Defending Data

Privacy Protection, Independent Researchers, and Access to Social Media Data in the US and EU

CENTER FOR DEMOCRACY & TECHNOLOGY

KONRAD ADENAUER STIFTUNG

January 2023

The **Center for Democracy & Technology** (CDT) is a 28-year-old 501(c)3 nonpartisan nonprofit organization that fights to put democracy and human rights at the center of the digital revolution. It works to promote democratic values by shaping technology policy and architecture, with a focus on equity and justice. The organization is headquartered in Washington, D.C. and has a Europe Office in Brussels, Belgium.

Defending Data:

Privacy Protection, Independent Researchers, and Access to Social Media Data in the US and EU

Author

Caitlin Vogus

With Contributions by

Asha Allen, Eleftherios Chelioudakis, Samir Jain, Aleksandra Kuczerawy, Jake Laperruque, Alex Lawrence-Archer, Laureline Lemon, Rachel Levison-Waldman, Emma Llansó, Iverna McGowen, Greg Nojeim, Eric Null, Nicole Ozer, Cassie Roddy-Mullineaux, Dhanaraj Thakur, and Joris van Hoboken. Design by Tim Hoagland.

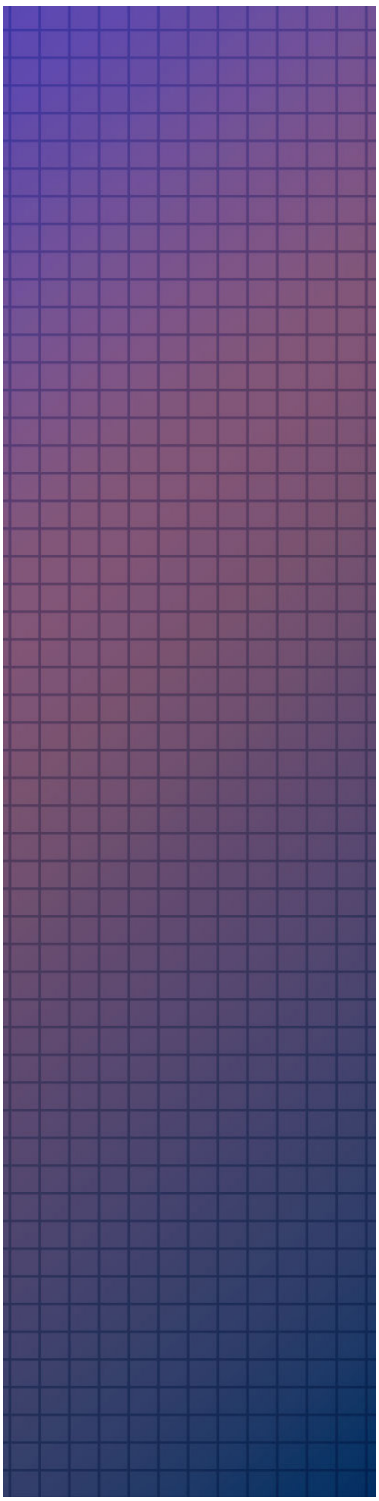
Acknowledgements

This report was made possible by support from the **Konrad-Adenauer-Stiftung**. The views, conclusions and recommendations expressed in this report are solely those of its author(s) and do not reflect the views of the Konrad-Adenauer-Stiftung, its offices, or its employees. This publication is solely intended for information purposes. It may not be used by political parties or by election campaigners or supporters for the purpose of election advertising. This applies to federal, state and local elections as well as elections to the European Parliament.

January 2023

Table of Contents

Executive Summary



Access to social media data by independent researchers is at the forefront of efforts to improve tech transparency. Article 40 of the European Union's Digital Services Act (DSA) requires providers of very large online platforms and very large online search engines to provide vetted researchers with access to data, subject to certain conditions. In the United States, lawmakers are considering several bills, which vary in their details, that would require social media companies to provide data to researchers.

But social media data is a rich source of information not only for researchers; law enforcement agencies are also often interested in obtaining social media data. In some cases, these demands are lawful and justified. However, law enforcement personnel have also used social media data in the past for illegitimate purposes such as monitoring protestors, dissidents, and members of religious or racial minorities.

Legal requirements that social media companies make data available to independent researchers should not inadvertently become tools for unjustified and increased law enforcement surveillance of social media users. This report examines existing legal protections for stored social media data in the US and EU and how they might be impacted by researcher access to social media data.

In the US, there is a significant risk that disclosure of social media data to independent researchers may make it easier, as a matter of law, for law enforcement personnel to access that data and surveil social media users. In particular, the uncertainty of how courts may apply the Fourth Amendment's third-party doctrine and gaps in the Stored Communications Act may allow law enforcement agencies to have more access to data that is disclosed to independent researchers.

We also find that legal protections that allow journalists and others who disseminate information to the public to resist law enforcement demands for information may not always apply to academic researchers or may provide only limited protections.

In contrast, in the EU disclosure of social media data to independent researchers likely will not impact the legal requirements governing the disclosure of data to law enforcement found in the General Data Protection Regulation, European Convention on Human Rights, EU Charter of Fundamental Rights, and Law Enforcement Directive and national implementing legislation. However, there remains a risk that, in practice, law enforcement personnel may find it easier to access social media data from researchers, regardless of legal protections.

In light of these increased risks, policymakers implementing or enacting requirements that social media companies share data with independent researchers in both the US and EU may want to consider the following mitigating measures:

1. Not allowing law enforcement agencies to qualify as vetted researchers in the legal regimes that establish access mechanisms.
2. Providing independent researchers with access to data through or at the social media company, rather than allowing the researcher to possess it.
3. Requiring researchers to destroy data after a certain time period or when their research has concluded.
4. Additional research to understand whether providing data to independent researchers will make law enforcement more aware of – and likely to demand access to – users' social media data.

The report recommends that US lawmakers, in particular, may want to consider:

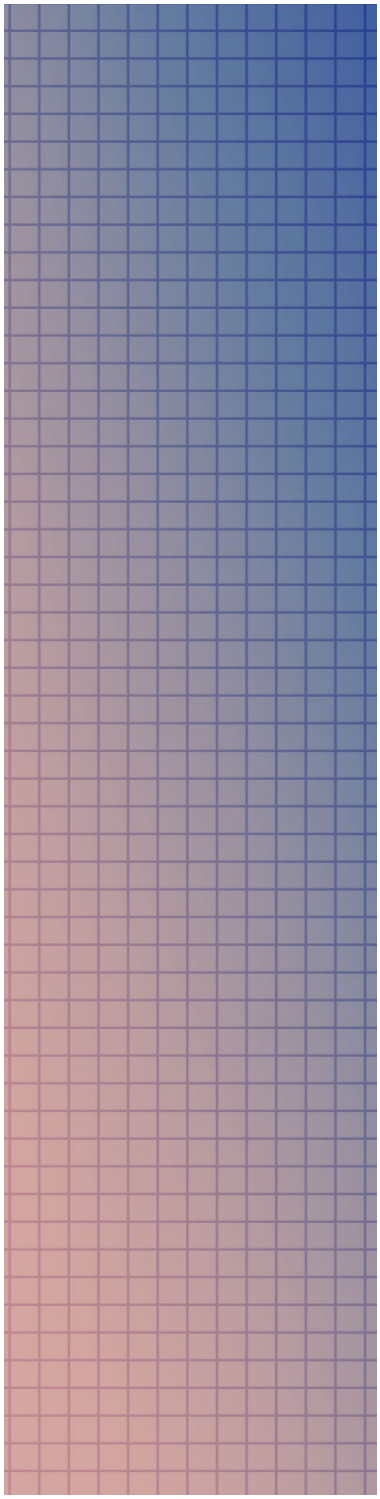
1. Restricting access to research tools that offer public data, such as APIs, to vetted researchers.
2. Requiring government entities to seek access to data subject to the Stored Communications Act only from providers of an electronic communications service or remote computing service subject to the Act's requirements, and not from a researcher when the researcher obtained such data under a researcher access to data law.

3. Requiring law enforcement agents to obtain a warrant, supported by probable cause, before they may access social media data obtained by researchers under a researcher access to data law, in addition to (2) (with respect to data not covered by the Stored Communications Act) or as an alternative to (2).
4. Prohibiting researchers from voluntarily disclosing data.
5. Enacting a federal shield law and expanding state shield laws to clearly cover researchers.

The report recommends that EU lawmakers, in particular, may want to consider:

1. Requiring data sharing agreements that prohibit researchers from sharing data with any other party unless legally obligated to do so.
2. Additional transparency obligations for social media platforms and independent researchers.

I. Introduction



Independent researchers' access to social media data is critical to understanding how social media impacts society. But social media data is a rich source of information not only for researchers; law enforcement agencies are often interested in obtaining information by and about social media users too.[1] In the past, some law enforcement agencies have sought to access tools that would allow them to monitor social media data that is collected in bulk, such as through Application Programming Interfaces (APIs), which are often used for data disclosures to researchers and others. [2]

In some cases, legal demands for social media data are lawful and justified. However, law enforcement personnel have also sought access to social media data in the past for illegitimate purposes, such

1 Transparency reports from Meta, Twitter, Google, and TikTok show growing numbers of requests or legal demands in recent years from governments around the world for user data. See also Adrian Shahbaz & Allie Funk, *Social Media Surveillance*, Freedom House (last visited Dec. 27, 2022); Rachel Levison-Waldman, *Directory of Police Department Social Media Policies*, Brennan Ctr. (last updated May 25, 2022).

2 For example, public records obtained by the American Civil Liberties Union of Northern California showed that social media intelligence platform company Geofeedia – which had access to user data from Facebook, Instagram, and Twitter – marketed itself to US law enforcement. Matt Cagel, *Facebook, Instagram, and Twitter Provided Data Access for a Surveillance Product Marketed to Target Activists of Color,* ACLU NorCal (Oct. 11, 2016). Other records showed that US fusion centers were accessing data from social media monitoring company Dataminr. Nicole A. Ozer, *Twitter Cuts Off Fusion Spy Centers' Access to Social Media Surveillance Tool*, ACLU NorCal (Dec. 15, 2016). Across the Atlantic, UK police and intelligence agencies used access to Twitter data intended for private sector clients and advertising companies, leading Twitter to cut off the government's access to its API. Natasha Lomas, *UK government irate at Twitter's surveillance API crackdown*, Tech Crunch (Apr. 26, 2017). Other types of government entities have also sought access to data intended for researchers. For example, as EU lawmakers debated Digital Services Act Article 40, NATO lobbied for a change in the law that would allow its researchers to access social media data. *Disinfo Update 08/02/2022*, EU Disinfo Lab (last visited Dec. 27, 2022).

as monitoring protestors, dissidents, and members of religious or racial minorities.[3] Social media data can be particularly sensitive and private. It includes data such as the contents of users' communications, photos, and videos; information about relationships and connections between users; and information about users' finances, health, sexual orientation, political beliefs, and other private matters.

Lawmakers in both the European Union and the United States have enacted or are considering proposals that would require social media companies (and, in some cases, some other tech companies) to provide researchers with certain data. To protect individuals' privacy and limit the chance of government abuse, lawmakers may also want to think about what steps they can take to ensure that these laws and proposals do not inadvertently allow law enforcement personnel to access social media data that they otherwise could not obtain.

In the EU, Article 40 of the Digital Services Act will require providers of "very large online platforms" or of "very large online search engines" to provide vetted researchers with access to data, subject to certain conditions. The DSA entered into force in November 2022, but many details around independent researchers' access to data through DSA Article 40 will be decided in delegated acts that have yet to be adopted.[4] Among other things, the delegated acts will likely address how to implement DSA Article 40 consistently with the General Data Protection Regulation (GDPR). The European Digital Media Observatory on Platform-to-Researcher Data Access Working Group has published a detailed report and proposed Code

3 In one recent example, federal authorities in the US monitored social media "to gather intelligence about nationwide protests and possible violence" following the US Supreme Court's decision to overturn *Roe v. Wade*. Jack Gillum, *DHS Agents Monitored Twitter After Roe Decision, FOIA Shows*, Bloomberg (Oct. 26, 2022). In another example, a 2018 report by the American Civil Liberties Union of Massachusetts revealed that the Boston Police Department used access to Geofeedia "to monitor the entire Boston Muslim community" and "track online speech associated with large protests," including tracking the terms "protest," "#blacklivesmatter," and "Ferguson." Nasser Eledroos & Kade Crockford, *Social Media Monitoring In Boston: Free Speech In The Crosshairs*, ACLUm (2018). In the Netherlands, the National Coordinator for Counterterrorism and Security has gathered social media data from activists to conduct general threat assessments. Naomi Appelman, *The Dutch government wants to continue to spy on activists' social media*, Racism & Tech. Ctr. (May 11, 2022).

4 For example, it is not clear whether certain data may be categorically excluded from access by independent researchers, or what mechanisms very large online platforms may be expected or required to use to provide researchers with access to data. It is also not clear whether researchers outside the EU will be permitted to apply for the status of "vetted researchers" or whether data created by or about users outside the EU may be provided to researchers.

of Conduct with recommendations about how platforms can share data with independent researchers in compliance with the GDPR.[5] Many expect it will form the basis of at least some of the delegated acts implementing DSA Article 40.

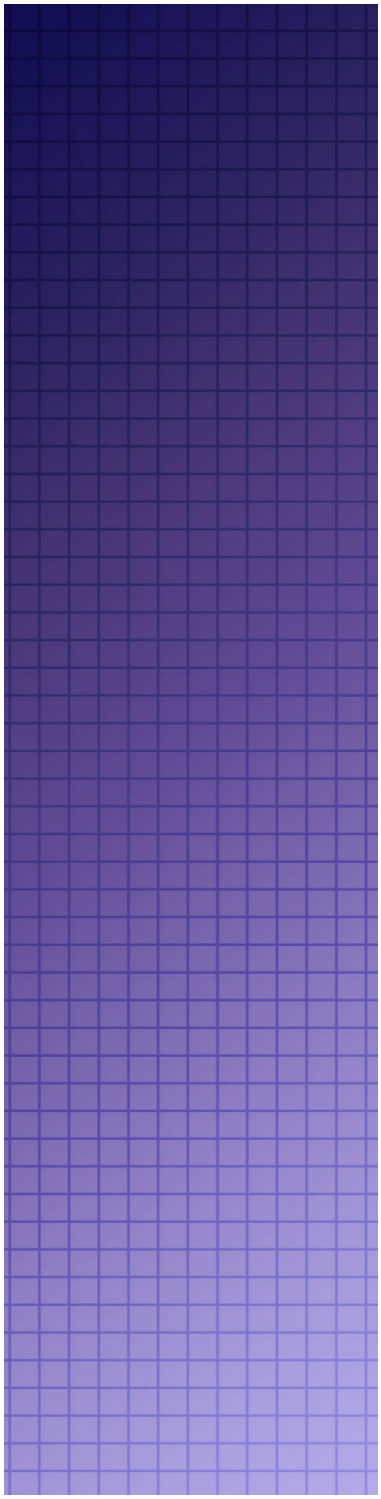
Legislative efforts in the US to provide independent researchers with greater access to social media data are not as far along. US lawmakers are considering a variety of bills that would require social media companies to provide data to researchers, with every proposal varying on the details. Prominent examples include the Platform Accountability and Transparency Act (PATA), the Digital Services Oversight and Safety Act (DSOSA), the Kids Online Safety Act (KOSA), and the Social Media Disclosure and Transparency of Advertisements (DATA) Act.[6]

This report examines the existing limitations under the law on law enforcement access to stored social media data in the US and EU and how laws that require social media companies to share data with independent researchers might impact them. Where there are significant risks of increased law enforcement surveillance created by researcher access to social media data, we recommend mitigating measures that policymakers implementing or enacting such requirements may want to consider. If done correctly, independent researcher access to social media data has the potential to enhance transparency and public understanding of social media without allowing an unjustified increase in law enforcement access to social media data.

5 *See* Report of the European Digital Media Observatory's Working Group on Platform-to-Researcher Data Access, European Digital Media Observatory (May 31, 2022) (hereinafter "EDMO Code").

6 For a more in-depth summary of DSA Article 40 and proposals in the US, *see* Caitlin Vogus, *Improving Researcher Access to Digital Data: A Workshop Report*, Ctr. for Democracy & Tech. (Aug. 2022) at Appendix (hereinafter "*Workshop Report*").

II. Scope of Report



In this report, we identify and discuss laws in the US and EU that control whether and how law enforcement personnel may access stored social media data and how the application of these laws may change when social media companies provide data to independent researchers under a legal requirement.

Legal mandates requiring social media companies to provide data to researchers raise a range of practical and legal concerns, especially relating to the protection of user privacy. Risks to user privacy and safety from the disclosure of social media data are real. The question of whether researchers should have access to data, what data they should have access to, and what kind of privacy and cybersecurity protections should be in place continue to be debated. There are also significant questions about how to vet researchers to weed out bad actors who may seek to take advantage of data access for commercial purposes or other inappropriate uses, such as influencing elections or gathering information to discredit political or intellectual opponents.[7] Past CDT reports have set forth recommendations intended to help companies and policymakers navigate some of these concerns.[8]

This report builds on that work by considering only the specific issue of whether laws providing social

7 For example, the consulting firm Cambridge Analytica used the private data of millions of Facebook users without their consent to provide psychological profiles of potential voters to political campaigns, including former-President Trump's 2016 campaign, intended to aid in campaign outreach and advertising. The firm purchased the data from a researcher affiliated with Cambridge University who claimed he was collecting it for academic purposes. Scott Detrow, *What Did Cambridge Analytica Do During The 2016 Election?*, NPR (Mar. 20, 2018).

8 *See* Gabriel Nicholas & Dhanaraj Thakur, *Learning to Share: Lessons on Data Sharing from Beyond Social Media*, Ctr. for Democracy & Tech. (Sept. 2022); *Workshop Report*, *supra* note 6.

media data to independent researchers make it easier for law enforcement entities in the US and EU to access that data. Because we conclude that such laws do increase the risk of unjustified law enforcement access to social media data, this report explains what steps US lawmakers might wish to consider to mitigate that risk if they decide to enact laws that require social media companies to provide data to independent researchers and what measures EU policymakers may want to consider for the delegated acts for DSA Article 40 or elsewhere in EU law.[9]

This report focuses on the disclosure of data by social media companies to independent researchers. By independent researchers, we mean researchers who are not affiliated with social media companies. This could include academic researchers, researchers from civil society, and journalists. By social media companies, we mean online intermediaries that host user-generated content primarily for the purpose of sharing that content with others. This would include, for example, social networking sites and applications as well as messaging services.[10]

The data that could be made available to independent researchers may be related to an identified or identifiable individual or it may be non-identifiable information. It may include content that is publicly available for anyone on the internet or data that is not publicly available. For example, a law requiring social media companies to provide data to independent researchers could require companies to permit researchers to access publicly posted Tweets, comments by users in a public or private Facebook group, or even (in theory)

9 Use and sharing of social media data for commercial purposes also raises other significant privacy risks. This report focuses on law enforcement use of social media data using laws, methods, or tools intended for independent researchers, and, as a result, it does not discuss commercial data uses. Lawmakers in the US and EU may wish to continue to consider how commercial practices that rely on selling or buying user data should be limited, including the sale of social media data to law enforcement. *See* Carey Shenkman et al., *Legal Loopholes and Data for Dollars: How Law Enforcement and Intelligence Agencies are Buying Your Data From Brokers*, Ctr. for Democracy & Tech. (Dec. 2021).

10 While DSA Article 40 applies beyond social media companies and bills in the US may also apply more broadly if adopted, we focus on social media companies here because both independent researchers and law enforcement personnel may have a particular interest in social media data and because this data can contain particularly sensitive information about users.

users' direct messages.[11] It may also require social media companies to disclose aggregate or de-identified information to researchers, such as information about the number of views a social media post receives.[12]

By "law enforcement" agencies, we mean government agencies responsible for investigating and charging individuals with criminal offenses. This report does not consider the use of social media data provided to independent researchers by other government entities such as immigration authorities and intelligence agencies, *i.e.*, agencies involved in the gathering of data for national security, military, or foreign policy objectives. Lawmakers may also want to consider whether and how laws constraining the collection of social media data by intelligence agencies and other government entities would be impacted by requiring social media companies to share such data with independent researchers.

This report considers three ways that law enforcement personnel may obtain access to research-related social media data held by social media platforms or independent researchers:

1. *Voluntary disclosure by platforms or researchers in response to informal requests;*
2. *Compelled disclosure by platforms or researchers in response to a legal order; and/or*
3. *Direct access to platform data by law enforcement through mechanisms designed for academic researchers.*

In this report, we consider the potential for law enforcement agencies to use tools made available to the general public, or to be openly registered as researchers for those access mechanisms, or to affiliate themselves with researchers as part of consortia with access to platform data. Covert use of access mechanisms designed for researchers – such as scenarios where law enforcement personnel impersonate researchers or researchers act as clandestine or unofficial agents of law enforcement – is a

11 Laws requiring social media companies to provide data to independent researchers may need to specify that companies are required to provide only data to which the company has access. *See*, *e.g.*, DSA Article 40(5)(a) (explaining that providers may ask that a data access request be amended on the grounds that "they do not have access to the data"). For example, an encrypted messaging service cannot provide the content of users' messages to researchers, because it does not have access to that data.

12 For a more detailed discussion of the types of social media data that researchers seek access to, *see* Workshop Report, *supra* note 6.

serious concern but is outside the scope of this report.[13] Of course, law enforcement officials may have legitimate reasons for seeking access to social media data to investigate crime, and there are lawful ways for law enforcement personnel to obtain that access in both the US and EU. However, in this report, we consider whether researcher access to data laws may allow law enforcement personnel to access social media data in an unjustified or abusive manner.

Additionally, we consider four possible models for how independent researchers will be given access to social media data in law.[14]

1. The social media company is required to make data publicly available to anyone.

Under this approach, the law would require a social media company to make certain data publicly available for anyone to access, including researchers. A social media company may use an API to make data available to anyone, or simply publish certain datasets or information. This is the approach taken, for example, in the EU's Code of Practice on Disinformation for access to "non-personal data and anonymised, aggregated or manifestly-made public data" that is relevant to researching online disinformation.[15] In the US, both PATA and DSOSA, for example, would also potentially require social media companies to disclose certain data or information publicly.

2. The social media company makes data available to specific independent researchers, and the researchers possess the data.

This approach would require social media companies to make certain data available only to researchers who meet specific criteria. The researcher would be given a copy of the data to use

13 While not addressed in this report, lawmakers may also want to consider the threat of covert or unofficial government access to social media data shared with independent researchers. For example, in 2021, the Hungarian government transferred control of public universities in the country to "quasi-public foundations led by close allies of the country's prime minister, Viktor Orban." Benjamin Novak, *Hungary Transfers 11 Universities to Foundations Led by Orban Allies*, N.Y. Times (Apr. 27, 2021). This raised concerns about academic freedom and government control over institutions of higher education. In addition, academics have been among the targets of the Pegasus spyware. Wagdy Sawahel, *Academics are among the alleged targets of Pegasus spy software*, Univ. World News (Aug. 9, 2021).

14 Additional details on how DSA Article 40 will operate and how bills in the United States would function are available in the section, *Comparing Independent Researcher Access to Data Mandates*.

15 *The Strengthened Code of Practice on Disinformation 2022*, European Comm'n, at Measure 26.1 (last visited Dec. 27, 2022).

for their research, and the data would be held and controlled by the researcher. This approach is exemplified in some existing voluntary initiatives by social media companies to share data with researchers. For example, the Twitter academic API allows researchers to convert results into a CSV format and import them into a compatible database, like Excel.[16]

3. The social media company makes data available to a third party to administer independent researchers' access to the data.

Here, the law would require social media companies to make certain data available to a third party, like a university or research consortium. The third party would determine which researchers may access or use the data. It could give approved researchers a copy of the data, or allow them to access it through the university or research consortium. Social Science One is a prominent example of this method of providing researchers with access to data, in a voluntary context.

4. The social media company makes data available to specific independent researchers, but the company holds the data and makes it available for access in a company-controlled environment.

Under this approach, the law would require social media platforms to make certain data available only to researchers who meet specific criteria. The data would be held by the social media company, and the company would give the researcher access to the data for research purposes, but the researcher would not be permitted to copy the data or possess it.

The use of virtual or physical clean rooms is one way to achieve this method of access. A virtual clean room is a digital environment that would "permit researchers to import their own data, perform research analyses, and export the results of their analyses," while preventing them from exporting the social media data itself.[17] A physical clean room is a space to which physical access is restricted and "[d]ata analysis takes place on designated machines that are secured by encryption and disconnected from the internet."[18] Virtual or physical clean rooms would allow a researcher

16 *Tools and guides to support your work*, Twitter (last visited Dec. 27, 2022) ("Learn about five methodsways [sic] you can use to turn a JSON payload into a CSV; CSVs are a great way to review individual fields in Twitter data.")

17 EDMO Code, *supra* note 5, at Part II, para. 6.3.1.

18 *Id*. at Part II, para. 6.3.2.

to access, inspect, and analyze data, but not possess it. If this method of access is used, it is important that the social media company's ability to control how researchers use and analyze the data within the clean room environment is limited to only what is necessary to ensure the security of the data. This limit is necessary to ensure the independence of research.

This report also considers four potential options with respect to notice to or consent from users for the sharing of their social media data with independent researchers, and the privacy implications of each. Lawmakers could:

1. *Not require social media companies to give any notice to users that their data will be shared with researchers.*
2. *Require social media companies to give notice to users, but not require them to offer users the option to opt-in or opt-out of sharing data with researchers.*
3. *Require social media companies to allow users to opt-out from sharing data with researchers.*
4. *Require social media companies to affirmatively require users to opt-in to share data with researchers.*

////

Comparing Independent Researcher Access to Data Mandates

DSA Article 40 is the first major legislation requiring some online services to make data available to independent researchers. In the US, members of Congress have recently proposed at least five bills with provisions about researcher access to data held by online services: The Platform Accountability and Transparency Act (PATA), Social Media DATA Act, Digital Services Oversight and Safety Act (DSOSA), and Kids Online Safety Act (KOSA). The details of DSA Article 40 and the US bills have differences and similarities across five key metrics.[19]

1. What is the method for vetting independent researchers, and is data access limited to academics?

Each of these laws would require researchers to meet certain criteria to obtain data, but their vetting criteria and processes differ. Tiered access systems, in which vetted academic researchers receive greater access to data than other researchers or the public, are common. Under DSA Article 40 and in most US bills, both academics and researchers at nonprofit organizations may access some or all of the data the law would require social media companies to provide. US bills often delegate the vetting of researchers to a government agency, such as the National Science Foundation or Federal Trade Commission. Under DSA Article 40, the Digital Services Coordinator of Establishment will decide whether to grant a researcher the status of "vetted researcher." However, the DSA's delegated acts must also consider the creation of "independent advisory mechanisms" to facilitate researcher access to data, which could include a third party vetting body.

2. Are there limits on the types of data that would be accessible to "researchers," specifically, or the types of research that may be conducted?

Yes, DSA Article 40 and all of the US bills limit what kind of data will be made available to researchers or the type of research that can be conducted. These limits may be narrow or broad. For example, the Social Media DATA Act would require the disclosure of advertising data

19 For a more detailed comparison, *see* Caitlin Vogus, *Independent Researcher Access to Social Media Data: Comparing Legislative Proposals*, Ctr. for Democracy & Tech. (Apr. 21, 2022).

only. In contrast, PATA and DSOSA would require the disclosure of data necessary to study social media platforms. DSA Article 40 requires disclosure of data for the sole purpose of research that contributes to the "detection, identification and understanding of systemic risks" in the EU or assessments of risk mitigation measures. DSA Article 40 and many of the US bills do not exclude any specific categories of user data from disclosure (such as direct messages). Categorical exclusions of data could be enacted in future delegated acts or implementing regulations.

3. Which online services must make data available?

DSA Article 40 and most of the US bills attempt to target large companies by requiring that a platform reach a certain size before it must disclose data to independent researchers. In addition, DSA Article 40 and the US bills are not limited to social media platforms. DSA Article 40, for example, specifically includes "very large online search engines," and US bills often define the platforms they would cover to include services that host user-generated content more generally.

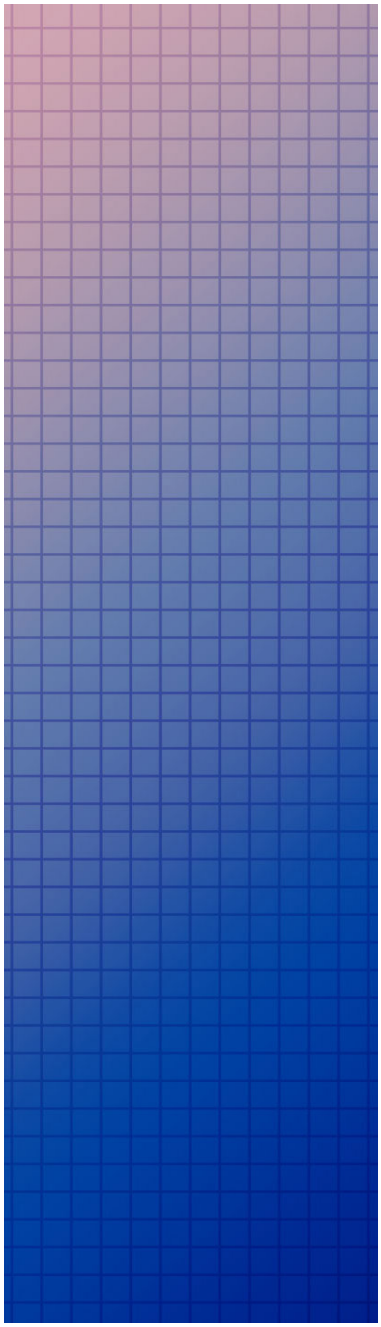
4. What privacy and security safeguards would there be for data made available to researchers?

DSA Article 40 and most US bills contain some safeguards for privacy and cybersecurity, but it is common to leave the details for future delegated acts or regulations. For example, the DSA's delegated acts must consider "the technical conditions under which data may be shared with researchers" as well as "the protection of confidential information, in particular trade secrets, and maintaining the security of [the online] service." In the US, DSOSA, for example, would require the FTC to issue regulations regarding privacy-protecting techniques, what information security standards should be in place, and other privacy and security measures.

5. Is there a safe harbor for independent methods of data access?

Most of the US bills contain a safe harbor from civil or criminal liability for independent researchers who face legal claims arising from their use of platform data. The details of the safe harbor, such as when a researcher may invoke it and the claims to which it applies, differ in each proposal. In contrast, the DSA does not contain a safe harbor for independent researchers.

III. US Law Enforcement Access to Social Media Data Shared with Researchers



Although social media companies are not currently required to share data with independent researchers under US law, lawmakers are actively considering imposing such requirements.[20] This section examines the implications of a potential mandatory researcher access to social media data law in the US on three key areas of law governing compelled disclosures of stored social media data to law enforcement: the Fourth Amendment of the US Constitution, the Stored Communications Act, and constitutional and statutory protections from compelled disclosures to law enforcement for those who disseminate information to the public. State constitutions and statutes may provide additional protections against compelled disclosures of stored social media data to law enforcement.[21] However, except for state shield laws, this report focuses on federal law. State law should be examined in more detail in future research.

////

20 *See Comparing Independent Researcher Access to Data Mandates.*

21 *See, e.g.*, Cal. Const. Art. I, § 13; California Electronic Communications Privacy Act (requiring state law enforcement to obtain a warrant to access almost all electronic communication information); Stephen E. Henderson, *Beyond the (Current) Fourth Amendment: Protecting Third-Party Information, Third Parties, and the Rest of Us Too*, 34 Pepp. L. Rev. 975, 989 (2007) (listing states that reject the federal third-party doctrine).

A. The Fourth Amendment

Law enforcement personnel in the US typically use two methods in the United States to compel access evidence: a subpoena or a warrant.[22] (They may also use other surveillance authorization methods established by the Stored Communications Act, like a Section 2703(d) order,[23] which are discussed in more detail in Section III.B.) A warrant authorizes investigators to physically search a place for evidence and take the evidence they find, while a subpoena directs a person in possession of evidence to produce it to investigators.[24]

Law enforcement officials must meet a higher legal standard to obtain a warrant than to obtain a subpoena. A warrant requires a court order. To obtain a warrant, the government must specify "the place to be searched and the person or thing to be seized."[25] In addition, the warrant must be supported by "probable cause," *i.e.*, a "fair probability that contraband or evidence of a crime will be found in a particular place."[26] In contrast, a subpoena does not require a court order; it can be issued by a grand jury or even a government agency. And a subpoena does not require the government to demonstrate probable cause; typically the requested data only needs to be relevant to an investigation, rather than likely include evidence of wrongdoing. A grand jury subpoena may require the recipient to turn over evidence as long as it is not "too sweeping in its terms 'to be regarded as reasonable.'"[27]

Because a higher legal standard applies to the issuance of a warrant than to a subpoena, social media users will generally have greater legal protection from law enforcement searches and seizures of their private data if the government is required to obtain a warrant than if the government is merely required to

22 Other types of surveillance orders include wiretap orders and various types of demands made pursuant to the Foreign Intelligence Surveillance Act.

Law enforcement may also ask researchers to voluntarily disclose data to them, and the Fourth Amendment does not prohibit law enforcement from making that request or researchers from complying with it. As discussed in Section III.B., the Stored Communications Act also does not prohibit researchers from voluntarily disclosing data to law enforcement personnel.

23 A Section 2703(d) order is sometimes described as a "mix between a subpoena and a search warrant." Orin S. Kerr, *A User's Guide to the Stored Communications Act, and a Legislator's Guide to Amending It*, 72 Geo. Wash. L. Rev. 1208, 1219 (2004) (hereinafter "*A User's Guide to the SCA*").

24 Orin Kerr, *Does Carpenter Revolutionize the Law of Subpoenas?*, Lawfare (June 26, 2018).

25 U.S. Const. Amend. IV.

26 *Illinois v. Gates*, 462 U.S. 213, 238 (1983).

27 *United States v. Dionisio*, 410 US 1, 11-12 (1973) (quoting *Hale v. Henkel*, 201 U.S. 43, 76 (1906)).

obtain a subpoena. As a result, it is important to understand when the government is required to obtain a warrant, and whether and how providing social media data to independent researchers could impact that requirement.

With some limited exceptions, the Fourth Amendment requires the government to obtain a warrant when it engages in a "search" or "seizure."[28] If there's no search or seizure, then no warrant is required.[29] A "'search' occurs when an expectation of privacy that society is prepared to consider reasonable is infringed."[30] To determine whether a search has occurred, courts typically apply a two-prong test: (1) whether there is a subjective reasonable expectation of privacy; and (2) if there is, whether that expectation is one that society would recognize as reasonable.[31] The first prong, often referred to as the "subjective prong," is generally less important, and courts often focus more on the second prong, often referred to as the "objective prong," asking whether the expectation of privacy is "objectively reasonable."[32]

Accordingly, whether the Fourth Amendment requires law enforcement to obtain a warrant to access social media data that is provided to independent researchers would turn on the two-prong test: Do social media users have a subjective *and* objective expectation of privacy in their data that is provided to researchers? The answer could differ depending on whether the data is publicly available to anyone on the internet or not. For data that is not publicly available, the answer could also depend on how courts apply the so-called "third-party doctrine" in this context.

1. The Fourth Amendment and publicly available data.

Some data that could be made available to researchers is also publicly available to anyone on the internet.[33] For example, Twitter's academic API gives researchers access to publicly posted Tweets. Some proposed legislation in the US would potentially require

28 *Katz v. United States*, 389 U.S. 347 (1967).

29 *See Illinois v. Andreas*, 463 U.S. 765, 771 (1983).

30 *United States v. Jacobsen*, 466 U.S. 109, 113 (1984).

31 *Katz*, 389 U.S. at 360 (Harlan, J. concurring).

32 *See* Orin S. Kerr, Katz *Has Only One Step: The Irrelevance of Subjective Expectations*, 82 U. Chi. L. Rev. 113, 113 (2015).

33 As CDT has previously explained, it can be difficult to precisely define when social media data is "public" or not. For purposes of this discussion, we are referring to data that is not restricted by a log in requirement or any other manner, and, as a result, is publicly available to anyone on the internet.

social media platforms to make certain public data available to researchers or members of the general public.[34]

There is no subjective or objective expectation of privacy in information that is voluntarily made public.[35] As a result, law enforcement traditionally is not required to obtain a warrant before they may access such information. This means that law enforcement would likely not be required to obtain a warrant to access social media posts that a user voluntarily makes public.[36] Indeed, it is well documented that government officials in the US monitor public social media posts by those suspected of crimes, potential terrorist threats, immigrants or visitors, and others.[37] Similarly, if law enforcement officials could access a tool, like an API, that provides publicly available data to researchers, they likely would not be required to obtain a warrant to use it.

However, the Supreme Court has suggested more recently that the collection of publicly available information may, in some circumstances, constitute a search. For example, in *United States v. Jones*, five Justices agreed that long-term GPS monitoring of a person's movements on public streets may violate an individual's reasonable expectation of privacy.[38] In *Carpenter v. United States*, the Supreme Court explained that "[a] person does not surrender all Fourth Amendment protection by venturing into the public sphere."[39] In that case, the Court held that requiring a telephone provider to disclose 7 days or more of a customer's cellphone records – allowing law enforcement to obtain "an all-encompassing

---

34 For example, PATA would require the FTC to consider whether to require certain social media platforms to disclose statistically representative samplings of public content and other information about high profile public accounts to the public.

35 *Katz*, 389 U.S. at 351 ("What a person knowingly exposes to the public, even in his own home or office, is not a subject of Fourth Amendment protection").

36 Rachel Levinson-Waldman, *Government Access to and Manipulation of Social Media: Legal and Policy Challenges*, 61 Howard L.J. 523, 533 (2018) (citing *People v. Harris*, 949 N.Y.S.2d 590, 595 (City Crim. Ct. 2012) appeal dismissed, 2013 WL 2097575 (N.Y. App. Term 2013) ("If you post a tweet, just like you scream it out the window, there is no reasonable expectation of privacy.")).

37 Rachel Levinson-Waldman et al., *Social Media Surveillance by the U.S. Government*, Brennan Ctr. (Jan. 7, 2022).

38 565 U.S. 400, 415 (2012) (Sotomayor, J., concurring); *id.* at 430 (Alito, J., concurring). However, the majority's decision in *Jones* that the warrantless GPS monitoring of a vehicle's movements on public streets was a search that violated the Fourth Amendment ultimately rested on its holding that the government's physical intrusion on private property is a search, and not on the individual's reasonable expectation of privacy. *Id.* at 404-05.

39 585 U.S. ___, 138 S. Ct. 2206, 2217 (2018).

record of the holder's whereabouts" – is a search under the Fourth Amendment. [40]

Under this more recent case law, it is possible that a court could conclude that law enforcement officials' monitoring of public social media content collected in bulk (such as through an API) is a search that would require a warrant under the Fourth Amendment if the monitoring could reveal sensitive information such as a comprehensive record of a user's location over a significant period of time or detailed information about their associations. On the other hand, some courts may rely heavily on the fact that public social media data is voluntarily disclosed to the public to conclude that a warrant is not required.

In sum, under current precedent, law enforcement officers are generally not required to obtain a warrant to monitor public social media data. Recent precedent, however, raises the possibility that this analysis could change if law enforcement personnel were to monitor public social media content collected in bulk in ways that could reveal sensitive information. This means that if social media companies make publicly available data more accessible to researchers through an API or other tool that is also open to other members of the public, courts may or may not create new precedent holding that the Fourth Amendment requires law enforcement officials to obtain a warrant to use that tool themselves.

2. The Fourth Amendment, non-publicly available data, and the third-party doctrine.

Lawmakers could also require social media companies to provide independent researchers with access to social media data that is "private," *i.e.*, not publicly available to anyone with an internet connection. The question of whether a "search" occurs if law enforcement were to compel access to such data from a researcher – and therefore whether a warrant is required to do so – is more complicated. Even if a user has a subjective expectation of privacy in such data (a determination that may vary depending on whether the user is aware of or consents to the sharing of data with the researcher), courts could conclude that a user lacks an

40 *Id.* at 2217 & 2217 n.3.

objective expectation of privacy in data shared with researchers under the **"third-party doctrine."**[41]

Whether courts will apply the third-party doctrine in this context will depend on many factors – such as the sensitivity of data, the scale and ease of government collection, and how voluntary the act of sharing was – that are being actively evaluated by courts in different settings. Because the third-party doctrine is currently in flux, it is difficult to say how courts would apply it to social media data disclosed to independent researchers. This means there is a risk that at least some courts could decide that law enforcement personnel are not required to obtain a warrant to access non-public social media data disclosed to researchers, depending on the exact circumstances of the disclosure.

The third-party doctrine provides that "a person has no legitimate expectation of privacy in information he voluntarily turns over to third parties."[42] As a result, "the Government is typically free to obtain such information from the recipient without triggering Fourth Amendment protections."[43] The third-party doctrine arose largely out of two cases decided by the Supreme Court in the 1970s.[44]

In *United States v. Miller*, the Court held the Fourth Amendment does not require the government to get a warrant to obtain the banking records (such as checks and deposit slips) from the defendant's bank accounts.[45] The Court concluded that these records were not private papers but rather "business records" of the bank.[46] The Court said that the defendant lacked a legitimate expectation of privacy in such records, which it said "contain only information voluntarily conveyed to the banks and exposed to

41 Another layer of complexity could arise in determining whether and how the Fourth Amendment applies to aggregate or de-identified data. For purposes of this discussion, we assume that the data at issue is personally identifiable. Depending on whether and how the Fourth Amendment applies generally to aggregate or deidentified data, this discussion of the third-party doctrine could also apply to such data, but we do not address that issue here.

42 *Smith v. Maryland*, 442 U.S. 735, 743–44 (1979).

43 *Carpenter*, 138 S. Ct. at 2216.

44 Another line of cases, often referred to as the "undercover informant cases," also informed the development of the third-party doctrine. In those cases, the Court held that a person does not have a reasonable expectation of privacy in information she voluntarily shares with a government agent, even unknowingly. *See, e.g., United States v. White*, 401 U.S. 745 (1971) (explaining that "[i]nescapably, one contemplating illegal activities must realize and risk that his companions may be reporting to the police.")

45 425 U.S. 435 (1976).

46 *Id.* at 440.

their employees in the ordinary course of business."[47] As the Court explained, "The depositor takes the risk, in revealing his affairs to another, that the information will be conveyed by that person to the Government."[48]

Three years later, in *Smith v. Maryland*, the Supreme Court held that the Fourth Amendment does not require a warrant before the government may install a pen register at a telephone company, which is a device that records the telephone numbers that a particular customer dials.[49] The Court rejected the defendant's claim that the pen register invaded his legitimate expectation of privacy, explaining that telephone customers know that the numbers they dial are conveyed to the telephone company and recorded "for a variety of legitimate business purposes."[50] Relying on *Miller*, the Court concluded that the defendant "voluntarily conveyed numerical information to the telephone company" by using his phone "and 'exposed' that information to its equipment in the ordinary course of business."[51] As a result, the Court held that the pen register was not a "search" that triggered the Fourth Amendment's warrant requirement.[52]

As many commentators have noted, the third-party doctrine has raised significant issues in the digital age.[53] Users routinely share highly sensitive and private information online, and, because of the nature of the internet, must share this data with third party intermediaries.[54] As a result, under the third-party doctrine, a user's disclosure of data to a social media company itself could preclude the user from asserting a reasonable expectation of privacy over that data, even before any data is shared with an independent researcher.[55] However, because some more recent court decisions have suggested that at least certain data may continue to be protected by the Fourth Amendment even when it is shared

47 *Id.* at 442.

48 *Id.* at 443.

49 442 U.S. at 745–46.

50 *Id.* at 743.

51 *Id.* at 743-44.

52 *Id.* at 745-46.

53 *See, e.g., A User's Guide to the SCA, supra* note 23; Tonja Jacobi & Dustin Stonecipher, *A Solution For The Third-Party Doctrine In A Time Of Data Sharing, Contact Tracing, And Mass Surveillance*, 97 Notre Dame L. Rev. 823 (2022); Brian Mund, *Social Media Searches And The Reasonable Expectation Of Privacy*, 19 Yale J.L. & Tech. 238 (2017).

54 *A User's Guide to the SCA, supra* note 23, at 1210–13.

55 Mund, *supra* note 53.

with online intermediaries,[56] it is important to consider whether courts might conclude that additional sharing of that data with a researcher abrogates Fourth Amendment protections if government officials were to seek that data from the researcher.

It is possible that at least some courts could hold that under the traditional application of the third-party doctrine as explained in *Miller* and *Smith*, disclosure of nonpublic social media data to independent researchers destroys the social media user's legitimate expectation of privacy in that data, meaning that the government could access that data from the researcher without a warrant. Especially if a user opts-in affirmatively to sharing their data with a researcher, a court may conclude that the user has voluntarily disclosed it to the researcher and therefore lacks an expectation of privacy in it. A court may decide that a user lacks an expectation of privacy in data she shares with a researcher even if she shares it with the understanding that it will only be used for research purposes.[57]

But even if a user is offered the option to opt-out of sharing data with a researcher and fails to do so, or simply knows that a social media company shares data with a researcher and continues to use that service, some courts could decide that the third-party doctrine as interpreted by *Miller* and *Smith* still applies. Just as the customer in *Smith* knew he was transmitting data to the telephone company and that the telephone company was recording it, courts may conclude that a social media user who fails to opt-out or who knows and passively acquiesces in data sharing has "voluntarily" shared their data with researchers as well.

In recent years, however, the third-party doctrine has been in flux, and changes to the doctrine may provide more legal protections to users whose data is shared with researchers. Courts are increasingly recognizing that certain disclosures to third parties do not destroy the reasonable expectation of privacy, and holding that the Fourth Amendment requires law enforcement officials to obtain a warrant in those cases. For example, the Massachusetts Supreme Court has held that the third-party doctrine does not per se preclude a reasonable expectation of privacy in data posted on a social media service, as long as a user takes protective measures that support a reasonable expectation of privacy, such as setting

56 These decisions are discussed in more detail later.

57 *See Miller*, 425 U.S. at 443 (holding that the third-party doctrine applies "even if the information is revealed on the assumption that it will be used only for a limited purpose").

and enforcing strict privacy settings on the user's account. [58]

Two leading cases, *United States v. Warshak* and *Carpenter v. United States*, have greatly informed the development of the third-party doctrine and called into question whether disclosure of data to third party intermediaries destroys the reasonable expectation of privacy.

In *United States v. Warshak*, the Sixth Circuit held that the defendant had both a subjective and objective expectation of privacy in the content of his emails, despite the fact that they were "disclosed" to a third party, his internet Service Provider.[59] As a result, the court concluded that the government violated the Fourth Amendment by compelling his ISP to disclose them without obtaining a warrant.[60] The court in *Warshak* distinguished *Miller* on two grounds: First, *Miller* involved "simple business records, as opposed to the potentially unlimited variety of 'confidential communications' at issue here."[61] And second, unlike the bank in *Miller*, the ISP was a mere "intermediary" for the defendant's emails, and not their intended recipient.[62]

In *Warshak*, the court analogized emails to other traditional forms of communication, such as phone calls and letters, where courts have recognized a reasonable expectation of privacy. It also rejected the government's argument that the defendant had no reasonable expectation of privacy in his email content because his ISP "contractually reserved the right to access [his] emails for certain purposes," concluding that "the mere ability of a third-party intermediary to access the contents of a communication cannot be sufficient to extinguish a reasonable expectation of privacy."[63] However, the court noted that "if the ISP expresses an intention to 'audit, inspect, and monitor' its subscriber's emails, that might be enough to render an expectation of privacy unreasonable."[64] Based on this analysis, the *Warshak* court held that the third-party doctrine did not allow the government to obtain the content of the defendant's emails without a warrant.

---

58 *Commonwealth v. Carrasquillo*, 179 N.E.3d 1104 (2022).

59 631 F.3d 266 (6th Cir. 2010).

60 The government had compelled their disclosure without a warrant under the Stored Communications Act.

61 *Warshak*, 631 F.3d at 288.

62 *Id.*

63 *Id.* at 286.

64 *Id.* at 287.

In *Carpenter v. United States*, the US Supreme Court cast further doubt on the third-party doctrine and its continuing application in the digital age. In that case, prosecutors obtained the cell phone records for petitioner Timothy Carpenter, a suspect in a series of robberies, without a warrant.[65] The government compelled Carpenter's wireless carriers to disclose cell site location information (CSLI) for Carpenter's cellphone during the four-month period when the string of robberies occurred.[66] Prosecutors used the records to show that Carpenter's phone was near four of the robbery locations when the robberies occurred, and Carpenter was convicted.[67] Carpenter appealed, arguing the Fourth Amendment required prosecutors to obtain a warrant to access these records. The Supreme Court agreed, holding that "an individual maintains a legitimate expectation of privacy in the record of his physical movements as captured through CSLI. The location information obtained from Carpenter's wireless carriers was the product of a search."[68]

The *Carpenter* Court "declined to extend" the third-party doctrine of *Smith* and *Miller* "to cover these novel circumstances."[69] According to the majority, the holdings of *Smith* and *Miller* depend at least in part on "'the nature of the particular documents sought.'"[70] The Court determined that cell-site records were a "qualitatively different category" of records from telephone numbers and bank records because they can reveal "a detailed and comprehensive record of the person's movements."[71] The Court also rejected the idea that CSLI was "voluntarily expos[ed]" to cellphone providers.[72] Cellphones are pervasive and indispensable in modern life, and CSLI is recorded "by dint of [a cell phone's] operation, without any affirmative act on the part of the user beyond powering up," the Court said.[73] "As a result, in no meaningful sense does the user voluntarily 'assume[ ] the risk' of turning over a comprehensive dossier of his physical movements."[74]

---

65 *Carpenter*, 138 S. Ct. at 2212. The government received court orders compelling disclosure of the cell phone records under the Stored Communications Act.

66 *Id.*

67 *Id.* at 2212–13.

68 *Id.* at 2217, 2221.

69 *Id.* at 2217.

70 *Id.* at 2219.

71 *Id.* at 2216–17.

72 *Id.* at 2220.

73 *Id.*

74 *Id*.

Courts are continuing to resolve whether and how the third-party doctrine applies to electronic records under current precedent. *Carpenter* moved the third-party doctrine away from a bright line application, in which any disclosure of any records to a third party destroys the reasonable expectation of privacy.[75] Instead, *Carpenter* and *Warshak* (which is controlling precedent only in the Sixth Circuit but may be persuasive elsewhere) offer other factors that courts may consider in applying the third-party doctrine and that may be relevant to determining whether social media data disclosed to researchers loses its Fourth Amendment protections.

First, both the Supreme Court in *Carpenter* and the Sixth Circuit in *Warshak* looked at the type of records sought in determining whether the third-party doctrine applies. In both cases, the courts decided that records that may be particularly revealing of personal information – the "sensitive information" contained in the contents of emails in *Warshak* and the "detailed and comprehensive records" of a person's movements revealed by CSLI in *Carpenter* – are entitled to greater Fourth Amendment protection than the "business records" at issue in *Miller*.

Second, both courts considered, in one way or another, the voluntariness with which a person shared their data with the third party. In *Warshak*, the Sixth Circuit emphasized that an ISP acts as a mere intermediary, and not the intended recipient, of the contents of emails. In *Carpenter*, the Supreme Court explicitly held that cellphone users do not voluntarily share their CSLI with providers since it is collected automatically when the phone is turned on.

Applying these two factors to social media data sheds light on how courts might apply the third-party doctrine to data shared with researchers.

Courts would first have to consider whether users maintain a reasonable expectation of privacy in the information they share with social media companies. As noted previously, it is not clear how courts will resolve this question, given the evolution of the third-party doctrine. However, at least some courts may conclude that, like the cellphones at issue in *Carpenter*, social media services are pervasive and indispensable, and so disclosure of data to them is not "voluntary." Some courts may also hold that at least some data held by social media companies is as sensitive and revealing as

75 Orin Kerr, *Understanding the Supreme Court's Carpenter Decision*, Lawfare (June 22, 2018).

CSLI.[76] As a result, they may hold that the third-party doctrine does not apply to at least some kinds of data shared with social media companies and that users can still have a reasonable expectation of privacy in it.

If a court concludes that users maintain a reasonable expectation of privacy in data disclosed to a social media company, it would then have to consider whether the third-party doctrine applies to the further disclosure of that data to an independent researcher. The resolution of this question is also uncertain. Under *Carpenter*, courts may again consider both the nature of the records sought and the voluntariness of the disclosure to the researcher.

The first factor considers the nature of the records sought. A law requiring social media companies to provide independent researchers with access to data could apply to a variety of data, and this data will differ in its level of sensitivity. If the court has already concluded that disclosure of the data to the social media company did not destroy the user's reasonable expectation of privacy, it likely would have already determined that the data at issue is sensitive.

As a result, the second factor, which considers whether a user voluntarily shared the data with a third party researcher, may be decisive.[77] As discussed previously, courts could conclude that a user who affirmatively opts in to sharing data with a researcher has voluntarily disclosed that data to the researcher and therefore has less or no reasonable expectation of privacy in it. At the other

76 The outcome of the court's analysis may vary depending on the type of data at issue. *Carpenter* and *Warshak* teach that the more sensitive the data, the more likely it is that the Fourth Amendment applies even if it is disclosed to a third party. Thus, some courts may treat the content of private social media messages or posts available to only a small number of recipients akin to email in *Warshak*, which would suggest that a warrant is required before law enforcement can obtain them, even though they are disclosed to a social media company. Similarly, some courts may consider individualized geolocation data as sensitive as the CSLI sought in *Carpenter* and also require a warrant for law enforcement access, even if the data is disclosed to a social media company. In contrast, some courts may decide that users do not have a reasonable expectation of privacy in the contents of social media posts that are available to a large number of people (though not the entire public). Similarly, some courts may decide that aggregate data or metadata about social media content – while still potentially quite revealing, especially if it can be reidentified – should not receive Fourth Amendment protection under the third-party doctrine, if it is disclosed to third parties. *See United States v. Forrester*, 512 F.3d 500, 511 (9th Cir. 2008) (holding that there is no reasonable expectation of privacy in email to/from addresses and IP addresses).

77 Unlike the ISP in *Warshak*, researchers would not be a mere "intermediary," since their access to social media data is not necessary to facilitate the communication of the data from the sender to the recipient.

extreme, a social media company could provide data to researchers without the consent of their users and with little or no notice. While such an approach may not be ethical, advisable, or legal,[78] it may lead a court to conclude that the user has not "voluntarily" disclosed the data to a researcher because the user takes no affirmative act to share the data beyond using the social media service.[79]

In between these two circumstances are situations in which a social media company gives users notice that their nonpublic data will be or may be shared with researchers or notice and the opportunity to opt out. Again, it is not clear how courts will treat the voluntariness of disclosure of data to researchers if the user is offered the opportunity to opt out. In *Carpenter*, the Court emphasized that it is not possible to use a cell phone without providing CSLI.[80] In contrast, users can use social media services even if they opt out of providing data to researchers. As a result, some courts may conclude that failure to opt out of sharing data with researchers is a voluntary disclosure that destroys a user's reasonable expectation of privacy in the data. In addition, courts may consider the nature of the notice and/or opt-out option. If the notice or opt-out option expresses a definitive intent to allow researchers to "audit, inspect, and monitor" specific social media data, "that might be enough to render an expectation of privacy unreasonable."[81]

In sum, if social media companies provide independent researchers with data that is not publicly available, whether the Fourth Amendment would require law enforcement officials to obtain a warrant to compel access to that data from researchers will turn on the courts' application of the third-party doctrine. The third-party doctrine is currently undergoing changes and reinterpretations that make it uncertain exactly how courts will apply it in different circumstances. However, it is possible that at least some courts could hold that law enforcement officials are not required to obtain a warrant to access social media data that has been disclosed to researchers.

////

78 For example, the California Consumer Privacy Act requires a business that controls the collection of a consumer's personal information to disclose "whether that information is sold or shared."

79 *Carpenter*, 138 S. Ct. at 2220.

80 *Id.*

81 *Warshak*, 631 F.3d at 287.

B. Stored Communications Act

Part of the Electronic Communications Privacy Act of 1986, the Stored Communications Act (SCA) is a federal law that governs the privacy of communications in electronic storage, such as emails or other internet communications. Congress enacted the SCA in response to the third-party doctrine and the concern that users may not maintain a "reasonable expectation of privacy" in internet communications that necessarily must be sent through third party intermediaries.[82]

As most relevant here, Section 2702 of the SCA restricts a provider of communication services from voluntarily disclosing the contents of certain communications and other information about their customers, subject to certain exceptions.[83] Section 2703 of the SCA establishes the legal process a governmental entity must use to compel a provider to disclose the content records or subscriber or customer records in different circumstances.[84]

Importantly, the SCA applies only to providers of "remote computing service" (RCS) or "electronic communications service" (ECS).[85] The law defines RCS as "the provision to the public of computer storage or processing services by means of an electronic communications system."[86] It defines ECS as "any service which provides to users thereof the ability to send or receive wire or electronic communications."[87] A single provider can be both an ECS provider and RCS provider, depending on the services they offer. Courts have held that social media

82 *A User's Guide to the SCA, supra* note 23, at Section I.

83 18 U.S.C. § 2702.

84 *Id.* § 2703.

85 As the Ninth Circuit has explained, the SCA "focused on two types of computer services that were prominent in the late 1980s: electronic communications services (e.g., the transfer of electronic messages, such as email, between computer users) and remote computing services (e.g., the provision of offsite computer storage or processing of data and files)." *Graf v. Zynga Game Network, Inc.* (*In re Zynga Privacy Litig.*), 750 F.3d 1098, 1103 (9th Cir. 2014).

86 18 U.S.C. § 2711(2). An "electronic communications system" is not the same as an electronic communications *service*, or ECS. The SCA defines an electronic communications system broadly as "any wire, radio, electromagnetic, photooptical or photoelectronic facilities for the transmission of wire or electronic communications, and any computer facilities or related electronic equipment for the electronic storage of such communications." *Id*. § 2510(14).

87 *Id.* § 2510(15).

services such as Facebook,[88] Myspace,[89] Twitter,[90] YouTube,[91] and WhatsApp[92] are ECS providers, RCS providers, or both, depending on the circumstances and what services they are providing.

In general, under Section 2702, an RCS or ECS cannot disclose the contents[93] of an electronic communication to any person or entity.[94] Section 2702 also prohibits an RCS or ECS from disclosing any other "record or other information pertaining to a subscriber to or customer of such service" to a governmental entity.[95] However, section 2702 provides certain exceptions to both of these prohibitions. The exceptions permit an ECS or RCS provider to voluntarily disclose the content of electronic records and subscriber or customer records with the consent of the customer

88 *Ehling v. Monmouth-Ocean Hosp. Serv. Corp.*, 961 F. Supp. 2d 659, 667 (D.N.J. 2013) (holding that Facebook is an ECS provider when it allows users to send private messages or make wall posts).

89 *Crispin v. Christian Audigier, Inc.*, 717 F. Supp. 2d 965, 981–82 (C.D. Cal. 2010) (holding that Facebook and Myspace are ECS providers with respect to wall posts or comments that are "restricted in some fashion" and for private messages); *id*. at 990 (holding, in the alternative, that Facebook and Myspace are RCS providers with respect to wall posts and comments).

90 *Shenwick v. Twitter, Inc.*, No. 16-cv-05314-JST (SK), 2018 U.S. Dist. LEXIS 22676, at *7 (N.D. Cal. Feb. 7, 2018) (holding that Twitter is an ECS with respect to direct messages).

91 *Viacom Int'l Inc. v. YouTube Inc.*, 253 F.R.D 256, 264–65 (S.D.N.Y. 2008) (holding that YouTube is an RCS provider when it allows users to upload videos and mark them as "private").

92 *In re United States of Am. for PRTT Order for One WhatsApp Account for Investigation of Violation of 21 U.S.C. § 841*, No. 18-pr-00017, 2018 U.S. Dist. LEXIS 43599, at *18 (D.D.C. Mar. 2, 2018) (holding that by "provid[ing] users with the ability to send and receive electronic communications to each other," . . . through the WhatsApp Messenger application, WhatsApp is providing an electronic communications service") (citation omitted).

93 The "contents" of a communication are defined as "any information concerning the substance, purport, or meaning of that communication." 18 U.S.C. § 2510(8).

94 *Id*. § 2702(a)(1), (2). "An ECS provider is prohibited from divulging only 'the contents of a communication while in electronic storage by that service.' 18 U.S.C. § 2702(a)(1). 'Electronic storage' is '(A) any temporary, intermediate storage of a wire or electronic communication incidental to the electronic transmission thereof; and (B) any storage of such communication by an electronic communication service for purposes of backup protection of such communication.' *Id*., § 2510(17). By contrast, an RCS provider may not divulge the content of any communication received by electronic transmission that is carried or maintained on its service for a customer or subscriber 'solely for the purpose of providing storage or computer processing services to [the] subscriber or customer, if the provider is not authorized to access the contents of [the] communications for purposes of providing . . . services other than storage or computer processing.' *Id*., § 2702(a)(2)." *Crispin*, 717 F. Supp. 2d at 973.

95 18 U.S.C. § 2702(a)(3). The U.S. Department of Justice "has taken the position that § 2702's prohibition on voluntary disclosures does not apply to sharing aggregate, de-identified non-content data with the government so long as it does not identify or otherwise provide information about any particular subscriber or customer." *Overview of Governmental Action Under the Stored Communications Act*, Cong. Research Serv. (Aug. 3, 2022).

or subscriber.[96] The exceptions also provide that both the contents of an electronic communication and subscriber or customer records may be disclosed if authorized by Section 2703.[97]

Section 2703 establishes how government entities can compel providers to disclose certain electronic records using a warrant, a subpoena, or a Section 2703(d) order.[98] The requirements differ depending on a variety of factors, including: (1) whether the government seeks to compel records from an ECS or an RCS; (2) how long the records have been in storage; (3) whether the government seeks the contents of electronic communications or subscriber or customer records; (4) whether the government will give notice to the user or not. In general, Section 2703 "requires a warrant for new ECS communications content (held for 180 days or less) and less robust protection for older content and non-content information."[99]

Because the SCA applies only to an ECS provider or RCS provider, its provisions governing the voluntary disclosure of records and the compelled disclosure of records to law enforcement would not apply to an independent researcher who gains access to those records.[100] This has two consequences for law enforcement access to such records.

First, the SCA would not forbid a researcher from voluntarily disclosing contents of communications or subscriber or customer information to governmental entities.[101] In other words, unlike an ECS or RCS, a researcher who wanted to share such data with law enforcement would be free to do so under the SCA.

---

96 18 U.S.C. §§ 2702(b)(3), 2702(c)(2).

97 *Id.* §§ 2702(b)(2), 2702(c)(1).

98 A Section 2703(d) order "is something like a mix between a subpoena and a search warrant. To obtain the order, the government must provide 'specific and articulable facts showing that there are reasonable grounds to believe' that the information to be compelled is 'relevant and material to an ongoing criminal investigation.'" *A User's Guide to the SCA, supra* note 23, at 1219.

99 *Overview of Governmental Action Under the Stored Communications Act*, *supra* note 95.

100 Note that, under the SCA as currently written, ECS and RCS providers cannot voluntarily disclose contents of an electronic communication to researchers, see 18 U.S.C. § 2702(a)(1), (a)(2), but they are permitted to voluntarily disclose noncontent records, *see* 18 U.S.C. § 2702(a)(3) (prohibiting voluntary disclosure of subscriber or customer information only to a "governmental entity").

101 *See Wesley College v. Pitts*, 974 F.Supp. 375, 389 (D. Del. 1997) ("[A] person who does not provide an electronic communication service [or a remote communication service] can disclose or use with impunity the contents of an electronic communication unlawfully obtained from electronic storage.") (citing 18 U.S.C. § 2702(a)).

Second, law enforcement agencies would not have to follow the requirements of Section 2703 to compel disclosure of the contents of an electronic communication or information about a subscriber or customer from a researcher. This means, for example, that law enforcement may not be required to seek a warrant to compel access to the contents of electronic communications that a researcher receives from an ECS or RCS,[102] even though the SCA would require them to seek a warrant to compel those same records from the ECS or RCS itself. Rather, law enforcement can use a subpoena to seek the contents of an electronic communication from a researcher; as explained previously, a subpoena provides less protection than a warrant because it does not have to be approved by a judge and can be issued based on mere relevance of the information sought to an investigation.

In sum, once a social media company discloses a user's data to a researcher, the data loses the protections against law enforcement access afforded by the SCA. The researcher could disclose the data to law enforcement voluntarily even though the social media company from which the data was obtained could not. If the researcher refuses a law enforcement request to voluntarily disclose the information, law enforcement officials could often compel its disclosure from the researcher with an easier-to-obtain legal process than they would have had to use to compel its disclosure from the social media company from which the researcher originally obtained the user's data.

////

C. Constitutional and Statutory Protections for those who disseminate information to the public

While the SCA does not apply to researchers, certain constitutional and other statutory provisions may protect researchers from law enforcement demands. In particular, the constitutional and statutory reporter's privilege and the Privacy Protection Act of 1980 (PPA) may shield independent researchers, at least to an extent, from law enforcement demands for data that researchers obtain from social media companies.

Researchers who receive law enforcement demands for online data may have a privilege against compelled disclosure of their records, known as the "reporter's privilege." The reporter's privilege is a legal protection that allows a journalist – or other people

102 As discussed in Section III.A., the Sixth Circuit has held that a warrant is required to access the contents of emails "that are stored with, or sent or received through, a commercial ISP." *Warshak*, 631 F.3d at 288. However, this holding is not binding outside the Sixth Circuit and may also be distinguishable when the contents of electronic communications are shared with third party researchers.

who disseminate information to the public – to refuse to provide information in response to a legal demand. The privilege arises out of the recognition that the free flow of information to the public would be impeded if those who disseminate that information could routinely be called to testify or provide information obtained in the course of newsgathering.[103] The application and scope of the reporter's privilege varies depending on whether it is based on the First Amendment or a state shield law.

The First Amendment reporter's privilege. As of 2021, every federal court of appeals in the US except for the Seventh and Eighth Circuits had recognized a reporter's privilege arising from the First Amendment.[104] Although different courts interpret the privilege differently, in general, the First Amendment reporter's privilege protects against the compelled disclosure of confidential, unpublished information.[105] The First Amendment reporter's privilege is qualified, meaning that it can be overcome – and the holder of the privilege can be compelled to turn over information – if certain conditions are met.[106]

Although the privilege is referred to as the "reporter's privilege," some courts have interpreted it to apply to more than just journalists employed by the traditional news media. These courts apply a "functional test," examining the actions taken by the person invoking the privilege to determine whether the privilege applies. For example, the Second Circuit has held that the First Amendment reporter's privilege could be claimed by anyone with "the intent to use material – sought, gathered or received – to disseminate information to the public."[107] Similarly, the Ninth Circuit has held that a book author could invoke the First Amendment reporter's privilege, as "the critical question for deciding whether a person

103 *See Von Bulow v. Von Bulow*, 811 F.2d 136, 142 (2d Cir. 1987); *United States v. Cuthbertson*, 630 F.2d 139, 147 (3d Cir. 1980).

104 *Introduction to the Reporter's Privilege Compendium*, Reporters Comm. for Freedom of the Press (last updated Nov. 5, 2021).

105 Lee Levine et al., *Newsgathering and the Law* at 18.07[1] (5th Ed. 2018). Some courts have also applied the qualified First Amendment reporter's privilege to nonconfidental information, though overcoming the privilege may be easier with respect to nonconfidential information than confidential information. *See Gonzales v. Nat'l Broad. Co.*, 186 F.3d 102, 104, 106 (2d Cir. 1998).

106 For example, in the Second Circuit, the First Amendment reporter's privilege can be overcome in a criminal case if the party seeking to compel subpoenaed documents makes "a clear and specific showing that the subpoenaed documents are 'highly material and relevant, necessary or critical to the maintenance of the claim, and no obtainable from other available sources." *In re Petroleum Products Antitrust Litigation*, 680 F.2d 5, 7–8 (2d Cir. 1982) (per curiam).

107 *Von Bulow*, 811 F.2d at 144–45.

may invoke the journalist's privilege is whether she is gathering news for dissemination to the public."[108]

At least one federal court of appeals, the First Circuit, has specifically recognized that "[a]cademicians engaged in pre-publication research should be accorded protection commensurate to that which the law provides for journalists" under the First Amendment.[109] In *Cusumano*, the First Circuit affirmed the denial of a motion by Microsoft to compel the confidential "notes, tape recordings and transcripts of interviews, and correspondence with interview subjects" of two professors who wrote a book about the "'browser war'" waged between Netscape and Microsoft."[110] As the Court explained, "scholars too are information gatherers and disseminators" entitled to First Amendment protections from compelled disclosure of their confidential materials.[111]

Despite the fact that academics and other researchers may be entitled to invoke the First Amendment reporter's privilege when they gather information for dissemination to the public, the privilege offers limited protection from compelled disclosure of information in criminal cases. The Supreme Court has held that journalists enjoy no First Amendment privilege to refuse to testify before a state or federal grand jury.[112] In addition, lower courts have more often recognized the First Amendment reporter's privilege in civil cases than criminal.[113]

State Shield Laws. In addition to the First Amendment reporter's privilege, 40 states and the District of Columbia have enacted statutory reporter's privileges, also known as "shield laws."[114] State constitutional provisions may also encompass a reporter's privilege. Shield laws "generally give greater protection to journalists than

108 *Shoen v. Shoen*, 5 F.3d 1289, 1293 (9th Cir. 1993).

109 *Cusumano v. Microsoft Corp.*, 162 F.3d 708, 714 (1st Cir. 1998).

110 *Id.* at 711.

111 *Id.* at 714.

112 *Branzburg v. Hayes*, 408 U.S. 665 (1972). Because the Court's decision in *Branzburg* was limited to the narrow question of whether journalists can be compelled to testify before a grand jury consistent with the First Amendment, lower courts have continued to interpret the First Amendment to provide a qualified reporter's privilege in other contexts.

113 *Introduction to the Reporter's Privilege Compendium*, *supra* note 104. The First, Second, Third, Ninth, Eleventh, and DC Circuits have recognized a reporter's privilege arising from the First Amendment in criminal cases. *Reporter's Privilege Compendium*, Reporters Comm. for Freedom to the Press at Section III.C.2 (last visited Dec. 27, 2022).

114 *Introduction to the Reporter's Privilege Compendium*, *supra* note 104.

the state or federal constitution" though their protections vary from state to state.[115] Some provide an absolute privilege, at least in certain circumstances; others provide a qualified privilege that can be overcome if certain statutory criteria are met.[116] Most apply to confidential information, while others also apply to nonconfidential information.[117]

State shield laws also vary with respect to who may invoke them. Some shield laws apply only to journalists who work for the traditional press.[118] Others are broader and may include academics, civil society organizations, and other researchers who disseminate information to the public.[119] For example, the Nebraska shield law protects any "medium of communication" which "shall include, *but not be limited to*, any newspaper, magazine, other periodical, book, pamphlet, news service, wire service, news or feature syndicate, broadcast station or network, or cable television system."[120] At least one state, Delaware, specifically includes "scholars" and "educators" within its shield law.[121]

The Privacy Protection Act. The Privacy Protection Act of 1980 (PPA) is a federal law that makes it unlawful for a government officer investigating a criminal offense to search for or seize any

115 *Id.*

116 *Id.* ("In 16 states and the District of Columbia, the privilege for confidential sources is absolute, meaning it cannot be overcome, despite the circumstances. . . . In the remaining states, the privilege is qualified.").

117 "Every state in the country recognizes legal protections for a journalist's confidential sources, except two . . . .Some state shield laws, like those in California, Illinois, and New York, also protect non-confidential information, though the protections are sometimes weaker." *Id.*

118 *See, e.g.*, Ala. Code § 12-21-142 (limiting protection of the Alabama shield law to "persons engaged in, connected with, or employed on any newspaper, radio broadcasting station or television station, while engaged in a newsgathering capacity."); Ariz. Rev. Stat. § 12-2237 (limiting protection of the Arizona shield law to "a person engaged in newspaper, radio, television or reportorial work, or connected with or employed by a newspaper or radio or television station . . . .")

119 *See, e.g., Quigley v. Rosenthal*, 43 F. Supp. 2d 1163 (D. Colo. 1999) (holding that the Anti-Defamation League may assert the privilege under the Colorado Shield Law because it gathers news and publishes periodicals, books and pamphlets), *aff'd in part and rev'd in part on other grounds*, 327 F.3d 1044 (10th Cir. 2003); *Cukier v. Am. Med. Ass'n*, 259 Ill. App. 3d 159, 630 N.E.2d 1198 (1994) (applying the Illinois Shield Law to a medical journal and its editor); *Louisiana v. Fontanille*, 1994 La. App. LEXIS 191, *7 (La. App. 5th Cir. 1994) (granting a qualified reporter's testimonial privilege to an investigative nonfiction book author).

120 Neb. Rev. Stat. § 20-145(2) (emphasis added).

121 Del. Code Ann. tit. 10 § 4320 (defining "reporter" as "any journalist, scholar, educator, polemicist, or other individual" who earns their principal livelihood by disseminating information to the general public or who obtained the information sought while serving as an agent, assistant, employee, or supervisor of a "reporter.")

work product[122] or documentary materials[123] from people engaged in the dissemination of information to the public.[124] In other words, the PPA prohibits law enforcement from using a search warrant to obtain these materials. The PPA can be enforced through a civil cause of action.[125]

The PPA was enacted in the wake of a Supreme Court decision upholding the search of the *Stanford Daily*'s newsroom by the police for pictures of a protest at the Stanford University Hospital that had turned violent.[126] The Senate Report accompanying the bill stated that it was enacted to afford "the press and certain other persons not suspected of committing a crime with protections not provided currently by the Fourth Amendment."[127] However, the PPA's protections are not limited to journalists. Rather, the PPA applies to work product or documentary materials held by any person who has the purpose "to disseminate to the public a newspaper, book, broadcast, or other similar form of public communication."[128] This includes academics, among others.[129] As a result, a researcher from academia or civil society who obtains social media data and plans to disseminate information to the public can invoke the PPA.

Importantly, however, the PPA does not bar law enforcement from obtaining work product or documentary materials entirely. Rather, the PPA simply "requires law enforcement agencies to rely on the cooperation of the media or subpoenas duces tecum[130] to obtain

122 "Work product material means materials (other than property used to commit a criminal offense) that are to be communicated to the public and contain the authors' impressions, opinions, conclusions, or theories." *Guest v. Leis*, 255 F.3d 325, 340-41 (6th Cir. 2001); *see also* 42 U.S.C. § 2000aa-7(b).

123 "Documentary materials" include "materials like notes, photographs, or tapes, other than things possessed for use in a criminal offense." *Guest v. Leis*, 255 F.3d 325, 340-41 (6th Cir. 2001); *see also* 42 U.S.C. § 2000aa-7(a).

124 42 U.S.C. § 2000aa(a), (b). The PPA provides for several exceptions to its "no search and seizure" provisions, such as when the person being searched is the suspect in a crime. *See* 42 U.S.C. § 2000aa(a)(1).

125 42 U.S.C. § 2000aa-6.

126 *See Zurcher v. Stanford Daily*, 436 U.S. 547 (1978).

127 S. Rep. No. 96-874, at 4 (1980), reprinted in 1980 U.S.C.C.A.N. 3950.

128 42 U.S.C. § 2000aa.

129 Jose M. Sariego, *The Privacy Protection Act of 1980: Curbing Unrestricted Third Party Searches in the Wake of* Zurcher v. Stanford Daily, 14 U. Mich. J.L. Reform 519, 535 (1981) (citing H.R. Rep. No. 96-1064, at 5 (1980)).

130 A subpoena duces tecum is a subpoena that requires a person to produce documents.

such documentary materials."[131] Despite the fact that a warrant (unlike a subpoena) must be supported by probable cause, the PPA assumes that a subpoena protects the interest of the person from whom documents are demanded more than a warrant for several reasons.[132]

The idea behind the PPA's subpoena requirement is that requiring law enforcement to use a subpoena gives the subpoenaed party the opportunity to object in court *before* turning over documents to law enforcement; in contrast, subjects of warrants have no opportunity to object before the warrant is executed.[133] The subject "may be able to quash the subpoena altogether, or at least modify it to release only a minimum of information."[134] This is especially important when coupled with the protections afforded by the First Amendment reporter's privilege or state shield laws. The PPA gives the subpoenaed party the chance to resist turning over documents that are privileged under the First Amendment or state shield law *before* their premises are searched and the documents are seized, rather than after the fact.[135]

In sum, reporter's privileges or shield laws may protect independent researchers who obtain social media data from compelled disclosure of that data to law enforcement. However, the protections of the First Amendment reporter's privilege and state reporter's privileges vary and often provide only a qualified privilege against disclosure. Independent researchers may also be able to invoke the PPA to prevent law enforcement from obtaining a warrant allowing seizure of social media data. However, under the PPA, law enforcement can still obtain a subpoena to compel researchers to provide this data.

131 *Citicasters v. McCaskill*, 89 F.3d 1350, 1353 (8th Cir. 1996); *see also Davis v. Gracey*, 111 F.3d 1472, 1481 (10th Cir. 1997) ("The PPA requires law enforcement officers, absent exigent circumstances . . . to rely on subpoenas to acquire materials intended for publication. . . .").

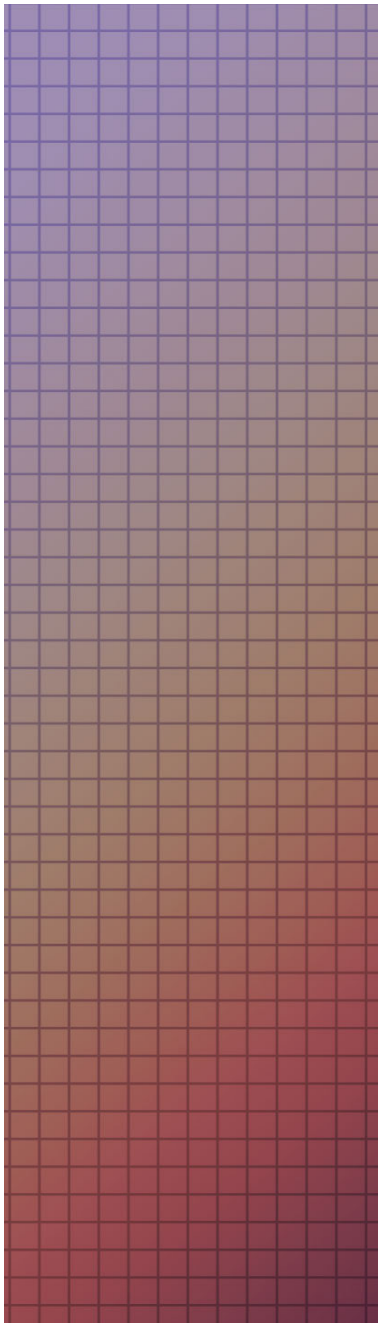
132 *See* Elizabeth B. Uzelac, *Reviving the Privacy Protection Act of 1980*, 107 Nw. U.L. Rev. 1437, 1460 (2013) ("Implicit in the Act's prohibition of searches and seizures is the assumption that a search – even with a warrant's probable cause requirement – was inferior to a subpoena duces tecum to protect the interests that the drafters of the statute had in mind.").

133 Sariego, *supra* note 129, at 530.

134 *Id.*

135 Other commonly cited benefits of the PPA's "no search" requirement include prohibiting law enforcement from physically searching the subpoenaed party's premises and "rummaging" through items, which may include confidential information that is not the subject of the law enforcement investigation, id. at 526–27, 529, and ensuring that law enforcement cannot prevent publication of information by removing the only copy of work product or documentary materials when executing a search warrant. *Id*. at 530.

IV. EU Law Enforcement Access to Social Media Data Shared with Researchers



In the EU, DSA Article 40 will establish a process under which researchers can apply for the status of "vetted researcher" and for access to specific data from a very large online platform or search engine.[136] To obtain the status of vetted researcher, DSA Article 40 requires that the researcher must be affiliated with a research organization, independent from commercial interests, conducting research in line with the permissible purposes under the DSA, and capable of meeting data security and confidentiality requirements; the researcher must also disclose their source of funding, demonstrate the necessity and proportionality of the data they seek, and commit to making their research results publicly available for free.[137]

If the researcher meets the criteria set forth in the DSA, the Digital Services Coordinator of Establishment[138] can order the very large online platform to provide the researcher with access to the data within a reasonable time period. Only certain research is permitted, however: The purpose of the research must be to contribute to the detection, identification, and understanding of systemic risks in the EU identified in the DSA and to assess the adequacy, efficiency, and impacts of the risk mitigation measures set forth in the DSA.

136 The DSA defines a very large online platform or search engine as those with 45 million or more average monthly active recipients of the service in the EU. DSA Article 33(1).

137 DSA Article 40(8).

138 The Digital Services Coordinator of Establishment is "the Digital Services Coordinator of the Member State where the main establishment of a provider of an intermediary service is located or its legal representative resides or is established. DSA Article 3(n).

Delegated acts will fill in further details about how DSA Article 40 will be implemented in practice. The delegated acts must address: the technical conditions under which data may be shared with researchers and the purposes for which the data may be used; how to share data with researchers consistent with the GDPR; the "relevant objective indicators, procedures and, where necessary, independent advisory mechanisms in support of sharing of data taking into account the rights and interests of the providers of very large online platforms or of very large online search engines and the recipients of the service concerned, including the protection of confidential information, in particular trade secrets, and maintaining the security of their service."[139]

This section first examines how the GDPR controls the voluntary sharing of social media data that independent researchers obtain under DSA Article 40 with law enforcement. It then considers the application of four sources of law on the compelled disclosure to law enforcement of social media data by independent researchers: the European Convention on Human Rights,[140] the EU Charter of Fundamental Rights, the Law Enforcement Directive (LED), and member state law, using the examples of France, Greece, and Ireland.

////

139 DSA Article 40(13).

140 The European Convention on Human Rights is not an EU instrument; it establishes the human rights that are protected in all countries that belong to the Council of Europe (COE). Because all members of the EU are also members of the COE, the European Convention on Human Rights places limits on the ability of EU law enforcement entities to compel access to social media data.

A. General Data Protection Regulation

The GDPR, Regulation (EU) 2016/679, governs the general processing[141] of personal data[142] by both platforms and independent researchers.[143] The GDPR provides specific exemptions and derogations for processing personal data for research purposes.[144] The GDPR also requires that the processing of personal data for research purposes must be subject to appropriate safeguards for the rights and freedoms of the data subject.[145] Under the GDPR, the data controller – or the individual or entity which "determines the purposes and means" of processing[146] – has the primary legal obligation to comply with the GDPR's rules.

The GDPR is most relevant to considerations of whether social media platforms or independent researchers may *voluntarily share* data with law enforcement.[147] Voluntarily sharing personal data with law enforcement agencies, whether done by a social media platform or an independent researcher, would be an act of processing that must comply with the GDPR.[148] In certain limited circumstances, the GDPR permits a social media platform or researcher to voluntarily share data with law enforcement agencies. To the extent that a social media platform or researcher voluntarily shares data with law

141 "Processing" data under the GDPR means "any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means." GDPR Article 4(2). This term is broad and applies to the sharing of personal data between online platforms, academic researchers, and law enforcement (including making data available for access), as well as any use of that data by law enforcement.

142 The GDPR defines personal data broadly, as "any information relating to an identified or identifiable natural person [a 'data subject']." GDPR Article 4(1). While the regulation does not precisely define what information can identify a natural person – and European data protection authorities have differed in how they interpret what information is identifying – in general, "personal data" is a broader category than "personally identifiable information" and will include most of data that might be made available by platforms to academic researchers.

143 For purposes of this discussion, we assume that the platforms and researchers are both subject to the GDPR. The GDPR applies to "a company or entity that processes personal data as part of the activities of one of its branches established in the EU, regardless of where the data is processed;" or "a company established outside the EU and is offering goods/services (paid or for free) or is monitoring the behaviour of individuals in the EU." *Who does the data protection law apply to?*, European Comm'n (last visited Dec. 30, 2022).

144 *See* GDPR Articles 5(1)(b), 5(1)(e), 9(2)(j), 14(5)(b), 17(3)(d), 21(6) and 89(2).

145 GDPR Article 89(1).

146 GDPR Article 4(7).

147 The GDPR permits data controllers to share personal data with law enforcement in response to a legal demand. *See* GDPR Article 6(1)(c) (providing that processing is permitted where it is necessary to comply with a legal obligation). Whether and how law enforcement can compel disclosure of data from social media platforms or researchers is governed primarily by the Law Enforcement Directive and member state law. *See supra* Section IV.D.

148 *See* GDPR Article 4(2) (defining "processing").

enforcement in a way that breaches the GDPR, the affected data subjects would be entitled to either (i) complain to a data protection supervisory authority (each member state has its own supervisory authority), and/or (ii) seek an 'effective judicial remedy' before the courts (e.g. a claim for damages (Articles 77-79 GDPR).

Importantly, the GDPR's requirements for voluntarily sharing data with law enforcement do not depend on whether the sharing is done by a social media platform or an independent researcher when both are subject to the GDPR. Whichever entity is the data controller is obligated to follow the GDPR's requirements. Both social media platforms and independent researchers can be considered data controllers.[149]

The GDPR requires that the processing of personal data be "lawful, transparent and fair."[150] That is, processing of personal data may only take place where at least one of the lawful bases set out in Article 6 GDPR applies, the controller must explain its processing to data subjects, and personal data must not be processed in ways that are unfair to data subjects. The GDPR also requires, among other things, that personal data not be processed for new purposes incompatible with those for which it was collected ('purpose limitation'). Chapter 3 of the GDPR creates a range of legal rights for data subjects in relation to the processing of their personal data, including the right to object to processing in certain circumstances.

GDPR Article 6: Lawful bases for voluntarily transferring data to law enforcement. The GDPR requires the data controller – whether it is a social media platform or an independent researcher – to establish that a lawful basis for sharing data with law enforcement applies. The GDPR sets out six lawful bases on which personal data may be processed, two of which could be the basis under which social media platforms or researchers could voluntarily share data with law enforcement agencies: Article 6(1)(d) and Article 6(1)(f).[151] Article 6(1)(d) permits processing "in order to protect the vital interests of the data subject or of another natural person." Article 6(1)(f) permits processing that is "necessary for the purposes of the legitimate interests pursued by the controller or by a third party."

149 *See* EDMO Code *supra* note 5, at Part I, para. 34–38.

150 GDPR Article 5(1)(a).

151 The GDPR also includes heightened protections for certain data designated as "special categories of personal data." GDPR Article 9. Social media platforms or researchers could not process special category data by voluntarily sharing it with law enforcement unless the processing meets one of the specific exemptions set out in GDPR Article 9 or member state law. In practice, this means it would be significantly more difficult – but not impossible – for platforms or researchers to voluntarily share special category data with law enforcement.

It is well established that the prevention, detection, and investigation of crimes is a legitimate interest for data controllers to pursue,[152] though this interest is not without limitation.[153] In most cases of voluntary sharing of data with law enforcement, social media platforms or researchers would likely rely on Article 6(1)(f) as the lawful basis. In addition to identifying a legitimate interest, Article 6(1)(f) requires that the social media platform or researcher establish that the processing is necessary for that interest, i.e. that the requested data is relevant to the investigation, and balance the legitimate interests against the rights and freedoms of the data subject(s). Thus, whether Article 6(1)(f) permits a social media platform or researcher to voluntarily share data with a law enforcement agency will vary depending on the circumstances and whether Article 6(1)(f)'s criteria for necessity and balancing of interests are satisfied.

GDPR Article 5: Purpose limitation. Article 5(1)(b) GDPR requires personal data to be "collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes." Controllers may have multiple purposes for collecting data, and the purposes of data collection should be communicated to data subjects at the point of data collection, such as through a privacy policy. If controllers do not disclose to data subjects that personal data may be voluntarily shared with law enforcement at the point of collection, this sharing will be considered processing for a new purpose.

Processing for a new purpose is lawful only if the new purpose is not 'incompatible' with the original purpose of data collection. GDPR Article 6(4) sets forth the criteria a controller must consider in determining whether a new purpose is compatible with the original

152 In the only extant EU-level guidance on legitimate interests, the Article 29 Working Party confirmed that "the notion of legitimate interest could include a broad range of interests, whether trivial or very compelling." *Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC*, Article 29 Working Party at p.24 (Apr. 9, 2014). The wording of Article 6(1)(f) makes it clear that the legitimate interest may be pursued by a 'third party' (in this case, the law enforcement agency requesting voluntary disclosure). The fact that GDPR Article 23(d) permits member states to derogate from parts of the GDPR in the interests of the prevention, detection and investigation of criminal offenses – whilst not directly applicable to Article 6(1)(f) – indicates the weight which is given to this interest in the GDPR regime. Finally, in the case *Ryneš* C-212/13 (2014), the European Court of Justice confirmed that the use of private CCTV cameras for the prevention of property crime was lawful processing on the basis of legitimate interests.

153 For example, a social media company that provides evidence of criminal activity to the law enforcement authorities for the prevention, detection and investigation of crimes would likely fall within the scope of GDPR Article 6(f). In contrast, actively processing data to detect any such activity, such as by using advanced data analytics, would likely not be covered by Article 6(1)(f).

purpose for which data was collected. In considering whether a new purpose is compatible, controllers must take into account:

> (a) any link between the purposes for which the personal data have been collected and the purposes of the intended further processing;
> (b) the context in which the personal data have been collected, in particular regarding the relationship between data subjects and the controller;
> (c) the nature of the personal data, in particular whether special categories of personal data are processed, pursuant to Article 9, or whether personal data related to criminal convictions and offences are processed, pursuant to Article 10;
> (d) the possible consequences of the intended further processing for data subjects;
> (e) the existence of appropriate safeguards, which may include encryption or pseudonymisation.[154]

Whether voluntary sharing of personal data with law enforcement breaches the principle of purpose limitation is therefore very fact specific. What is clear, however, is that there could be at least some circumstances in which the principle does not prevent voluntary sharing with law enforcement.

GDPR Article 21: The right to object. When the lawful basis for the processing of personal data relies on GDPR Article 6(1)(f), the GDPR also provides data subjects with a right to object to the processing of their personal data.[155] If a data subject objects to the processing of their personal data, "the controller shall no longer process the personal data unless the controller demonstrates compelling legitimate grounds for the processing which override the interests, rights and freedoms of the data subject or for the establishment, exercise or defence of legal claims."[156] The right to object applies prospectively only – it requires the controller to stop processing personal data, but it does not affect the lawfulness of processing that took place and cannot undo the transfer of data from a social media platform or researcher to law enforcement that has already occurred.

In sum, the GDPR's obligation to ensure the lawful processing of personal data applies to both social media platforms and

154 GDPR Article 6.

155 GDPR Article 21.

156 *Id.*

researchers when they act as data controllers, and both face the same general obligations under the GDPR to ensure that their voluntary sharing of personal data with law enforcement is lawful. The GDPR does not prohibit a social media company or independent researcher from voluntarily sharing data with law enforcement officials in all circumstances, nor does it permit them to do so in all circumstances. A social media company or independent researcher may have a lawful basis under the GDPR to share personal data with law enforcement when, for example, a social media user is thought to be at risk of self-harm,[157] but may not have a lawful basis to share highly-sensitive and intrusive personal data with law enforcement when, for example, the data would be used only to investigate a minor crime like jaywalking.[158]

////

B. The Convention on Human Rights

The European Convention on Human Rights (the Convention) establishes the human rights that are protected in all countries that belong to the Council of Europe (COE). COE member states must comply with the Convention, and claims that a state has violated Convention rights can be brought before the European Court of Human Rights (ECtHR).

Article 8 of the Convention provides a right to respect for private and family life, and, in particular, "correspondence." It states:

> 1. Everyone has the right to respect for his private and family life, his home and his correspondence.
>
> 2. There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.

Article 8 has been interpreted to constrain whether and how law enforcement agencies may process data. The ECtHR has held that the processing of personal data by a state body may – and usually will – constitute a prima facie interference with the Article 8 right.[159] The personal data need not be confidential or secret for

157 GDPR Article 6(1)(d).

158 GDPR Article 6(1)(f) (requiring a balancing of interests).

159 *See, e.g., Amann v. Switzerland* [GC], 2000, § 65.

its processing by law enforcement to interfere with Article 8.[160] The Article 8 right is not absolute, and interference with it through government processing of data may be permitted in certain limited circumstances.

To determine whether the Article 8 right has been interfered with by law enforcement processing, the ECtHR considers whether the individual has a reasonable expectation of protection for his private and family life in the given context. Again, the secrecy or confidentiality of data is not critical to the determination of whether an individual has a reasonable expectation of protection for his private and family life. For example, the ECtHR held in *Perry v the UK*, 2003 that an unusual use of CCTV by police breached the Article 8 right, as it went beyond what an individual would expect in a public place. Similarly, in *P.G. and J.H. v. the United Kingdom*, 2001, the ECtHR held that the processing by the state of even data that is fully in the public domain, where it is stored systematically, is an interference with the Article 8 right.

As a result, it seems unlikely that the prior sharing of personal data with independent researchers by itself would reduce an individual's reasonable expectation of protection for their private life (in contrast to the US system, *see supra* Section III.A). An individual who understands that a platform will share their data with a researcher for research purposes would not also expect law enforcement to have access to and make use of that same data. Because the two purposes are too disconnected and the consequences for the data subject are qualitatively different, the individual's reasonable expectation of protection for their private life vis-a-vis law enforcement would not be impacted by the sharing of their data with independent researchers. Even if a person explicitly consents to personal data being shared with researchers, this would be unlikely to undermine their expectation of respect for private life by the state. The ECtHR has found that individual actions can indicate that the expectation has been given up, but only in relatively extreme cases.[161]

Although law enforcement's processing of personal data can interfere with the Article 8 right, the right is not absolute. Interference is permissible if it is "in accordance with the law and is necessary in a democratic society in the interests of national

160 *See Benedik v. Slovenia*, 2018; *Catt v the UK*, 2019 (finding that the collection of publicly available information is processing interfering with the Article 8 right).

161 *See Axel Springer AG v. Germany* [GC], 2012 (finding that the individual had "sought the limelight" by giving a number of press interviews about their private life).

security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others." Thus, to be permissible under Article 8, law enforcement processing of personal data must (1) have a specific basis in clear and accessible national law, (2) be in pursuit of a legitimate aim, such as the prevention, detection or investigation of crime and (3) be "necessary in a democratic society." This third factor is a fact-specific inquiry that requires law enforcement processing not to be disproportionate or excessive relative to the law aims pursued.[162] Application of these three criteria, however, will not be impacted by the sharing of personal data with independent researchers. There is no ECtHR precedent, for example, suggesting that law enforcement processing of data would be any more proportionate because of prior sharing with independent researchers.

In sum, Article 8 of the Convention limits whether and how EU law enforcement agencies may process personal data. Whether the Article 8 right applies to a given act of processing depends on the individual's reasonable expectation of protection for his private and family life in the given context. Data need not be secret or confidential for an individual to have a reasonable expectation of protection of it. Sharing data with independent researchers is unlikely to lead to the conclusion that law enforcement may process that data without respecting the limits imposed by Article 8. However, the Article 8 right is qualified, meaning that it can be overcome – and law enforcement processing of personal data can be permitted – if law enforcement processing of personal data: (1) has a specific basis in clear and accessible national law, (2) is in pursuit of a legitimate aim, such as the prevention, detection or investigation of crime and (3) is "necessary in a democratic society." Again, the application of these three criteria is not likely to be impacted by the sharing of personal data with independent researchers.

////

162 *See Khelili v. Switzerland*, 2011 and *L.L. v. France*, 2006.

C. The EU Charter of Fundamental Rights

The EU Charter of Fundamental Rights (the Charter) enshrines the fundamental rights of people in the EU.[163] Article 7 of the Charter, which governs the right to respect for private and family life, has a similar scope and meaning to Article 8 of the Convention. Article 8 of the Charter establishes the right to the protection of personal data. Article 8 of the Charter states:

> 1. Everyone has the right to the protection of personal data concerning him or her.
>
> 2. Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified.
>
> 3. Compliance with these rules shall be subject to control by an independent authority.

The Charter is an important safeguard against indiscriminate law enforcement access to data and mirrors many of the protections covered under the Convention and the GDPR. As with Article 8 of the Convention, the right to protection of personal data under Article 8 of the Charter is not impacted by the source of the data, or whether a platform has previously shared it with researchers.

However, the Charter also adds additional dimensions of protection. For one thing, the Convention establishes the minimum threshold of protection; EU law may provide for more extensive protection. In addition, as long as the EU has not acceded to the Convention, the Convention does not constitute a legal instrument that has been formally incorporated into EU law.[164] In contrast, the Charter is binding on Member States when they are implementing or applying EU law and on the institutions and bodies of the EU.

As a result, the Charter will apply when Member States are implementing DSA Article 40, and when they are implementing or applying the Law Enforcement Directive (*see* Section IV.D) or the GDPR.[165] In effect, this means that these laws be seen "through the

163 The Charter became legally binding in 2009, and so predates the GDPR.

164 *EU Fundamental Rights Agency Handbook, 'Applying the Charter of Fundamental Rights of the European Union in law and policy making at national level,* 2020.

165 In what is a novelty in EU law, the DSA also obligates companies to respect the EU Charter directly. The risk assessments that VLOPs must carry out are in line with rights protected under the Charter, and DSA Article 14 provides that online platforms must respect the Charter in their terms and conditions.

lens of" or interpreted "so consistently as possible" with the rights in the Charter, including the right to the protection of personal data and the right to respect for private and family life.

Furthermore, because the Charter is directly binding on EU agencies and institutions, it places further restrictions on the ability of agencies such as Europol (law enforcement) or Frontex (migration) to request access to data as 'researchers' under DSA Article 40. It is less clear, however, whether a third party employed by an EU agency to carry out its research would be bound by the Charter directly in the same way as the agency. This will depend upon the contractual agreement with that third party.

Finally, the Charter provides for a qualitative assessment of judicial review. As was mentioned earlier, one of the conditions of access to data under the GDPR is access to a judicial remedy. GDPR Article 78(1) and (2) recognises the right of each person to an effective judicial remedy, in particular, where the supervisory authority fails to deal with his or her complaint. Recital 141 of the GDPR also refers to the "right to an effective judicial remedy in accordance with Article 47 of the Charter" in circumstances where that supervisory authority does not act where such action is necessary to protect the rights of the data subject. This could be particularly relevant in jurisdictions where the rule of law is under threat.[166] The Charter adds a layer of assurance beyond the GDPR by requiring that access to judicial review must be effective in practice in the jurisdiction concerned in order for the handing over of data to be considered lawful.

In sum, the Charter provides an additional – substantially overlapping – layer of protection of personal data to those established by the Convention rights and GDPR, and the Charter's right to protection of personal data in Article 8 is not lessened by the sharing of data with researchers. The Charter also provides additional protections for personal data, beyond those in the convention, and places additional restrictions on how law enforcement may process personal data provided to independent researchers, because it is directly binding on Member States when they are implementing or applying EU law and on the institutions and bodies of the EU.

////

166 *See* C-156/21 *Hungary vs. Council and EP*, as an example of the CJEU elucidating what Art. 47 of the Charter entails.

D. The Law Enforcement Directive and Member State Law

The Law Enforcement Directive (EU) 2016/680 (LED), governs the processing of personal data by competent authorities[167] in the EU for the purposes of the "prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security."[168] The LED requires that law enforcement processing be lawful – *i.e.*, "necessary" for the performance of a task carried out by the competent authority for law enforcement purposes, and based on EU or member state law[169] – and fair. It also requires member states to legislate for the principles of purpose limitation and data minimisation in the context of law enforcement processing.[170]

The LED sets out a framework for the processing of personal data and must be implemented through national legislation. The LED sets a "floor" for individual rights in the protection of personal data. However, the national law governing law enforcement processing of data will vary among member states, which can enact more extensive protections than what the LED requires. For purposes of this report, we examined implementing legislation in France,[171] Greece, and Ireland, to analyze differences and similarities across member states and whether sharing social media data of different kinds with independent researchers would impact law enforcement access to that data under national legislation. France, Greece,[172]

167 Competent authorities include law enforcement agencies. The LED defines "competent authorities" as "any public authority competent for the prevention, investigation, detection or prosecution of criminal offences [...] including the safeguarding against and the prevention of threats to public security" or "any other body or entity entrusted by Member State law to exercise public authority and public powers for the purposes of the prevention, investigation, detection or prosecution of criminal offences including the safeguarding against and the prevention of threats to public security."

168 LED para. 7.

169 *Id.* at Article 8.

170 *Id.* at Article 4.

171 *See* Title III of The French Data Protection Act, Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés (FR), Légifrance, accessed on 28 September.

172 *See* The Hellenic Data Protection Act, i.e. Law 4624/2019, The Hellenic Republic, Law 4624/2019: measures for implementing Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data, and transposition of Directive (EU) 2016/, EN Translation by the Hellenic Data Protection Authority, August 2019.

and Ireland[173] have all transposed the law enforcement directive in national legislation.

France. The French Code of Criminal Procedure[174] governs compelled disclosure of data for preliminary and judicial investigations by law enforcement authorities. The most common types of police investigations are preliminary investigations and investigations for flagrant offenses. They are conducted by the public prosecutor, who reports to the Ministry of Justice. At the end of the preliminary investigation, the prosecutor may close the case, propose alternatives to prosecution, open a judicial investigation by referring the case to an investigating judge, or summon the accused to appear before the criminal court. Judicial investigations are conducted by an investigating judge, who is independent and has sole control over the investigation. They are opened either at the request of the public prosecutor or at the initiative of a victim. Judicial investigations represent 5% of criminal cases.

Law enforcement access to stored data, including information from a computer system, is subject to certain safeguards and constitutional principles, which vary according to whether data is sought prior to a judicial investigation, during one, or following detection in the course of a crime's commission. These safeguards may include, for example, limits on the duration of the investigation, the types of crimes for which information may be demanded, or supervision of the investigation by a judge or magistrate.[175]

---

173 *See* the Data Protection Act 2018. The Data Protection Act 2018 serves to repeal the Data Protection Act, 1988, and the Data Protection (Amendment) Act, 2003, except for provisions relating to the processing of personal data for the purposes of national security, defence, and international relations of the State. The 2018 Act, together with the previous data protection legislation, is collectively known as the "Data Protection Acts 1988 – 2018."

174 Code de procédure pénale (FR), Légifrance, accessed on 28 September 2022.

175 *See* Article 99-3 and 99-4 of the Code of Criminal Procedure (governing judicial investigations of criminal cases and for certain offenses); Article 77-1-1 of the Code of Criminal Procedure (governing preliminary investigations); Articles 60-1 and 60-2 of the Code of Criminal Procedure (governing investigations following detection *in flagrante delicto*, *i.e.*, a crime that is in the process of being committed or which has just been committed).

In 2021, the French Constitutional Council deemed the provisions related to preliminary investigations unconstitutional. Conseil Constitutionnel. Decision no. 2021-952 QPC of 3 December 2021, EN translation by the Constitutional Council. The French government amended the provisions, in the new Law n° 2021-1729 of December 22, 2021. Law n. 2021-1729 of 22 December 2021 for trust in the justice system. Loi n° 2021-1729 du 22 décembre 2021 pour la confiance dans l'institution judiciaire.

Importantly, however, it makes no difference to the legal analysis whether the data sought by law enforcement for judicial, preliminary, or *in flagrante delicto* investigations is held by a social media company or by a researcher. In addition, there are no authorities that suggest that a judge would consider the extent to which the personal data is 'private' or has been previously shared with other entities when deciding whether or not to order disclosure.

Greece. Compelled disclosure of personal data to law enforcement in Greece is governed by the Hellenic Constitution, Law 2225/1994 on Freedom of Correspondence and Communications,[176] Presidential decree 47/2005 on lifting of confidentiality of Communications,[177] and the provisions of the Code of Criminal Procedure.[178]

Article 19 of the Hellenic Constitution broadly protects the confidentiality of communications, including the exchange of ideas, news, and opinions that takes place within a context of intimacy and confidentiality.[179] This can include electronic communications. Article 3 of Greek Presidential Decree 47/2005 provides that electronic communication encompasses any type of communication that uses the internet as a channel, as well as the access to a website or to a database.[180]

---

176 The Hellenic Republic, Law 2225/1994 on the freedom of correspondence, communications and other provisions, as amended by Law 4947/2022.

177 The President of the Hellenic Republic, Presidential Decree 47/2005 on the Procedures and Technical and Organisational safeguards for lifting the confidentiality of communication and to safeguard it, 2005.

178 The Hellenic Republic, Criminal Procedure Code, as amended by Law 4620/2019.

179 The Hellenic Parliament, The Hellenic Constitution: Article 19, November 2019, and The Prosecutor of the Supreme Criminal Court of Greece, Opinion on the lifting of confidentiality in telecommunications and Clarifications on the lifting of the confidentiality of internet telecommunications, September 2009.

180 The highest administrative court in Greece – the Council of State – and the highest criminal court in Greece – Areios Pagos – have issued conflicting opinions on when Article 19 protects electronic communications and when electronic communications fall outside the scope of Article 19. The Council of State has held that "electronic communications" includes stored communications and metadata generated in connection with the communication itself. The Council of State, Decision 1593/2016, 2016. In contrast, Areios Pagos has held that Article 19 applies only while the communication is in transit. Once the recipient becomes aware of the content of the message, Areios Pagos has held that a stored communication falls within Articles 9 and 9A, which protect the inviolability of home and of private life and Protection of Personal Data, respectively, and not Article 19. While these conflicting decisions may complicate law enforcement investigations and individuals' assertion of their constitutional rights, whether law enforcement may access stored communications under either Article 19 or Article 9 and 9A does not depend on whether it has been shared with independent researchers.

The Greek Code of Criminal Procedure specifies the circumstances under which the confidentiality of communications may be lifted and law enforcement may gain access to them. Articles 4-5 of Law 2225/1994 cover requests from prosecution, investigative and pre-trial authorities to providers of electronic communications services in general.[181] The provisions distinguish between two different scenarios: a) lifting of communications' confidentiality for reasons of national security (article 3 of Law 2225/1994),[182] and b) lifting of communications' confidentiality for the detection and prosecution of a specific list of serious crimes (article 4 of Law 2225/1994).[183]

For investigations of serious crimes, the confidentiality of communications may be lifted only if the competent judicial council makes a reasoned finding that the investigation of the case or the establishment of the place of residence of the accused is "impossible or substantially difficult" without the lifting of the confidentiality of the communication. The Greek Code of Criminal Procedure establishes certain safeguards that must be in place when confidentiality is lifted, such as limits on whose communications may be targeted and approval or supervision of an order lifting confidentiality by a judicial authority.

As in France, however, the legal determination of whether confidentiality may be lifted would not vary depending on whether law enforcement seeks access to electronic communications from a social media company or an independent researcher, and it would not be impacted by the prior disclosure of electronic communications to an independent researcher.

Ireland. Irish law has no standard regime for search warrants or disclosure orders. Rather, warrants or disclosure orders may be issued under hundreds of different statutes. While the precise extent of law enforcement's power to compel disclosure of electronic data will be context-specific, they are generally very expansive. As in France and Greece, however, the type of entity – *i.e.* a platform or researcher – from whom the personal data is sought

181 The Prosecutor of the Supreme Criminal Court of Greece, Opinion on the lifting of confidentiality in telecommunications and Clarifications on the lifting of the confidentiality of internet telecommunications, September 2009.

182 According to the Hellenic Authority for Communication Security and Privacy, the vast majority of orders are granted for reasons of national security, rather than for serious crimes. The Hellenic Authority for Communication Security and Privacy, Annual Report 2020, p.57, 2021. Nevertheless, the focus of this report is on law enforcement criminal surveillance as opposed to national security or intelligence services' surveillance.

183 The Hellenic Republic, Law 2225/1994 on the freedom of correspondence, communications and other provisions, as amended by Law 4947/2022, Articles 3-4.

is not relevant to compelling access. Nor is it relevant whether the personal data has been previously shared.

Most search warrants in Ireland are issued by a District Court Judge or (less often) a peace commissioner. However, in a small number of cases, a warrant may be issued by a senior officer of the An Garda Síochána (the Irish police, or AGS). The search warrant provision with the widest application is section 10 of the Criminal Justice (Miscellaneous Provisions) Act 1997, as amended by the Criminal Justice Act 2006. It provides that the District Court may issue a search warrant where there are reasonable grounds for suspecting that evidence of or relating to an arrestable offense (an offense carrying at least 5 years imprisonment on conviction) is to be found at a place. This standard is relatively low, except that the suspected offense must have a degree of seriousness. The search must take place within one week of the warrant being issued.

Search warrants can also apply to electronically stored data. For example, section 7 of the Criminal Justice (Offences Relating to Information Systems) Act 2017 applies to certain cybercrime offenses. It explicitly permits police to require the disclosure – by a person present during the search – of a password/encryption key; access to information; or production of the information itself, for the purpose of both examination and seizure of information. Law enforcement agencies can also obtain search warrants to compel organizations to hand over data. The powers would generally include the ability to enter and search premises for relevant evidence and to seize and retain evidence and could also include the right to question persons on the premises and require their assistance, including requiring the decryption of communications data.

Disclosure orders allow law enforcement agencies to require persons to show them any material in their possession that is likely to be of "substantial value" in the context of certain criminal investigations or proceedings. This can include electronically stored information. Disclosure orders can require the subject of the order to give a password necessary to examine the information or to produce the information in a form in which it is, or can be made, legible and comprehensible. These requirements can be used to require the decryption of electronic data. Generally, disclosure orders are issued by a District Court Judge.

Finally, it is interesting to note that voluntary disclosures of stored communications from online service providers to law enforcement officials appear common in Ireland. According to a 2017 Report by Privacy International, law enforcement officials frequently request disclosure of stored communications on a voluntary basis under Irish data protection law, rather than pursuant to a search warrant, disclosure order or similar powers.[184]

In sum, the Law Enforcement Directive sets a baseline of minimum requirements for the processing of personal data by law enforcement agencies in Europe, which must then be implemented through national legislation. An examination of statutes governing the compelled disclosure of personal data to law enforcement in Greece, France, and Ireland reveals a variety of mechanisms for compelling such data, standards that law enforcement must meet in order to compel disclosure, and mechanisms for protecting privacy. However, common among all these regimes is that it is irrelevant to the legal analysis whether law enforcement seeks to compel access to personal data from a social media company or from an independent researcher, and whether the data has been previously shared with an independent researcher would not give law enforcement any additional legal grounds to gain access to the data under national law.

184 *The Right to Privacy in Ireland*, Privacy International & Digital Rights Ireland, at para. 45 (Sept. 2015).

V. Recommendations

Analysis of the laws governing voluntary or compelled disclosure of stored social media data to law enforcement in the US and EU reveals several ways in which law enforcement personnel could try to gain greater access to this data if social media companies are required to provide it to researchers.

In the US, the Fourth Amendment's third-party doctrine creates uncertainty about whether law enforcement personnel would be required to obtain a warrant to compel researchers to divulge social media data disclosed to them under such a law. For example, a court could conclude the third-party doctrine means that law enforcement officers are not required to obtain a warrant to access a user's direct messages once they are disclosed to an independent researcher. The limited application of the Stored Communications Act to providers of electronic communication services and remote computing services could allow for voluntary or compelled disclosure of social media data from researchers to law enforcement in a manner not currently possible. For example, an independent researcher who obtains users' non-public Facebook posts could voluntarily disclose those posts to law enforcement, even though the SCA would prohibit Facebook from doing so. And statutory protections for those who disseminate information to the public may need to be strengthened to ensure they protect researchers coming from academia and civil society, as well as journalists.

In the EU, legal protections for individuals' privacy appear to depend less significantly on whether data has been disclosed to independent researchers. The GDPR constrains the voluntary sharing of data with law enforcement by social media platforms and independent researchers alike, permitting it only in

certain circumstances that do not vary depending on whether data has been disclosed to researchers. The protections of Article 8 of the Convention on Human Rights, Articles 8 and 47 of the EU Charter also are not impacted by disclosure of data to independent researchers, nor are the requirements of the LED or member state law, at least in the three member states examined by this report. Nevertheless, in practice, the disclosure of data to independent researchers may still permit law enforcement to access more data than they can currently. For example, researchers may lack the knowledge or resources necessary to challenge unjustified law enforcement demands for user data.

Policymakers in the US and EU may wish to consider various steps to lessen the possibility of increased law enforcement access to social media data – particularly unjustified access – that could result from mandated researcher access to social media data.

The first set of recommendations is for policymakers in both the US and EU. These recommendations are aimed at ensuring that law enforcement personnel cannot directly access data intended for researchers by qualifying as "researchers" under the law and that researchers' access to data does not become a tool for unjustified law enforcement surveillance.

The second set of recommendations is for policymakers in the US, mainly to address the third-party doctrine and Stored Communications Act provisions that could provide law enforcement officials with greater access to data than currently permitted.

The third set of recommendations is for policymakers in the EU. These recommendations are intended to help ensure that researchers adequately adhere to GDPR requirements intended to restrict the voluntary sharing of data and to inform data subjects of when their data is shared so they can exercise their right to object.

////

A. For Policymakers in Both the US and EU

1. Consider disallowing law enforcement agencies from qualifying as vetted researchers in the legal regimes establishing access mechanisms.

In the EU, the GDPR sets out a special regime for processing for "scientific or historical research purposes"[185] with a range of qualified exemptions to GDPR obligations. DSA Article 40's regime for researcher access to platform data will compel certain platforms to provide data to "vetted researchers" and is intended to operate in tandem with GDPR rules on research processing. However, neither GDPR Article 89 nor DSA Article 40 contain provisions that would necessarily prevent a law enforcement agency from acting as a "vetted researcher."

Similarly, in the US, various bills would make certain social media data available only to vetted researchers. However, for the most part, these bills lack provisions that would exclude law enforcement agencies or personnel from being qualified as vetted researchers.[186] For example, DSOSA would require organizations to host researchers who may access social media data, and it would limit "host organizations" to 501(c)(3) organizations or institutions of higher education. Neither definition, however, categorically excludes law enforcement officers or those affiliated with law enforcement.

Law enforcement and other governmental agencies in both the US and EU have a demonstrated interest in gathering data from social media, and, in some instances, this data gathering may arguably have similarities to research.[187] Therefore, it is possible that law enforcement agencies may wish to gain direct access to the mechanism established in DSA Article 40 – and similar ones – intended for academic researchers.[188] Most bills in the US would similarly leave open the possibility that law enforcement agencies could qualify as "vetted researchers." This is particularly the case where law enforcement agencies might form part of research consortia, or where a body has a dual purpose or status, such as a training academy for police officers that carries out both research and law enforcement functions.

185 GDPR Article 89.

186 One exception is PATA, which excludes from its definition of "qualified researcher" any researcher affiliated with a Federal, State, local or Tribal law enforcement or intelligence agency.

187 *See supra* notes 2–3.

188 As noted in Section II, surreptitious access by law enforcement via their impersonation of qualified researchers is a serious risk, but is outside the scope of this report. Policymakers and researchers may want to also evaluate the risks of, and potential safeguards against, covert law enforcement access to social media data through access mechanisms intended for researchers.

In response, policymakers in both the US and EU may wish to consider limiting access mechanisms only to "vetted researchers" and requiring that researchers seeking the status of "vetted researcher" may not be personnel at law enforcement agencies, organizations that are acting on behalf of law enforcement agencies (such as contractors or organizations funded by law enforcement), or organizations acting in a law enforcement capacity or with a purpose similar to law enforcement.[189] Policymakers could consider broadly defining law enforcement agencies to include any government agency authorized to engage in or supervise the prevention, detection, investigation, or prosecution of any violation of civil or criminal law.

In the EU, the delegated acts implementing DSA Article 40 could define vetted researchers to exclude personnel at law enforcement agencies. The EDMO Code outlines one possible definition of scientific research that would exclude access to platform data by law enforcement agencies.[190] In the US, policymakers could include language in mandated researcher access to data laws that prohibits personnel of law enforcement agencies or affiliated with law enforcement agencies from obtaining the status of "vetted researcher."

In addition, policymakers may also want to consider designing access mechanisms described in laws or regulations so that it is always clear which named individuals are permitted (legally and technically) to access the social media data at issue. This would reduce the risk of data 'leaking' from one research organization to connected organizations or individuals, who might be involved in law enforcement. Again, the EDMO Code sets out an approach of this kind, which requires a research plan that identifies the principal investigator and "each known additional individual researcher (or required qualifications where recruitment is ongoing) that will access the dataset."[191]

---

189 Although outside the scope of this report, policymakers may also wish to consider excluding intelligence agencies, military agencies, and national security agencies from the category of "vetted researcher."

190 EDMO Code, *supra* note 5, at Preamble, para. 13. The EDMO Code would require the parties involved in voluntary sharing of platform data to ensure that the proposed activity is "qualifying research" under the Code. This case-by-case determination includes consideration of, among other things, the entities that will carry out the research. The EDMO Code provides that the entity must not carry out any of the functions of law enforcement, intelligence services, or defense or upholding of national security.

191 *Id.* at Part II.

2. Consider providing independent researchers with access to data through or at the social media company, rather than allowing the researcher to possess it.

Law enforcement agencies in both the US and EU may be more likely to demand social media data from independent researchers than social media companies if they believe that it is easier to obtain data from researchers than companies. As a practical matter, researchers may be less able to resist unjustified law enforcement demands for user data than social media companies. Individual researchers may not know that they can bring legal challenges to attempts to compel access to their data, and they may not have the financial resources necessary to do so. In addition, some researchers may have close relationships with law enforcement agencies and wish to actively cooperate with them. In short, researchers' willingness and ability to resist unjustified law enforcement demands for user data may vary greatly depending on the researcher and the circumstances.

In addition, while the law enforcement agencies in the EU have no greater legal basis to demand data from independent researchers compared to social media companies,[192] the same may not be true in the US. Because Fourth Amendment and statutory protections may be less when the data is held by a researcher than when the data is held by a social media company in the US,[193] US law enforcement agencies may seek to compel data from independent researchers instead of companies.

However, a researcher can provide social media data in response to a law enforcement demand only if they actually have the data to turn over. Law enforcement personnel would not be able to access data from researchers if it is technically impossible for researchers to comply with any order for compelled disclosure, even if they wanted to.

Lawmakers in the United States and EU may therefore want to consider facilitating independent researchers' *use* of platform data for research, while not allowing it to be transferred to – or stored by – the researchers. This would mean that researchers simply will not have data to turn over in response to a law enforcement demand. A researcher could access data through a virtual or physical clean room at a social media company, for example, in which they are allowed to inspect and analyze data but are not

192 *See supra* Sections IV.B–D.

193 *See supra* Sections III.A and B.

allowed to take it with them. Policymakers in the US and EU could consider requiring that all data – or just certain sensitive data that may be of greatest risk of unjustified law enforcement demands for access – will be accessible to researchers only in a clean room or another environment that will allow them to use the data but not possess it.[194]

Some researchers may oppose restrictions on their ability to possess data necessary to conduct their research. Researchers have expressed concerns about giving social media companies too much control over the types of data and methods of providing researchers with access to data, because they believe it undermines the independence and accuracy of their research.[195] In addition, allowing researchers to access but not possess data would prevent them from sharing data with other researchers who want to attempt to reproduce their studies and replicate their results, which is often a high priority for researchers.[196] Policymakers may want to consider other ways of addressing those concerns, such as through independent auditing of clean room data or requirements that social media platforms preserve data provided to one researcher and make it available to subsequent researchers seeking to replicate the research.

3. Consider requiring researchers to destroy data after a certain time period or when their research has concluded.

If policymakers in the US or EU decide to allow researchers to possess at least some social media data as part of a legal regime facilitating independent researchers' access to social media data, they may want to consider requiring researchers to destroy data periodically or when their research ends. As with the previous recommendation, this recommendation seeks to make it technically impossible for researchers to turn over data in response to an unjustified law enforcement demand. Requiring researchers to destroy data when it is no longer needed would ensure that researchers do not become a vast repository of social media data. This requirement would also bring a legal framework requiring disclosure of social media data to researchers into alignment with privacy best practices such as data minimization.

---

194 Indeed, the EDMO Code, *supra* note 5, proposes the use of digital clean rooms for high risk research processing.

195 *See* *Workshop Report*, *supra* note 6, at 15, 18–19.

196 *See* Luke Hutton & Tristan Henderson, *Making Social Media Research Reproducible*, Standards and Practices in Large-Scale Social Media Research: Papers from the 2015 ICWSM Workshop (last visited Dec. 30, 2022).

Again, however, some researchers may object to this requirement because it makes it more difficult to share data with other researchers to replicate results. Reproducibility of research may still be possible even if researchers must delete data if researchers maintain enough information about the type of data they requested, and the format or manner in which they requested it, to allow other researchers to make the same request to social media companies. Lawmakers may want to consider encouraging or requiring social media companies to make data available to researchers for replication purposes.

4. Consider further study of whether providing data to independent researchers will make law enforcement more aware of – and likely to demand access to – users' social media data.

Law enforcement agencies can make applications for compelled disclosure of data only if they are aware of the existence and availability of this data. Their applications are more likely to be successful the more they know about the data that might be available, as they will be better able to explain or document why the personal data could be relevant to an investigation.

Researcher access to social media data in either the US or EU will presumably result in more public information about what data social media companies hold and the insights that can be generated from it. This is to be expected when researchers carry out and publish research using the data that they access from social media companies – indeed, it is substantially the point of doing this research.[197] Greater law enforcement awareness of available data may also result from the laws that enable independent researchers to better access social media data, which may include requirements that social media companies publish codebooks or other descriptions of what data is available to researchers.[198] As a result, access to data by researchers may make law enforcement agencies more aware of what types of information they could find useful to

197 Lawmakers may wish to consider the risk that some independent researchers acting in bad faith may purposefully publish research that "tips off" law enforcement that they should pursue certain illegitimate investigations, even if the published research does not share the underlying data and researchers are restricted from voluntarily sharing the underlying data with law enforcement. Because this report focuses on the risk of law enforcement obtaining data from independent researchers, we do not address this risk in depth. However, lawmakers may want to carefully consider the criteria used to vet independent researchers and whether to limit the types of data available even to vetted independent researchers as potential ways to mitigate this risk.

198 *See*, *e.g.*, DSOSA at Sec. 10(d) (requiring submission of "data dictionaries" to the FTC); EDMO Code, *supra* note 5, at Part II, Section 2 (recommending publication of codebooks).

compel from social media companies and make them better able to formulate legal demands for data. Again, while there are legitimate reasons for law enforcement to make lawful demands to access social media data to investigate crime, unjustified and abusive law enforcement demands are also possible.

Policymakers may want to consider whether providing independent researchers with access to data could also make law enforcement more aware of and better able to access that data, and whether that result could have negative impacts on freedom of expression, assembly, religious exercise, and other rights. Mitigation of the risk of these potential negative impacts is likely to be difficult, as the publication of research findings is fundamental to the purpose of facilitating researcher access to social media data. However, other safeguards, discussed elsewhere in these recommendations, can help ensure that greater law enforcement awareness of available social media data does not turn into law enforcement abuse of this information.

Future research should examine whether law enforcement agencies are making more unjustified demands for social media data or are more successful in the demands they make following a legal mandate for social media companies to provide data to independent researchers. Such research could be conducted in EU countries following the effective date of DSA Article 40, for example.

////

B. For Policymakers in the US

1. Consider restricting access to research tools that offer public data, such as APIs, to vetted researchers.

As discussed previously, some social media platforms offer APIs or other tools that allow researchers access to publicly available content in an aggregate format. Some proposed legislation in the US would potentially require social media platforms to make data, including content, from high-profile public accounts or other public accounts available to the public using an API or other methods.[199]

APIs can be enormously powerful research tools because they allow their users to sort and filter publicly available content to drill down on specific types of information. For example, the Twitter API can give access to historical data or real-time Tweets. It can allow users to examine Tweets from a specific account, followers of a specific account, or "everyone who is talking about a certain topic,"

199 *See*, *e.g.*, PATA at Sec. 10.

including by allowing users to filter Tweets based on keywords.[200] It can use geo-filtering to identify Tweets from a certain geographic location.[201]

Just as researchers benefit from these tools, law enforcement could also find them extremely useful for monitoring online public conversations about current events or specific individuals, or even posts by specific individuals.[202] Law enforcement agencies already review public social media content manually or using specialized software, and while some of this monitoring could be for legitimate investigatory purposes, other monitoring could be for illegitimate ends.[203] Access to APIs or other databases of public social media content could increase unjustified law enforcement use of this data by making it even easier to compile, filter, and analyze. For this reason, some social media companies explicitly prohibit the use of their APIs for surveillance purposes.[204]

Because current precedent suggests that there is no reasonable expectation of privacy in social media content that users voluntarily post publicly,[205] the Fourth Amendment would likely not restrict law enforcement's ability to monitor that content through an API or other tool intended for researchers.[206] Lawmakers considering legislation that would require social media companies to offer APIs or other research tools for accessing publicly available information may want to consider including language in the statute that would ensure that the enhanced access to information that they would provide through the legislation benefited only researchers. One method would be to require that the researchers who can benefit

200 *See also* Listen for Important Events, Twitter (last visited Dec. 30, 2022) (explaining how the Twitter API can be used to "listen for important events from public Tweets as those events unfold in real-time."); *Explore a user's Tweets and mentions with the Twitter API v2*, Twitter (last visited Dec. 30, 2022) (explaining how the Twitter API can be used to "retrieve the public Tweets composed by, or mentioning a [specific] user").

201 *Building high-quality filters for getting Twitter data*, Twitter (last visited Dec. 30, 2022).

202 Access to many parts of the Twitter API is currently limited to those with developer or academic research accounts and is not provided to law enforcement.

203 *See supra* notes 2–3.

204 *See, e.g.*, Chris Moody, *Developer Policies to Protect People's Voices on Twitter*, Twitter (Nov. 22, 2016); *Meta Platform Terms*, Meta (last visited Dec. 30, 2022).

205 *See supra* Section III.A.1.

206 As discussed in Section III.A.1., some recent Supreme Court decisions have held that a warrant may be required, in some circumstances, to access information that an individual reveals publicly. However, because it is not clear how courts will apply these holdings to publicly posted social media content, lawmakers may want to clarify by statute the standards under which law enforcement officials may access and use tools that provide bulk access to such content.

from a right to access APIs and similar tools must meet certain criteria that would identify them as researchers, as opposed to making these tools widely available to the public generally,[207] including law enforcement.

In the alternative, if lawmakers wish to require social media companies to make such APIs or other tools available to the general public, they could provide that law enforcement personnel may access such tools only to investigate serious crimes such as terrorism, for example. Alternatively, lawmakers could consider at least prohibiting law enforcement agencies from using APIs or other required tools for certain purposes. Lawmakers could bar law enforcement from using them to monitor First Amendment-protected activity, such as the use of social media to organize a protest or engage in constitutionally-protected speech. They could also prohibit law enforcement from monitoring individuals using these tools based on a person's membership in a protected class, such as race, religion, or national origin.

In practice, however, such use restrictions are often difficult to enforce. Law enforcement officials may evade use restrictions by proposing dual purposes for investigations or simply concealing an investigation's purpose. If lawmakers do make APIs or other tools for accessing aggregate public social media data available to law enforcement officials with use restrictions, they may also want to consider requiring law enforcement agencies to report on their use of these tools to the public to ensure transparency and accountability. In addition, lawmakers may want to require independent audits of law enforcement use of these tools, such as through reports by the U.S. Government Accountability Office, to ensure that law enforcement personnel are abiding by the use restrictions.

2. Consider requiring government entities to seek access to data only from providers of an ECS or RCS directly under the SCA.

If lawmakers in the US allow researchers to possess at least some social media data as part of a researcher access to data law, they may also want to consider restricting law enforcement agencies' ability to compel disclosure of that data from researchers. The

207 This is the approach that EU lawmakers have adopted in DSA Article 40, which limits access to data, including through online databases or APIs, to vetted researchers. In addition, the GDPR would already make it unlawful for a platform to make its users' personal data generally available to anyone by way of a public tool for access. Furthermore, the Convention, LED, and national implementing legislation would place significant constraints on how law enforcement could use such tools, even if they were available.

SCA's critical limits on law enforcement demands for electronic communications data apply only when the data are held by an ECS or RCS. Because they would not apply to a researcher who holds such data, law enforcement could compel content and non-content data directly from a researcher without complying with the SCA's requirements.

Lawmakers could consider closing this loophole allowing an end-run around the SCA when drafting a statute that requires social media platforms to give independent researchers access to data. Lawmakers could include a provision in that law stating that if an ECS or RCS provides data covered by the SCA to a researcher, a governmental entity may not seek access to that data from the researcher. In effect, this would maintain the status quo under the SCA: the government entity would have to obtain that data from the ECS or RCS, subject to the SCA's restrictions, just as it would absent the researcher access mandated by the law. In addition, this requirement would ensure that independent researchers are not required to respond to law enforcement demands for data covered by the SCA. This is important because researchers may lack the knowledge and resources to respond to law enforcement demands with carefully limited disclosures (or, as discussed previously, to challenge law enforcement demands in their entirety).

3. Consider requiring law enforcement agents to get a warrant before they may access social media data obtained by researchers.

If lawmakers in the US allow researchers to possess at least some social media data as part of a researcher access to data law, they may want to consider requiring officials to obtain a warrant before they can compel a researcher to disclose it.[208] A warrant requirement would lessen the risk that law enforcement officials will gain access to social media data without adequate safeguards and oversight.

As discussed previously, the third-party doctrine makes it uncertain whether the Fourth Amendment's warrant requirement would apply

208 If lawmakers prohibit government entities from seeking access to data covered by the SCA from researchers, as discussed in Section V.B.2, they may still want to consider requiring law enforcement to obtain a warrant to compel disclosure of data that is not covered by the SCA from researchers. If lawmakers do not prohibit government entities from seeking access data covered by the SCA from researchers, they may want to consider requiring law enforcement to obtain a warrant to compel disclosure from a researcher any data she obtains through a researcher access to data law.

if law enforcement officials sought users' social media data from researchers.[209] Lawmakers may wish to consider removing this uncertainty in any law that would compel social media companies to provide independent researchers with access to data. Such a law could provide that, when researchers obtain social media data through the process or mechanism established by that law, law enforcement must obtain a warrant supported by probable cause to compel the researcher to disclose that data. This requirement would apply the Fourth Amendment's protections to data disclosed to researchers when sought by government officials from researchers, regardless of judicial interpretations of the third-party doctrine. Law enforcement officials who want to take advantage of any available lesser legal process that applies to certain data held by the social media company from which the researcher obtained the data can compel disclosure of that data using that lesser legal process directly from the social media company instead of from the researcher.

4. Consider prohibiting researchers from voluntarily disclosing data.

If lawmakers in the US allow researchers to possess at least some social media data as part of a researcher access to data law, they may also want to consider restricting a researcher's ability to voluntarily share data with law enforcement personnel and others. As discussed previously, the SCA's prohibition on voluntarily disclosing to law enforcement the contents of certain communications and other information about their customers applies only to an ECS or RCS. As a result, a researcher who possesses such data could voluntarily disclose it to law enforcement under current law.

Lawmakers may wish to consider prohibiting researchers who obtain social media data under the law from voluntarily disclosing the data to *any* third party, including law enforcement.[210] This would create an even stronger prohibition on the voluntary disclosure of data by researchers than that imposed by the SCA on an ECS or RCS provider, which allows an ECS or RCS provider to voluntarily disclose some types of data in some circumstances. An absolute prohibition on researchers voluntarily disclosing data to a third party would provide the strongest protection for the privacy of

209 *See supra* Section III.A.2.

210 Lawmakers could prohibit researchers from disclosing all of the data they receive, or they could limit it to certain categories of data, such as the contents of electronic communications or subscriber or customer records, as in the SCA.

users whose data is disclosed to researchers, while still maintaining existing mechanisms by which law enforcement can obtain such data directly from service providers.

Some researchers may object to this prohibition because it would prevent them from publicly publishing the data that underlies their research and from providing it to other researchers who may wish to replicate their work. As with other recommendations that would limit researchers' ability to possess or maintain data after their research is completed, policymakers could consider allowing researchers to maintain information about the type of data they requested, and the format or manner in which they requested it, and to share that information with other researchers, who can then request the same data from social media companies.

Alternatively, lawmakers may want to consider prohibiting researchers who obtain social media data under the law from voluntarily disclosing it to "governmental entities," as defined by the SCA. Consistent with the SCA, this would bar researchers from voluntarily disclosing the data they receive not only to law enforcement, but to every federal, state, and local government agency.

5. Consider a federal shield law and expansion of state shield laws to clearly cover researchers.

Lawmakers could also decrease the risk that law enforcement will seek to compel social media data from independent researchers by enacting a federal shield law that is broad enough to cover researchers. A federal shield law bill has been introduced several times in recent years in both the House and the Senate. The latest version, the Protect Reporters from Exploitative State Spying (PRESS) Act[211] would prohibit a federal entity from compelling disclosure of "protected information" from a "covered journalist" except in certain narrow, statutorily defined circumstances. The term "covered journalist" is defined broadly in a manner that should be interpreted to include researchers who intend to disseminate information on matters of public interest to the public.[212] In addition, "protected information" includes "any records, contents of a communication, documents, or information that a covered journalist

211 H.R. 4330, 117th Cong. (2022); S.2457, 117th Cong. (2022).

212 The bill defines the term "covered journalist" as "a person who gathers, prepares, collects, photographs, records, writes, edits, reports, or publishes news or information that concerns local, national, or international events or other matters of public interest for dissemination to the public."

obtained or created as part of engaging in journalism." This definition would likely include data researchers obtained from social media companies as part of their work to disseminate information on a matter of public interest to the public. Lawmakers concerned about law enforcement demands to compel researchers to disclose social media data may wish to consider enacting a federal shield law like the PRESS Act.

In addition, state lawmakers could consider expanding definitions in state shield laws to ensure they apply to researchers. They could do so by using a definition similar to that of "covered journalist" in the PRESS Act, by incorporating the functional definition of "reporter" that some courts have applied in the context of the First Amendment reporter's privilege, or by looking to other state shield laws with broader definitions. By expanding who may claim the privileges in a state shield law to include researchers, lawmakers could shield researchers from some or all law enforcement demands to compel social media data.

////

C. For Policymakers in the EU

1. Consider requiring data sharing agreements that prohibit researchers from sharing data with any other party unless legally obligated to do so.

The GDPR governs the voluntary sharing of personal data by both social media companies and independent researchers, and, as discussed previously, imposes the same requirements on both before they may voluntarily share data. Accordingly, providing social media data to independent researchers will not change the legal analysis of whether that data may be voluntarily shared with additional third parties.

However, as with compelled disclosure of data, there may be practical reasons that sharing social media data with independent researchers increases the risk that it will be voluntarily shared with law enforcement. Researchers may be more likely than social media platforms to interpret the GDPR to permit sharing or less aware of the restrictions that GDPR places on voluntary sharing of data. In addition, researchers may be more inclined to voluntarily share data when the GDPR permits it compared to social media companies, some of which may be more likely to insist on a legal demand for commercial reasons or because of concerns about negative public perception as a result of voluntary sharing.

In addition to considering technical measures that would make it impossible for researchers to voluntarily share data with law enforcement,[213] policymakers may also want to consider requiring that social media companies share personal data with researchers only pursuant to data sharing agreements, which would contain binding obligations on the researchers not to share the data they access with any other party unless legally obliged to do so. The GDPR already obligates any platform to have a data sharing agreement in place where it shares personal data with academics to facilitate their research.[214] Policymakers could consider reiterating this requirement in the delegated acts implementing DSA Article 40, to ensure it is compatible with the GDPR, and expand upon it to require that those data sharing agreements include a restriction on researchers voluntarily sharing data with third parties.

2. Consider additional transparency obligations.

The right to object under GDPR Article 21 may be an important safeguard against some types of law enforcement surveillance based on voluntary sharing with law enforcement. This right can only be exercised, however, when data subjects are *aware* of how their personal data is being used, making transparency from platforms and researchers vital. Consider a platform that has shared personal data with researchers, who go on to share it voluntarily with law enforcement. If data subjects are unaware of the first act of sharing with researchers, their prospects of exercising any data rights vis-à-vis the researchers are slim.

The GDPR provides for some exemptions to the principle of transparency where processing is carried out for scientific research purposes.[215] This presents a significant risk of the kind of transparency gaps described previously, in which data subjects are unaware of when and where their data has been shared. This problem is particularly acute because complete transparency is unlikely to be practical in the context of researchers accessing social media data, since companies will not know at the point of data collection which future research projects might rely on the data, or very large numbers of data subjects might be involved.

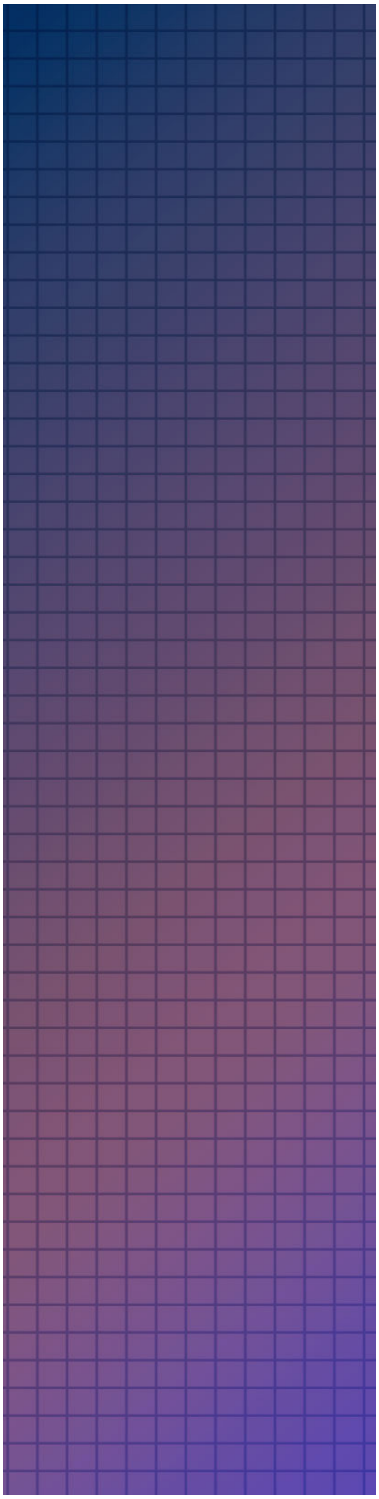
213 *See supra* Section V.A.2.

214 The EDMO Code, *supra* note 5, at Part I, section 2 and throughout, maps out an approach to researcher access to platform data consistent with the GDPR that puts a data sharing agreement at its heart.

215 GDPR Article 14. See the EDMO Code, *supra* note 5, at Part I, section 5 for a full analysis of these provisions.

Policymakers may wish to consider mitigating this risk by requiring social media companies and researchers to put in place pragmatic mechanisms for upholding the spirit of transparency. The EDMO Code proposes a range of options, including a system in which platforms provide an interface allowing a data subject to see at a glance which research projects (if any) their data has been shared with. In addition, researchers – if unable to provide transparency to data subjects directly – could make details of their personal data processing public on a dedicated website. Delegated acts implementing DSA Article 40 may wish to pay careful attention to how transparency and notification of data subjects that their data has been shared with researchers can be accomplished, to allow data subjects to exercise their right to object.

VI. Conclusion



Lawmakers have noticed both the importance of independent research in understanding how social media impacts society and the barriers that independent researchers face in accessing the data they need. The laws and proposals being enacted or considered in response that will require social media companies to provide vetted researchers with certain data would provide needed tech company transparency.

Yet, our review of legal protections for stored social media data in the US and EU suggests that these laws could unintentionally open the door to greater law enforcement surveillance of social media users, including in situations where such surveillance is not appropriate. Existing legal regimes governing law enforcement access to stored communications data often seek to strike a balance between protecting user privacy and permitting legitimate law enforcement access.

But these laws do not necessarily account for a world in which independent researchers may have access to vast amounts of highly sensitive user data. And disclosure of social media data to researchers creates practical risks of unjustified disclosures of data to law enforcement, because independent researchers may lack the resources and knowledge to resist abusive demands for user data.

Careful drafting of delegated acts implementing DSA Article 40 in the EU and of laws or regulations creating mandatory researcher access to social media data in the US can ensure that a proper balance is struck between providing researchers with access to social media data and preserving users' freedom from unwarranted law enforcement surveillance.

With the considerations identified in this report in mind, lawmakers can draft laws, regulations, and delegated acts requiring social media companies to provide independent researchers with access to the data they need to conduct research exploring how social media impacts us all, while maintaining strong legal protections for the privacy of stored social media data.

cdt.org

cdt.org/contact

Center for Democracy &
Technology
1401 K Street NW, Suite 200
Washington, D.C. 20005

202-637-9800

@CenDemTech

CENTER FOR
DEMOCRACY
& TECHNOLOGY

KONRAD
ADENAUER
STIFTUNG