regulated financial institutions; and
- the Children's Online Protection Act (COPPA), which addresses personal information of children 12 and younger.

The US regimes essentially are "harms-based" privacy protections, intended to mitigate, if not prevent, harms from misuse or unauthorized use of data collected in the relevant functional sectors, while allowing the collection by institutions and businesses of personal information without consent. This approach contrasts with the GDPR approach that built upon the European "rights-based" approach described above, dating back to the 1970s. This approach recognizes that individuals own their personal data and have the right to control it, and thus are entitled to hold accountable those that hold and choose to use that data.[3]

## Federal Data Privacy Legislation

In June 2022, a comprehensive federal consumer data privacy bill (the American Data Privacy and Protection Act or "ADPPA") was introduced in Congress, but despite bipartisan support in committee, that effort failed to advance to the House or Senate floors in the last Congress. The ADPPA would have established requirements on how companies handle personal data. It would have, among other things (*see* CRS Overview for more details):

- applied broadly to businesses and organizations operating in the United States, including nonprofits and common carriers, and imposed additional or different requirements on "large data holders" and "service providers";
- required companies to limit the collection, processing and transfer of personal data to that which is reasonably necessary to provide a requested product or service, subject to 17 enumerated exceptions;
- generally prohibited companies from transferring consumers' personal data without their affirmative express consent;
- established consumer data protections, including the right to access, correct and delete personal data, and rights to opt out of advertising;
- provided additional protections with respect to personal data of children under 17;
- prohibited companies from using personal data to discriminate based on specified protected characteristics and would have required "large data holders" to conduct algorithmic impact assessments;

[3] The literature also characterizes some approaches as "risk-based" (a variation on harms) and "accountability-based."

- provided for a delayed private right of action; and
- created a Bureau of Privacy at the Federal Trade Commission to enforce its provisions.

The perception that Congress would, as a practical matter, be unable to enact federal legislation gave the states the opening to enact state privacy legislation, and since then states generally have been embracing a GDPR "rights-based" regime. State-level efforts to enact comprehensive privacy legislation is at an all-time high, according to the IAPP (International Association of Privacy Professionals).

## State-Level Data Privacy Protection

As of today, the following have been enacted, with California having first enacted legislation (by ballot initiative) in 2018 and amended it in 2020, followed by Virginia and Colorado in 2021, Utah and Connecticut in 2022, and Indiana, Iowa, Montana and Tennessee this year:

- The California Privacy Rights Act became effective January 1, 2023, amending the California Consumer Protection Act.
- The Colorado Privacy Act becomes effective July 1, 2023.
- The Connecticut Personal Data Privacy and Online Monitoring Act becomes effective July 1, 2023
- The Indiana Consumer Protection Act becomes effective January 1, 2026.
- The Iowa Consumer Data Protection Act becomes effective January 1, 2025.
- The Montana Consumer Data Protection Act becomes effective October 1, 2024.
- The Tennessee Information Protection Act becomes effective July 1, 2024.
- The Utah Consumer Privacy Act becomes effective on Dec. 31, 2023.
- The Virginia Consumer Data Privacy Act became effective Jan. 1, 2023.

An additional 23 states are considering or have passed but not enacted privacy legislation. Yesterday, the Texas legislature passed HB 4 (the Texas Data Privacy and Security Act); it is awaiting signature by the Governor. A tremendously useful resource that tracks state privacy legislation is available from IAPP on its website, including its Legislation Tracker.

The Innovation, Data, and Commerce Subcommittee of the House Committee on Energy and Commerce held hearings in March on reviving the ADPPA, and a draft is expected imminently. One of the more controversial aspects of the initial versions of the ADPPA was the pre-emption of state law, which some argue would have had the effect of weakening protections afforded, for example, in California. (*See, e.g.,* letter from the California Governor and Attorney General, and the California Privacy Protection Agency.) That view though is not universally held. Another open issue was enforcement.

## Platform Transparency -- Section 230

There have been a plethora of legislative efforts to regulate the social media platforms. The focus of these efforts has been to address Section 230 of the Communications Decency Act of 1996, clause (c)(1) of which provides that "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider." In effect, while content creators may be liable for what they

post online, the platforms, when serving as hosts of that content, are not.[4] Section 230's 26 words[5] are viewed as the bedrock of the online industry in the United States, having enabled the unfettered monetization of personal information of online users through the deployment of opaque algorithms. Incidentally, Section 230 has no connection to online decency; in fact, the legislation initially was known as the "Internet Freedom and Family Empowerment Act."[6]

These days, Section 230 allows the platforms to moderate content based on their own standards, as they deem appropriate.[7] Needless to say, the evolution of technology has had a fundamental impact online speech since the heyday of online bulletin boards, circa 1996.

In recent years, efforts to regulate social media platforms have fallen victim to the culture wars, with Democrats broadly arguing that Section 230 allows platforms to host harmful hate speech, disinformation and other malign content with impunity, warranting modification of the blanket immunity to prompt greater responsibility to remove that content. Republicans broadly argue that platform content moderation policies have been misused to suppress free speech, amounting to censorship with a bias against conservatives. Essentially, Democrats have introduced legislation to reduce the Section 230 protections in specified instances, while Republicans have introduced legislation to force platform moderation policies to be "neutral."[8]

---

4 In 2018, Congress carved out an exception from Section 230's immunity shield for civil and criminal charges of sex trafficking. The legislation is the Allow States and Victims to Fight Online Sex Trafficking Act.

5 There is a second, and less publicized, prong to Section 230, which provides the basis for content moderation and website blockers. *See* note 7 below.

6 A concise history of Section 230 is available on Lawfare, in an article by Jeff Kosseff, "What's in a Name? Quite a Bit, if You're Talking about Section 230." A more fulsome guide by the same author appears in the Berkeley Technology Law Journal (Vol 37), "A User's Guide to Section 230, and a Legislator's Guide to Amending it (or Not)".

7 Some have argued that Section 230 in fact was intended to encourage the platforms to moderate content on their sites – in effect, a legislative response to a lawsuit, *Stratton Oakmont v. Prodigy Services Company*, which found that Prodigy Services could be held liable in a defamation action for speech on its site, because is attempted to moderate content. Around the same time, in *Cubby, Inc v. CompuServe*, the court granted summary judgment in a defamation action in favor of CompuServe, on the ground, in effect, that it took a hands-off approach to content on its site (akin to a bookstore, subject to protections afforded distributors). In effect, platforms were better off with no content moderation to speak of.

The basis for content moderation is the second clause of Section 230(c), which provides: "No provider or user of an interactive computer service shall be held liable on account of (A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or (B) any action taken to enable or make available to information content providers or others the technical means to restrict access to material described in paragraph (1)."

8 To keep track of the myriad bills, Slate, New America and Arizona State University, together with the Tech, Law, & Security Program at the Washington College of Law at American University and the Center on Technology Policy at the University of North Carolina at Chapel Hill launched the Section 230 Reform Hub. It is a valuable tool to track the myriad congressional efforts in this space, which incidentally fall into four categories, namely bills that:

In 2021, the Platform Accountability and Consumer Transparency Act ("PACT Act") was introduced on a bipartisan basis in the Senate; that effort failed. This was a revision of legislation introduced in the 2019-2020 congressional session, that also failed to advance. The Internet Platform Accountability and Consumer Transparency Act – S. 483 ("Internet PACT Act") was reintroduced yet again on a bipartisan basis in February. It would amend the Communications Act of 1934 by requiring social media companies to establish clear content moderation policies and would hold these companies accountable for content that violates their own policies or otherwise is illegal.

According to its sponsors, the Internet PACT Act would create more transparency by:

- requiring online platforms to explain their content moderation practices in an acceptable use policy that is easily accessible to consumers;
- implementing bi-annual reporting requirements for online platforms that include disaggregated statistics on content that has been removed, demonetized or deprioritized; and
- promoting open collaboration and sharing of industry best practices and guidelines through a National Institute of Standards and Technology-led voluntary framework.

The Internet PACT Act would hold platforms accountable by:

- requiring large online platforms to provide due process protections to consumers through a defined complaint system that processes reports and notifies users of moderation decisions within twenty-one days, and allows consumers to appeal content moderation decisions;
- amending Section 230 to require that large online platforms remove court-determined illegal content and activity within four days; and
- allowing smaller online platforms to have more flexibility in responding to user complaints, removing illegal content and acting on illegal activity, based on their size and capacity.

The Internet PACT Act would protect consumers by:

- exempting the enforcement of federal civil laws from Section 230 so that online platforms cannot use the provision's blanket immunity as a defense when federal regulators pursue civil actions online;
- allowing state attorneys general to enforce federal civil laws against online platforms; and
- requiring the GAO to study and report on the viability of an FTC-administered whistleblower program for employees or contractors of online platforms.

---

- repeal Section 230 in whole (only Republican-sponsored);
- restrict the types of activities protected by Section 230 (typically bipartisan);
- impose new obligations on platforms seeking to use the Section 230 defense (typically bipartisan); and
- seek to address perceived political biases/censorship (with a single exception, only Republican-sponsored).

On Thursday, less ambitious, but nonetheless critically important, bipartisan legislation (the Platform Accountability and Transparency Act – PATA) was reintroduced by Senators Coons, Cassidy, Klobuchar, Cornyn, Blumenthal and Romney to require the platforms to share more data with the public and researchers. This effort follows introduction of a similar bill in December 2022, which followed the release of a discussion draft in December 2021. There is no Section 230 element in the current version of PATA; the 2022 version of PATA (Section 8) conditioned Section 230 immunity on compliance with the access requirements of PATA.

PATA introduces three mechanisms to promote platform transparency:

- ***Researcher-specific data access***: platforms would be required to provide independent researchers approved by the National Science Foundation with access to platform data, subject to privacy and cybersecurity protections.
- ***Limited legal safe harbor for automated data collection***: a safe harbor would prevent platforms from suing or criminally accusing public interest researchers who use automated means to collect public-facing platform information, so long as they use appropriate privacy safeguards for the data they collect. Researchers report that the possibility of such liability is a significant obstacle to their ability to analyze platform behavior.
- ***Enhanced transparency through disclosures****:* PATA would require covered platforms to disclose certain information that would provide a much stronger understanding of what is happening on platforms that is currently opaque. Specifically, platforms would be required to report information about:
    - ***Viral content***: metrics about content that has gone viral or has been distributed from major public accounts (*e.g.,* data about the extent of dissemination, engagement, audience, and whether the content was recommended, amplified, or restricted).
    - ***Ad library****:* information about advertisers and ads they have run, and metrics about dissemination, reach, engagement and targeting criteria.
    - ***Algorithmic design***: a semi-annual description of the data used as inputs in ranking or recommendation algorithms and how that data affects the algorithm's output; information about each algorithm's optimization objective; information about how content is scored or ranked; and information about how companies assess new products.
    - ***Content moderation***: statistics about content that a platform took action against, broken down by the categories such as the policy that was violated; geographic and demographic factors; data about the number of times violating content was viewed; information about how violating content was identified; the extent to which violating content was recommended, amplified or restricted; and estimates about the prevalence of violating content.

A separate act – The Safeguarding Against Fraud, Exploitation, Threats, Extremism and Consumer Harms (SAFE TECH) Act was re-introduced in 2023 by Democratic sponsors in the Senate, the original bill having been introduced in February 2022. Under this proposed bill, platforms would:

- be unable to use Section 230 as a defense in cases related to ads or other content they are paid to make available;
- be unable to use Section 230 to avoid injunctive relief (*e.g.*, a court order compelling the company to take some action) if it fails to "remove, restrict access to or availability of, or prevent dissemination of material" that could cause "irreparable harm";
- be unable to use Section 230 in cases involving federal or state civil rights laws, federal or state antitrust laws, federal or state stalking, harassment or intimidation laws, or human rights law or in wrongful death actions; and
- need to prove that they are "a provider or user of an interactive computer service," and that they are "being treated as the publisher or speaker of speech provided by another information content provider."

One other avenue to address the power of the platforms has been the antitrust laws. Proposed legislation has included the American Innovation and Choice Online Act ("AICOA"), which would have limited the ability of the platforms to prefer their own products and services over their competitors' products and services, and the Open App Market Act, which would have limited app stores owned/controlled by companies with over 50 million US users from imposing anti-competitive restrictions. Neither bill passed in the 117th Congress, although both passed out of the Judiciary Committee.

Many attribute the failure of the AICOA to the inability to reconcile differences of opinion over content moderation, with Democrats keen to ensure that content moderation would be protected. In March, hearings were held on both pieces of legislation in the hope of resuscitating them. Separately, the Advertising Middlemen Endangering Rigorous Internet Competition Accountability (AMERICA) Act, which would prohibit large digital advertising firms from owning more than one part of the digital advertising market, was introduced at the end of March. It too had been introduced in the 117th Congress under the name the Competition and Transparency in Digital Advertising Act.

In April, Senators Graham and Blumenthal reintroduced bipartisan legislation (the Eliminating Abusive and Rampant Neglect of Interactive Technologies Act or EARN IT Act) to amend Section 230. The legislation is designed to encourage the platforms to address online child sexual exploitation, by removing the blanket immunity provided by Section 230 for violations of laws related to online child sexual abuse material. Similar to the other efforts outline above, prior attempts to pass the legislation had failed.

## Republican-Controlled State Legislative Efforts

In 2021, Republican-controlled states tried to end-run around the paralysis at the federal level (*see* Free Speech Challenges to Florida and Texas Social Media Laws). Florida passed SB 7072, designed to prevent platforms from banning (deplatforming) politicians. The target, according to Gov. Ron DeSantis, was discrimination against "freedom of speech as conservatives" in favor of "Silicon Valley ideology" A federal judge blocked the law, and an appeal is pending before the Supreme Court. A similar effort in Texas, HB 20, is also in legal limbo. These bills were in part a reaction to deplatforming of former President Trump following the January 6th insurrection, and were challenged by two trade groups (NetChoice and Computer & Communications Industry Association) on First Amendment grounds.

## The Supreme Court Takes a Pass

In May, in two eagerly-awaited Section 230 cases, *Gonzalez v. Google* and *Twitter v. Taamneh*,[9] the Supreme Court sidestepped the issue of whether Section 230 immunizes the platforms from lawsuits alleging that their recommendation algorithms promoted terrorist activities, leaving for another day the fate of a statutory provision that, in the words of Robert Barnes and Cat Zakrzewski, writing in the Washington Post, "has emerged as a lightning rod in the politically polarized debate over the future of online speech."

## Concluding Thoughts

This is very much of a watch-this-space moment for the future of both data privacy and platform transparency. The European Union, riding multiple waves of the Brussels Effect, has managed to legislate on both and, in the process, created a global standard, and a template. The GDPR enshrined data privacy protections that had evolved over time in Europe, though firmly aligned with its history and culture, and the DSA, writing on a clean slate, brought to an end an era of self-regulation of tech platforms, which so many had, and have, found wanting.

The growing patchwork of state data privacy legislation and the pass taken last month by the Supreme Court on Section 230 (not to mention the 51 bills tracked by the Section 230 Reform Hub to address Section 230, in one of four ways – admittedly covering initial and reintroduced bills) underscore the urgency of congressional action, and the current configuration on Capitol Hill means, of necessity, that it will have to be bipartisan. Protection of data privacy really should be uncontroversial, and therefore bipartisan. Regrettably, platform transparency has become hostage to culture war dysfunction. At the very least, in light of the failure of self-regulation, what should be bipartisan should be a public, data-driven understanding, based on mandated access by independent researchers, of how the platform algorithms operate and how they impact society (ranging from the effects on children, to mental health, national security and democracy).

My proposed solution for data privacy would be a reintroduced ADPPA, and my proposed solution for platform transparency would be the very recently reintroduced PATA, leaving Section 230 to be addressed another day.

* * *

**Mark S. Bergman**
**7Pillars Global Insights, LLC**
**Washington, D.C.**
**June 10, 2023**

[9] The *Gonzalez* case involved an allegation that YouTube contributed to the death of an American citizen killed in the 2015 terrorist attacks in Paris because its algorithms recommended a pro-ISIS video to viewers who watched similar material. The *Taamneh* case sought to hold Twitter liable for aiding and abetting an ISIS attack in Turkey by failing to block or remove pro-extremist messaging from its platform and that its algorithms matched ISIS-related content to users most likely to be interested in that content.