7

Privacy, Big Data, and Free Speech

The collection and processing of data is a necessary and inevitable aspect of operating a social media platform. Leaving aside business models (on which more later), it is simply impossible to run a platform without access to user data. Given the sheer volume of content on any popular platform (e.g., Facebook, Instagram, Twitter/X, YouTube, TikTok), platform operators have to make constant choices about what content to serve up or recommend to users. And unless they make those choices randomly (a truly horrific thought), those choices are necessarily made for users based on what the platform knows about individual users' interests, preferences, and life experiences (i.e., based on user data). In other words, the collection, storage, and processing of user data is an integral part of how platforms create a good – indeed, even tolerable – user experience. For that reason, no serious commentator or regulator proposes banning the collection and processing of user data.

In addition, for most social media platforms, their current business model necessarily requires the collection and processing of data. The reason for this, as noted in Chapter 3, is that for the major platforms, targeted advertising is their primary source of revenue. Take Meta, the owner of Facebook and Instagram. In required regulatory filings with the US Securities and Exchange Commission, Meta reported revenue of almost $135 billion in 2023; and as Meta itself acknowledges, "[s]ubstantially all of our revenue is currently generated from advertising on Facebook and Instagram" (reading the fiscal report suggests that 99 percent of Meta revenues are ad-based).[1] While precise comparative numbers are difficult to obtain regarding Twitter/X (because Twitter/X is no longer publicly traded, after Elon Musk's purchase of the company), press reports suggest that 70–75 percent of Twitter/X's 2023 revenue of

[1] *Form 10-K, Meta Platforms, Inc.* 60–61 (2024), https://d18rnop25nwr6d.cloudfront.net/CIK-0001326801/c7318154-f6ae-4866-89fa-f0c589f2ee3d.pdf.

approximately $3.4 billion was from advertising sales. And, it should be noted, advertising revenues at Twitter/X appear to have declined around 40 percent from 2022 to 2023, in the wake of Musk's takeover.[2] Even YouTube, which earns substantial revenues from subscription services such as YouTube Music, YouTube Premium, and YouTube TV, earned over twice as much ($31.5 billion)[3] from ad sales in 2023 than it earned from subscription services in which users pay for content ($15 billion).[4] And finally, while TikTok's comparable revenue numbers are hard to nail down (because TikTok is privately owned by ByteDance, a Chinese company), analysts appear to agree that the vast majority of TikTok's approximately $16 billion in US revenue in 2023 was derived from advertising.[5]

All of which is to say that targeted advertising is the revenue mainstay of the social media business, and a highly profitable mainstay at that. But effective targeted advertising of course requires user data. Just as user data tells platforms what content to serve up to individual users, so too the same data allows them to predict which advertisements are going to be most enticing to individual users, given their consumer preferences. This in turn means that user data permits the platform to match potential advertisers to their most potentially lucrative audience – something that is not possible to the same extent with traditional print and broadcast advertising. The enormous numbers quoted on the previous paragraph reflect this simple fact – the product that platforms sell to advertisers is extraordinarily valuable.

On the flip side, restricting the collection and/or use of data for targeted advertising has potentially perverse effects. The less user data that platforms have access to, the less precise their targeting, and so the less valuable the online advertising they sell is to advertisers – a point that Meta itself acknowledges in its regulatory filings.[6] But given that the costs of operating a social media platform are largely fixed, if platforms have to charge less for individual ads (because they provide less value to the ad buyer), they necessarily will

[2] Ashley Belanger, *Elon Musk's X Ad Revenue Reportedly Fell $1.5B This Year Amid Boycotts*, ARSTECHNICA (Dec. 13, 2023), https://arstechnica.com/tech-policy/2023/12/stop-comparing-xs-dismal-ad-sales-to-twitters-past-success-x-exec-says/.

[3] *Form 10-K, Alphabet, Inc.* 35 (2024), https://abc.xyz/assets/43/44/675b83d7455885c4615d848d52a4/goog-10-k-2023.pdf.

[4] Todd Spangler, *YouTube and Google Subscription Services Hit $15 Billion in 2023 Revenue*, VARIETY (Jan. 30, 2024), https://variety.com/2024/digital/news/youtube-google-subscription-services-annual-revenue-1235892210/.

[5] *TikTok's US Revenues Hit $16bn as Washington Threatens Ban*, FINANCIAL TIMES (March 15, 2024), www.ft.com/content/275bd036-8bc2-4308-a5c9-d288325b91a9.

[6] *Form 10-K, Meta Platforms, Inc*, *supra* n. 1, at 61–62.

have to sell a greater volume of advertising. But that, of course, degrades the quality of the service the platforms are providing, and so the user experience.

Finally, if we prohibited targeted advertising altogether, as the European Union (EU) has done with respect to minors under the age of eighteen in its recent Digital Services Act (DSA),[7] this realistically leaves only one business model open to platforms, which is to charge users for their services (after all, *someone* has to pay for those servers and coders and content moderators). But in fact, common sense suggests that paying for services is *not* what most users want.

Until recently, that this was the case was mainly speculation, albeit speculation consistent with common knowledge of human nature. But there is now an ongoing empirical experiment regarding whether users would rather watch advertising than pay for their social media accounts. In November of 2023, in response to increasingly strict EU regulations of data practices (including in the DSA and in its companion the Digital Markets Act), Meta began to offer, only in Europe, a paid, ad-free subscription to Facebook or Instagram. The price was €9.99 per month for a basic subscription, or €12.99 to use the Apple or Android App[8] (which at the August 2024 exchange rate is about $10.90 and $14.10). Furthermore, in response to complaints from European regulators, Meta later offered to halve the price for these services.[9] As of this writing (fall 2024), good data is not yet available about how many Europeans have switched to the ad-free model, but that no doubt will change. However, in the summer of 2024 European regulators (somewhat strangely) formally charged that this paid service option itself violated European law,[10] suggesting that the regulators themselves do not see even a fairly cheap monthly fee as an alternative to Meta's current advertising/user-data-based model.

So, the collection and processing of user data by social media platforms is inevitable, and in some sense desirable. But at the same time, as discussed in detail in Chapter 3, Big Data creates some serious social problems and

[7] *Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 (Digital Services Act)* Art. 28(2), https://eur-lex.europa.eu/eli/reg/2022/2065/oj; European Commission, Directorate-General for Communications Networks, Content and Technology, *The Digital Services Act (DSA) Explained – Measures to Protect Children and Young People Online* (2023), https://data.europa.eu/doi/10.2759/576008.

[8] Facebook Help Center, *Subscription for No Ads*, www.facebook.com/help/2620384466840 66.

[9] Andrew Hutchinson, *Meta Offers to Halve the Price of Its Ad-Free Subscription Package in the EU*, Social Media Today (March 19, 2024), www.socialmediatoday.com/news/meta-offers-to-halve-the-price-ad-free-subscription-package-in-eu/710782/.

[10] Kelvin Chan, *European Union Says Meta Breaking Digital Rules with Paid Ad-Free Option for Facebook and Instagram*, PBS News (July 1, 2024), www.pbs.org/newshour/world/european-union-says-meta-breaking-digital-rules-with-paid-ad-free-option-for-facebook-and-instagram.

concerns which, unlike many of the overblown fears noted in Chapters 1 and 2, appear to have a solid empirical foundation. And therein lies the conundrum that this chapter examines: We have good reasons to regulate social media platforms' data practices, but such regulation itself has the potential to harm not only the platforms but also their users and society at large. And, as we shall see, data privacy also sometimes raises serious free speech concerns. How to resolve these tensions is extremely difficult to figure out; but it is a problem that cannot be avoided.

7.1 DATA PRIVACY IN ACTION: THE GENERAL DATA PROTECTION REGULATION (GDPR) AND CALIFORNIA CONSUMER PRIVACY ACT (CCPA)

Given the many legitimate and serious privacy concerns raised by the day-to-day operations of social media platforms, it is unsurprising that many jurisdictions (though not the US Congress) have adopted privacy regulations directed at technology companies, including social media. Most importantly, long before the EU enacted the DSA, in May 2018 the EU's GDPR came into effect.[11] As Professors Meg Leta Jones and Margot Kaminski discuss, while the GDPR has been described in the United States as a law focused on consumer consent, this is in fact not entirely accurate.[12] Rather than being a traditional data privacy law on the American model (which does typically focus primarily on consent), the GDPR regulates *data* and data processing.[13]

The core of the GDPR, contained in Article 6, is a rule providing that holders of personal data may process it *only* for one of six listed reasons – all other processing is illegal.[14] And this restriction applies to all holders of data, not just the original collector or the firm with which the subject of the data has a relationship (contractual or otherwise).[15] Furthermore, while consent is indeed the first justification for data processing listed in Article 6, it is far from the only one. To the contrary, Jones and Kaminski argue that the sixth justification – that the "processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party"[16] – is the one that most

[11] General Data Protection Regulation 2016/679, 2016 O.J. (L 119), https://gdpr-info.eu/ (henceforth "GDPR").

[12] Meg Leta Jones and Margot E. Kaminski, *An American's Guide to the GDPR*, 98 Denver L. Rev. 93, 95 (2021).

[13] *Ibid*. at 106–08.

[14] GDPR, *supra* note 11, Art. 6(1); Jones and Kaminski, *supra* note 12, at 108.

[15] Jones and Kaminski, *supra* note 12, at 107.

[16] GDPR, *supra* note 11, Art. 6(1)(f).

firms are likely to rely upon, at least in part because consent requirements in the GDPR are far more onerous than under the typical American privacy law.[17] Finally, while the "legitimate interests" provision would appear to permit extensive data processing by tech firms in the course of selling advertising or other business processes, it is important to note that the ability to process data under this justification may be "overridden by the interests of fundamental rights and freedoms of the data subject,"[18] creating substantial uncertainty and leaving lots of scope for regulatory restrictions on data processing.

In addition to restricting data processing, the GDPR also grants important rights to the individual subjects of personal data. While a full description of those rights is not possible here, important elements include extensive rights of detailed notification regarding data collection, storage, and processing;[19] a right to access stored and/or processed data;[20] a right to correct inaccurate data;[21] and a right to object to continued processing of data by government entities or private entities under the "legitimate interest" justification discussed earlier, though, importantly, the right to object is *not* absolute and can be outweighed by "compelling legitimate grounds for the processing."[22] Most famously, the GDPR also codifies the "right to be forgotten,"[23] which had been recognized earlier by the Court of Justice of the European Union (CJEU), the top court in the EU, in a case called *Google Spain*.[24] This provision effectively permits data subjects to demand the erasure of data no longer needed for processing – though as with many GDPR "rights," this one is limited and can be overridden by, inter alia, "exercising the freedom of expression and information."[25]

There is little doubt that the GDPR, through these and many other provisions, establishes one of the most comprehensive data regulation regimes in the world. And because of the huge size of the European market, there are strong indications that firms, including social media platforms, subject to the GDPR have chosen to follow its basic provisions worldwide simply to

[17] *See* Jones and Kaminski, *supra* note 12, at 108–09.
[18] GDPR, *supra* note 11, Art. 6(1)(f).
[19] *Ibid*. Arts. 13, 14.
[20] *Ibid*. Art. 15.
[21] *Ibid*. Art. 16.
[22] *Ibid*. Art. 21.
[23] *Ibid*. Art. 17.
[24] Case C-131/12, Google Spain SL v. Agencia Española de Protección de Datos, 2014 E.C.R. 317 ("*Google Spain*"). For an excellent discussion of the *Google Spain* decision and its problems, see Robert C. Post, *Data Privacy and Dignitary Privacy*: Google Spain, *the Right to Be Forgotten*, *and the Construction of the Public Sphere*, 67 Duke L.J. 981 (2018).
[25] GDPR, *supra* note 11, Art. 17(3)(a).

avoid the costs of following diverse rules – the so-called Brussels Effect.[26] The GDPR also creates important personal rights that, if invoked, could return substantial power to individual data subjects. There is no question that these provisions, in combination, have forced firms to change their data practices in significant ways, most importantly by obtaining more frequent and prominent user consent. Furthermore, as a result of a huge GDPR-based fine imposed on Meta in 2023, the GDPR is likely to impact firms' ability to engage in cross-border data transfers.[27] It should be noted, however, that there is absolutely no evidence that the GDPR has substantially restricted or interfered with the fundamental operations and data practices of the major tech companies since it became effective in May of 2018.

Aside from the GDPR, probably the most important privacy protection statute of recent years is the CCPA, which was inspired by the GDPR though it is more limited in scope. The CCPA was first enacted in 2018, with an effective date of January 1, 2020; but in 2020 California voters enacted the California Privacy Rights Act (CPRA), which significantly amended the CCPA and created the CCPA in its current form.[28] The original CCPA's primary provisions give California consumers the right to request information about a firm's data collection, retention, and transfer practices,[29] a qualified right to have personal data deleted,[30] and a right to opt out of the sale of personal information.[31] The CPRA added to this a right to have data holders correct inaccurate data,[32] and a right to limit the use and disclosure of especially sensitive data such as financial information, geolocation data, and genetic data.[33] Since the adoption of the CCPA, nineteen other states (as of summer 2024) have also adopted data privacy laws, many of which are modeled on the CCPA and contain many of the same restrictions.[34]

Even as amended, the CCPA (and its imitators) do not contain the sorts of broad restrictions on data processing, or extensive consumer rights, contained

[26] Anu Bradford, The Brussels Effect: How the European Union Rules the World (2021).

[27] Adam Satariano, *Meta Fined $1.3 Billion for Violating E.U. Data Privacy Rules*, N.Y. Times (May 22, 2023), www.nytimes.com/2023/05/22/business/meta-facebook-eu-privacy-fine.html.

[28] Cal. Civ. Code §§ 1798.100–178.99-100. For a good, short summary of the CCPA, *see California Consumer Privacy Act (CCPA)*, Office of the Attorney General, State of California Department of Justice, https://oag.ca.gov/privacy/ccpa#.

[29] Cal. Civ. Code §§ 1798.110, 1798.115.

[30] Cal. Civ. Code § 1798.105.

[31] Cal. Civ. Code § 1798.120.

[32] Cal. Civ. Code § 1798.106.

[33] Cal. Civ. Code § 1798.121. Sensitive data is defined in Cal. Civ. Code § 1798.140(ae).

[34] *UD Data Privacy Guide*, White & Case (July 2, 2024), www.whitecase.com/insight-our-thinking/us-data-privacy-guide.

in the GDPR. But these laws, the CCPA most prominently (because of California's sheer size and the fact that it is home to most of the major American tech giants), have forced firms, including social media platforms, to provide consumers with access to their data and a means to delete or correct it.[35] And by incorporating a right to have data deleted, the CCPA and its progeny have arguably imported into the United States a version, albeit a watered down one, of the EU's "right to be forgotten."

7.2 TWO KINDS OF PRIVACY

Before getting into the merits of the GDPR, CCPA, and their brethren, it is useful to take a step back and consider what exactly it is that these laws are trying to accomplish. The obvious answer is that they seek to protect the privacy of individuals. But what exactly is "privacy?" In a deeply insightful article about the CJEU's *Google Spain* decision, which as noted earlier first created a Europe-wide right to be forgotten (which was later incorporated into the GDPR), Yale law professor (and former dean) Robert Post convincingly argues that the concept of privacy is widely misunderstood. In particular, he argues that in *Google Spain* the CJEU conflated two distinct forms of privacy.[36] And he further demonstrates that the CJEU's mistake, which has been carried over into Article 17 of the GDPR, raises serious questions about the scope and legitimacy of the right to be forgotten.

To understand Post's argument, one must deconstruct it a bit. First, let us consider the two forms of privacy he identifies. One form of privacy, the primary concern of the GDPR as well as its predecessor law that was invoked in *Google Spain*, is a right to control how one's personal data is stored and processed by bureaucratic entitles that collect such data, including large corporations such as Google and the major social media platforms. This particular kind of privacy does not have much to do with the *kind* of personal data being stored – it applies equally to innocuous information, such as one's phone number, and highly personal information like medical records. The primary goal in protecting this form of privacy, which we might call "data privacy," is an instrumental one, of ensuring that bureaucratic entities do not cumulate large hoards of data about individuals without need or reason.[37]

The other form of privacy Post identifies is entirely different. It is a deeply personal interest on the part of human beings in not having intensely personal

[35] *See*, e.g., *How Can You Delete Your Information?*, Facebook Privacy Center, www.facebook.com/privacy/dialog/delete-your-information/.

[36] Post, *supra* n. 24.

[37] *Ibid*. at 993–94, 1000–01.

information disclosed to the world. This form of privacy is dignitary in nature and turns entirely on the form of "data" at issue. Indeed, the deep question raised by this form of privacy, which we can call "dignitary privacy," is what kinds of things *are* so deeply personal that it is offensive and should be unlawful to publicly disclose such facts (which we presume are true) about other people. Dignitary privacy, it should be noted, is *not* the focus of modern data privacy statutes but rather is protected by other forms of law, including in the United States primarily by tort law (in particular, the tort of public disclosure of private fact).[38]

It is difficult in this limited space to do full justice to Post's subtle argument, but in short, he argues that in *Google Spain* the CJEU failed to recognize the existence of these two distinct forms of privacy and so ended up with an incoherent analysis. The *Google Spain* case was brought by a Spanish lawyer who was tied to an embarrassing financial episode many years in the past, the fact of which appeared in a Spanish newspaper. He claimed that one of the first links that appeared in a Google search of his name was to the online archives of that newspaper (which had been digitalized well after the original events). This, he said, caused him severe reputational harm, even though given the passage of time the original episode was now irrelevant. The CJEU accepted the lawyer's argument that linking to this archive violated EU privacy laws, because it constituted the processing of personal data in a way that was "irrelevant or no longer relevant." Therefore, the Court held, Google had a legal obligation to block links to the newspaper archive when the lawyer's name was searched. Notably, however, the Court held that the Spanish newspaper itself was *not* required to remove the relevant content from its online archive.[39]

Post convincingly argues that this reasoning is fundamentally flawed. The basic problem was that the main legal provisions that the CJEU invoked to justify its result was the GDPR's legal predecessor, which like the GDPR focused on data privacy, *not* dignitary privacy. As a result, the Court concluded that Google's legal error was in processing *any* out-of-date personal data in this way, regardless of whether it was highly private or embarrassing.[40] But this is obviously nonsense. Surely the *Google Spain* plaintiff would not have complained if a search of his name linked to a thirty-year-old marriage notice, or to an article about his high school athletic prowess. It was the *embarrassing* nature of the information that was crucial to his case, a point the Court simply ignored or missed.

[38] *Ibid*. at 991 and n.36.
[39] *Ibid*. at 995–98.
[40] *Ibid*. at 997–98 (*citing Google Spain*, para. 94).

More fundamentally, Post argues that the CJEU made a categorical error. Its reasoning regarding data privacy would make sense as applied to a database held by, say, a grocery store tracking customer purchases. It might even make sense as applied to Google's own collection, storage, and processing of data about its users (mainly for the purpose of selling targeted advertising). But to apply it to Google's search feature makes no sense. The whole purpose and very function of a search engine is to link to any and all information – which means that such "processing" can *never* be "irrelevant or no longer relevant." More profoundly, Post argues that in the modern world, Google plays a role similar to that played by newspapers in creating a public sphere of information within which public opinion can form, and so democratic politics can operate. And given that fact, it was particularly nonsensical for the CJEU to order Google to stop linking to the relevant newspaper article but permit the newspaper itself to keep the article accessible to the public – because after all, it was the newspaper that was hosting the offending content, not Google.[41]

None of which is to say, and Post does not claim, that disclosure of embarrassing and out-of-date facts can never pose privacy concerns. It certainly might, and while under US law it is almost impossible to win a tort claim under circumstances such as those in *Google Spain* because of the First Amendment, the same need not be true in Europe. But what Post is saying is that if such privacy concerns are to be addressed, it was profoundly mistaken to focus on Google rather than the underlying newspaper archive.

What is the implication of these arguments for social media platforms? To begin with, a sharp distinction needs to be made (as the CJEU failed to do) between such platforms' collection and processing of their users' data, and their hosting functions in disseminating user posts which might well contain personal data/information. The former activities, collecting user data and using it to engage in targeted advertising, fall squarely within the world of data privacy, and so are appropriate subjects of laws such as the GDPR and CCPA (which is not to say that those laws do not have their own issues, on which more later). But when platforms host user content, very different considerations arise.

This is because, even more than search engines like Google, modern social media platforms are the locus of public discourse and so the locus within which public opinion is formed. Therefore, for the same reasons that applying data privacy principles to a Google search is incoherent, so even more so with social media platforms in their hosting capacity. Such platforms are about facilitating *discourse*, and it is deeply problematic for any governmental actors,

[41] *Ibid*. at 1063.

whether EU regulators or US judges, to impose significant restraints on the subjects of that discourse, or to force platforms to police it for them. There are surely times when individual posters can be held liable for invasions of dignitary privacy – posting nonconsensual pornography being an obvious example. It may even make sense, as discussed in the next chapter, to narrow Section 230 slightly to impose obligations on platforms to police such awful content, in extreme and well-defined circumstances. But imposing wholesale obligations on platforms to protect privacy on their platforms makes little sense, even in the face of strong and legitimate dignitary privacy interests.

7.3 DATA PRIVACY AND PLATFORMS

Let us now consider the core of the privacy-based attacks on social media, which target not the hosting function but the tendency of social media companies to collect, store, and process enormous amounts of user data. To begin with, it should be reiterated that such data practices lie at the heart of the business models of most modern platforms. They are, in essence, what permits the platforms both to provide services at all and to provide them without charging users. Any data regulation that significantly interferes with those activities or business models almost certainly does more harm than good and is also likely to prove extremely unpopular with the social media-using public (which is to say, most of the public). As a consequence, even under the GDPR and CCPA, social media platforms continue to engage in their core data processing, either under consent theories or under the GDPR's "legitimate interest" exception discussed earlier.

On the other hand, there is no question that laws like the GDPR and CCPA have forced firms, including social media platforms, to take substantial steps to ensure transparency and data integrity. In most respects, this is undoubtedly a positive outcome. But it too has a significant downside: Such data protection steps are expensive. Obviously, few will shed tears over behemoths such as Facebook/Meta and YouTube/Google having to expend some of their seemingly limitless funds on data protection. It must be recognized, however, that expensive regulatory obligations inevitably act as barriers to entry, preventing startups and small firms from entering into these markets because they cannot afford the same levels of protections. And yet, aside from privacy, one of the prime complaints about the tech giants is their market power. Adopting regulatory regimes that accentuate that market power seems questionable policy.

One way to mitigate this concern might be to exempt small firms from data protection rules; but that is also somewhat problematic. For one thing, insofar as avoiding data regulation would reduce smaller platforms' costs, it might

permit them to provide a more seamless experience. But if so, this would have the perverse effect of incentivizing consumers to migrate to platforms that do not protect their data. And furthermore, regimes that exempt small actors from burdensome regulations have the (also perverse) consequence of discouraging them from growing beyond a certain point. But of course that also helps to entrench incumbent market power. Smaller platforms are less valuable to users than large platforms because the latter give access to larger audiences, a phenomenon called "network effects" in the economic literature.[42]

None of which is to say that protections for data privacy are a mistake. They clearly are not, and given the emerging consensus about the need for data protection (demonstrated by extensive regulatory steps taken around the United States and the world), we can expect regulatory initiatives to continue. But it is to emphasize that all regulation, even such seemingly innocuous steps as data transparency and protection rules, comes with cost. And sometimes regulation can have unexpected and potentially severe unintended consequences, as illustrated by a CJEU decision aggressively reading the GDPR to impose strict limits on data transfers between the EU and the United States, which potentially puts at risk $7.1 *trillion* dollars of transatlantic economic interactions.[43]

What this suggests is that, as with all regulatory initiatives in the complex and fraught space of new technologies, it is crucial that regulators remember to adopt a posture of humility. One aspect of the GDPR that is of particular concern in this respect is its flat ban (in Article 6) on all data processing, except for six specified reasons. Admittedly, many of the reasons are stated in relatively broad terms (notably the exception for pursuing the "legitimate interests" of the data processor), but there is still a degree of arrogance in assuming that regulators can predict the universe of possible, legitimate uses of data. Combine this with the fact that the GDPR explicitly states that the right to engage in data processing to pursue "legitimate interests" can be "overridden by the interests or fundamental rights and freedoms of the data subject," without specifying what those "rights and freedoms" are,[44] and one realizes that the GDPR introduces an enormous amount of uncertainty regarding permissibility of data processing. In a world of rapidly changing technology,

[42] *What Is the Network Effect?*, Wharton Online, Wharton School of Business, University of Pennsylvania (Jan. 17, 2023), https://online.wharton.upenn.edu/blog/what-is-the-network-effect/.

[43] Monika Zalnieriute, *Data Transfers after* Schrems II: *The EU-US Disagreements over Data Privacy and national Security*, 55 Vand. J. Transnat'l L. 1 (2022).

[44] GDPR, *supra* note 11, Art. 6(1)(f).

including notably the extraordinary recent growth of artificial intelligence, this seems precisely the wrong way to engage in regulation.

7.4 DATA IS SPEECH

Beyond the rather complex and technical issues discussed earlier, moreover, data regulations raise a more fundamental concern that should not be lost sight of. Restrictions on the use and dissemination of data are, in practice, restrictions on speech. In a law review article published some years ago, law professor Jane Bambauer posed the question, "is data speech?"[45] And her unequivocal conclusion was "yes," that data is indeed speech. Furthermore, in its leading opinion addressing regulation of data, a six-Justice majority of the US Supreme Court strongly endorsed (albeit in nonbinding language) the proposition that "information [i.e., data] is speech."[46] What this means is that at least in the United States, laws and regulations aimed at data practices implicate the First Amendment. What practical limitations does this impose on the ability of states (and eventually, presumably, Congress) to adopt data privacy legislation?

One implication seems clear: It would almost certainly be unconstitutional for the United States or any individual US state to adopt the strong form of a "right to be forgotten" established in the EU in the *Google Spain* case and the GDPR. The reason is that enforcement of the right to be forgotten is, in plain English, a direct restriction on speech. It forbids someone – in *Google Spain*, Google – from sharing information (i.e., speaking) because the information at issue is private, and so its disclosure causes dignitary harm. In the EU, with its relatively weak protections for speech vis-à-vis privacy and other social interests, this may be permissible (though it should be noted that even the GDPR recognizes that free speech principles will sometimes trump the right to be forgotten[47]). But in the United States, under the First Amendment as interpreted by the US Supreme Court, it probably is not.

Admittedly, the Court has never adopted a definitive rule regarding how to reconcile free speech and privacy concerns; but in a series of cases from the 1970s and 1980s, culminating in a decision from 2001 involving information obtained via wiretapping, the Court has consistently held that the right of the

[45] Jane Bambauer, *Is Data Speech?*, 66 Stan. L. Rev. 57 (2014).

[46] Sorrell v. IMS Health, Inc., 564 U.S. 552, 570 (2011). I have argued elsewhere that this is clearly correct analysis under current law. Ashutosh Bhagwat, Sorrell v. IMS Health: *Details, Detailing and the Death of Privacy*, 36 Vermont L. Rev. 855, 867 (2012).

[47] GDPR, *supra* note 11, Art. 17(3)(a).

press to publish private information trumps the privacy rights of individuals.[48] And while the US Supreme Court has never directly ruled on facts analogous to *Google Spain*, the California Supreme Court has. In a case from 2004 in which an individual sued for invasion of privacy when a television documentary disclosed his thirteen-year-old criminal conviction, that court held, relying on the US Supreme Court cases just mentioned, that the First Amendment flatly prohibited imposing liability for publishing information available in public records.[49] And notably, in doing so the California court overruled a previous decision, predating the key Supreme Court cases, permitting an invasion of privacy claim on similar facts.[50]

Aside from the right to be forgotten, the limitations that the First Amendment places on privacy regulation are less clear. It has been generally assumed that existing, pre-internet laws prohibiting the disclosure of sensitive information such as medical or financial records must be constitutional, even though such laws literally restrict "speech" based on its content (something which in other situations is presumptively unconstitutional). It must be admitted, however, that precisely why, as a technical legal matter, this is so remains unresolved. But nonetheless it seems likely that future courts will uphold laws prohibiting tech companies, including social media platforms, from disclosing user data to the public. Such data may not be as sensitive as, say, medical records, but data regarding what one posts, what one searches, etc., are surely still private matters.

It should be emphasized, however, that while as a predictive matter it is likely that courts will uphold prohibitions on personal data disclosure, First Amendment law in this area remains extremely unsettled. As such, the actual fate of laws regulating data disclosure remains uncertain. Furthermore, even if courts generally uphold such laws by analogy to historical privacy legislation, much more difficult questions arise when the disclosures being prohibited or punished involve information about public figures, especially public officials and other political figures. In those situations, the societal interest in having access to the information rises sharply, especially with respect to government officials and political candidates given the obvious relevance of such information to the democratic process. At the same time, there is an argument to be

[48] Cox Broadcasting Corp. v. Cohn, 420 U.S. 469 (1975) (disclosure of name of rape victim); Smith v. Daily Mail Pub. Co., 443 U.S. 97 (1979) (disclosure of name of juvenile defendant); The Florida Star v. B.J.F., 491 U.S. 524 (1989) (disclosure of name of rape victim); Bartnicki v. Vopper, 532 U.S. 514 (2001) (disclosure of contents of illegally intercepted cell phone conversation).

[49] Gates v. Discovery Communications, Inc., 34 Cal.4th 679, 101 P.3d 552 (2004).

[50] *Ibid*. at 563 n.9 (*overruling* Briscoe v. Reader's Digest Association, Inc., 4 Cal.3d 529 (1971)).

made that public officials and politicians, by voluntarily entering the political arena, have consented to a reduced right of privacy. This is not to say that even politicians have *no* reasonable expectations of privacy; of course they do, just reduced ones. As a result, whether in practice courts grant First Amendment protection to otherwise illegal disclosures of private information about public officials and political candidates is likely to turn on a case-by-case weighing of privacy rights against the social interest in public disclosure.

The transparency rights granted in the GDPR, CCPA, and other laws raise fewer serious First Amendment concerns. Requiring firms to disclose their data collection and storage practices, while technically a form of "compelled speech," seem likely to be upheld as routine commercial disclosures.[51] Similarly, granting individuals/users the right to access data about them held by firms and to seek correction of inaccurate data,[52] also do not seem to burden free speech in any tangible way.

What raises more difficult questions, however, is the right to erasure of data granted by both the GDPR and CCPA,[53] even as applied to data privacy (its application to dignitary privacy, as discussed earlier, is almost certainly unconstitutional). Requiring a firm to erase data that it is storing is not a literal restriction on speech (unlike a ban on disclosure). Nonetheless, requiring the erasure of data/deletion of information obviously makes it impossible for that data/information to be shared in the future. In that sense, such laws are parallel to prohibitions on making recordings of government officials such as police officers acting in the course of their duties, which have regularly been found to be unconstitutional.[54] Both types of laws have the direct and intended effect of disabling future speech. As such, there can be no doubt that laws requiring the deletion of data will *sometimes* violate the First Amendment.

Unfortunately, current First Amendment law does not provide clear answers to the question of when data erasure requirements are, or are not, permissible. Focusing on the impact of data erasure requirements on social media platforms (which are, after all, the subject of this book), there is an argument to be made that the First Amendment impact of such requirements is minimal. After all, as discussed in Chapter 3, the major platforms rarely if ever intentionally disclose user information, either publicly or to potential competitors in the advertising market, for business reasons. Of course, any deletion of user data has some impact on platforms' ability to sell targeted advertising, but such an indirect and minor impact on speech is unlikely to

[51] Zauderer v. Office of Disciplinary Counsel of Supreme Court of Ohio, 471 U.S. 626 (1985).
[52] GDPR, *supra* note 11, Arts. 15, 16.
[53] GPDR. Art. 17; Cal. Civ. Code § 1798.105.
[54] Ashutosh Bhagwat, *Producing Speech*, 56 Wm. & Mary L. Rev. 1029, 1038–44 (2015).

be found to trump users' legitimate data privacy interests. So, while data erasure requirements undoubtedly would, under certain circumstances, violate the First Amendment, their application to social media platforms is unlikely to do so unless the platform can make the unusual showing that the relevant data was intended, in the future, to be an integral part of some form of public communication by the platform.

In short, while the First Amendment does impose limits on some forms of privacy protection, such as protections for dignitary privacy, run of the mill data privacy provisions generally should survive constitutional scrutiny. And this is especially so as applied to social media platforms, because they are not themselves significant producers of content that is created by using and accessing data.

7.5 THINK OF THE CHILDREN

One final topic that requires some special attention is privacy protections for children, meaning minors under the age of eighteen. Children obviously raise special privacy concerns because of both their vulnerability to exploitation of various forms and their reduced ability to make decisions for themselves. The GDPR recognizes the latter point, for example, by providing that for children under the age of sixteen, consent to data processing must be provided by a parent or guardian.[55] But consent is only the tip of the privacy iceberg when it comes to children.

To begin with, it should be noted that when we refer to "children" or "minors," in the context of the major social media forms we are primarily talking about teenagers between the ages of thirteen and seventeen, because platforms such as Facebook, Instagram, and Twitter/X do not permit children under the age of thirteen from opening accounts.[56] Furthermore, in the United States, federal law prohibits the online collection of data regarding children under the age of thirteen without parental consent, making other privacy protections somewhat moot for that age group (since parental consent is prohibitively expensive to obtain in most situations).[57]

[55] GDPR, *supra* note 11, Art. 8(1).

[56] *Facebook Terms of Service* ¶ 3(1), www.facebook.com/terms.php; *Instagram Terms of Use*, https://help.instagram.com/581066165581870; X *Terms of Service* ¶ 1, https://x.com/en/tos. TikTok and YouTube are concededly different because they do not place age limits, and while they assert that minors under eighteen must have parental consent to use their services, there is no indication that this requirement is enforced.

[57] *Children's Online Privacy Protection Act*, 15 U.S.C. § 6502; *Children's Online Privacy Protection Rule*, 16 C.F.R. § 312.

Admittedly, even for services that have a thirteen-year-old age cut off, enforcement of this requirement has been uneven because platforms do not typically verify the age of users when they create accounts. In response to this concern (and others), in 2022 California adopted a statute, the California Age-Appropriate Design Code Act (CAADCA), which broadly regulates data practices for minors under the age of eighteen.[58] Among other things, the CAADCA requires online providers to estimate the age of potential child users, and adjust its services and default privacy protections accordingly. Strikingly, however, in late 2023 a federal district court in San Jose, California (in a case titled *NetChoice v. Bonta*) enjoined enforcement of much of the California statute. Regarding the age verification provision in particular, the court held that it violated the First Amendment because given the great burden that age verification requirements impose on online providers, the impact of this requirement would be to limit both children's *and* adults' access to some online content (we will return to other aspects of the CAADCA and this decision later).[59] The status and future of age verification requirements are thus very much in flux.

Turning now to specific data privacy policies aiming to protect children, let us start with targeted advertising. As noted at the beginning of this chapter, targeted advertising is the bedrock of social media platforms' business models, the thing which makes it possible for platforms to not charge users. At the same time, targeted advertising directed at children raises special concerns, because of their greater perceived vulnerability to manipulation. The EU's response to this, as noted earlier, has been to prohibit targeted advertising directed at minors, in its DSA.[60] This provision of the DSA, which has been implemented by the major platforms, has not proven terribly controversial. It should be noted, however, that the prohibition does have a secondary impact, of disincentivizing platforms from providing services specially directed at minors, since those services cannot be monetized under this rule.[61] This may well be a reasonable price to pay in order to protect children from manipulation, but it is a tradeoff nonetheless that should be recognized.

There is no such parallel prohibition on targeted advertising for minors under US law at a federal level. California's CAADCA, however, does have a provision that prohibited online businesses from profiling a child by

[58] Cal. Civ. Code §§ 1798.99.28–1798.99.40.

[59] *NetChoice, LLC v. Bonta*, 692 F.Supp.3d 924, 945–46, 950–52 (N.D. Cal. 2023) ("*Bonta*").

[60] *See supra* n. 7 and accompanying text.

[61] Platforms can be expected to continue to provide access to their general services to teenagers, because the marginal costs of doing so are trivial and teenagers will someday soon become adult customers.

default, which in effect prohibited targeted advertising directed at children.[62] As with that law's age verification requirements, however, the *NetChoice v. Bonta* court also held that this provision violated the First Amendment.[63] Furthermore, that decision was almost certainly correct. The US Supreme Court has made it clear that the First Amendment provides strong protections for speech directed at children.[64] Furthermore, the modern Supreme Court has substantially ratcheted up the amount of constitutional protection given to commercial advertising.[65] The combination of these trends makes any flat restriction on advertising to children highly vulnerable, as the *Bonta* court recognized.

Another area of child privacy protections where regulators have been active is data collection and storage. In the EU, as noted earlier, the GDPR requires parental consent before processing the data of children under the age of sixteen if consent is the justification for that processing – and processing is defined to include collection and storage of data.[66] While this is not an absolute bar to collecting children's data, it is a significant impediment. In the United States, federal law flatly bans collecting data for children under the age of thirteen without parental consent, which is generally not plausible to obtain. Finally, California's CAADCA also contains a broad prohibition stating that online providers may not "collect, sell, share, or retain any personal information" about minors under the age of eighteen, except for some narrow, specified purposes (in light of the federal prohibition, the CAADCA's restriction is only relevant to children between thirteen and seventeen years of age).[67]

Perhaps because of its limited application (only to teenagers), this provision too was struck down by the *NetChoice v. Bonta* court based on the conclusion that the inevitable impact of a restriction on data collection was to make it impossible to provide targeted content to teenagers. The court pointed out that such targeted content can be beneficial for minors, especially teenagers who are members of vulnerable subpopulations.[68] It is hard to imagine, however, that a court would reach the same conclusion regarding data collection targeting younger children, which remains effectively prohibited by federal law, given young children's greater need for privacy and the lesser value to them of targeted content.

[62] Cal. Civ. Code § 1798.99.31(b)(2).
[63] *Bonta*, 692 F.Supp.3d at 955–56.
[64] Brown v. Entertainment Merchants Assn., 564 U.S. 786 (2011).
[65] Lorillard Tobacco Co. v. Reilly, 533 U.S. 525 (2001); Sorrell v. IMS Health Inc., 564 U.S. 552 (2011).
[66] GDPR, *supra* note 11, Arts. 4(2), 8.
[67] Cal. Civ. Code § 1798.99.31(b)(3).
[68] *Bonta*, 692 F.Supp.3d at 956–57.

Finally, let us consider special limits on public disclosure of children's data. As noted earlier, prohibitions on disclosure of data are, on their face, direct restrictions on speech, and indeed content-based restrictions, which under current US constitutional doctrine are presumptively invalid. On the other hand, there is a long history of regulating the disclosure of highly sensitive information such as medical histories and financial details, whose constitutionality has not in the past been seriously questioned (at least as to data regarding nonpublic figures). But, it must also be acknowledged, the law in this area is seriously underdeveloped, especially in light of the Supreme Court's strong suggestion that data is speech.

With respect to data regarding *children*, however, it is very hard to believe that restrictions on disclosure would face any serious constitutional scrutiny. The reason, of course, is that children surely have a significantly heightened right and expectation of privacy, given their vulnerability. Furthermore, it is difficult to imagine a situation in which there would be a strong social interest in enabling public access to private facts regarding minors. After all, even famous minors such as child actors or British royals cannot seriously be said to have voluntarily consented to reduced privacy rights in the same way as adults who voluntarily enter the public sphere; even famous minors remain vulnerable to exploitation and long term emotional harms in a way that is categorically different from adults. So at least with respect to data disclosure laws, children probably are special and can legitimately be granted elevated protections.

7.6 PRIVACY AND HUMILITY

For all of these reasons, privacy regulation directed at social media platforms is often addressing a legitimate and serious problem and should not face the same skepticism as the many blatantly ideological attempts to regulate social media discussed in previous chapters. At the same time, especially in the United States, rights of free expression place some significant limits on legislative attempts to protect privacy. And furthermore, in the privacy arena no less than anywhere else, laws and regulations targeting rapidly evolving technologies can easily run up against the law of unintended consequences. As such, regulators should take care not to adopt excessively broad laws (as arguably the EU's GDPR is, in some respects); and more importantly, regulators should stand ready to reconsider their initiatives if unexpected and negative consequences of their actions emerge.