HAS *KATZ* BECOME QUAINT? USE OF BIG DATA TO OUTFLANK THE FOURTH AMENDMENT

Jeffrey L. Vagle[*]

INTRODUCTION

On December 14, 2010, a federal court, upon a government motion, entered an order pursuant to the Stored Communications Act ("SCA") requiring Twitter to turn over to the government subscriber information concerning the accounts of three Twitter users. The order demanded only "non-content" data: names, addresses, and all records of user activity, including dates, times, and IP address data for all subscriber activity since November 1, 2009.

The subscribers filed a motion to vacate the order on grounds that it was insufficient under the SCA and violated both the First and Fourth Amendments. The motion was denied by the magistrate judge.[1] The subscribers then filed objections to the magistrate judge's ruling.[2] The district judge denied the subscribers' objections, agreeing with the magistrate judge that the subscribers lacked standing to challenge the SCA-based order on non-Constitutional grounds. The court also rejected the subscribers' Fourth Amendment challenge, stating that "any privacy concerns were the result of private action, not government action," and thus the "mere

[*] Mr. Vagle is an associate with Pepper Hamilton LLP. J.D., Temple University Beasley School of Law; B.A., Boston University.

[1] *In re § 2703(d) Order*, 787 F. Supp. 2d 430 (E.D. Va. 2011). In her decision, the magistrate judge reasoned that since the order demanded only "records" and not the "contents" of their electronic communications, the subscribers had no standing to challenge the compelled disclosure under the SCA. Further, she held that the subscribers had no First Amendment claim involving non-content information, and they had no legitimate Fourth Amendment expectation of privacy in this information.

[2] *In re Application of the United States*, 830 F. Supp. 2d 114 (E.D. Va. 2011).

recording of . . . information by Twitter and subsequent access by the government cannot by itself violate the Fourth Amendment."[3]

The problems illustrated by this case are twofold. First, in the age of big data, the collection and analysis of "non-content" data can yield far more information about someone than was thought when the SCA was first drafted.[4] Properly applied, big data analytics can make record data more illuminating to the analyst than content, heightening concerns over reduced SCA protections for non-content data. Second, since this data is collected by third party providers, the government can obtain this data without dealing with Fourth Amendment protections,[5] possibly bypassing the courts altogether.[6] Furthermore, the government's focus on national security since 2001 has resulted in an increase in requests for such data, some of which remain unexamined due to government claims of state secrecy.[7] This essay argues that the nexus of ubiquitous computing and big data analytics has rendered existing standards of Fourth Amendment protection inadequate, and calls for a reexamination of these doctrines based on today's technologies.

---

[3] *Id*. at 132-33 (citing *U.S. v. Jacobsen*, 466 U.S. 109, 115-17) (1984)).

[4] The statutory definition of "content" is "any information concerning the substance, purport, or meaning of that communication." 18 U.S.C. § 2711(1). The SCA provides greater protection to the "contents of electronic communications" than to their non-content "records." 18 U.S.C. § 2073(a)-(c).

[5] Charlie Savage and Leslie Kaufman, *Phone Records of Journalists Seized by U.S.*, NY TIMES, May 13, 2013. This instance also illustrates the important First Amendment issues at stake.

[6] The government has solicited cooperation and assistance from multiple private companies under the auspices of national security. *See generally* David Kravets, *Court Upholds Google-NSA Relationship Secrecy*, WIRED, May 11, 2012; Brandan Sasso, *Supreme Court Lets AT&T Immunity Stand in Surveillance Case*, THE HILL, Oct. 9, 2012, *available at* http://thehill.com/blogs/hillicon-valley/technology/260951-supreme-court-lets-atat-immunity-stand-in-surveillance-case.

[7] James Bamford, *The NSA Is Building the Country's Biggest Spy Center (Watch What You Say)*, WIRED, Mar. 15, 2012.

MOSAIC THEORY AND THE AGE OF BIG DATA

In recent years, data storage capacities have increased by orders of magnitude, while associated costs have plummeted. Processing speeds have increased to the point that most people carry smartphones that are far more capable than the computers that sat on their desks a few years ago. These factors have combined to enable real time analysis of massive quantities of data, spurring research advances in fields as diverse as atmospheric science, genomics, logistics, and disease prevention.

These capabilities have not gone unnoticed by governments, which have employed big data analytics to reach previously unheard of dimensions of intelligence analysis.[8] These techniques have spilled over into domestic law enforcement, yielding some positive results[9] while at the same time posing new challenges to Fourth Amendment doctrine. And we are the ones supplying the data.

Most Americans own cell phones. We carry them everywhere, and are generally never more than a few feet from one at any time. We use them to send emails and text messages, post messages on Facebook or Twitter, take photos and share them with friends (or the world), and sometimes even to make calls. They are always on, and always on us. Most cell phone users understand that, in order for a cell phone to work, it must be in constant communication with the provider network. The information that is passed back and forth between the phone and the network includes subscriber and location information, and any content that you send or receive.

[8] Big data analytics is especially useful under the "mosaic theory" of intelligence gathering, which holds that small, disparate items of information, though individually of little or no use, can become meaningful when combined with other items of information by illuminating relationships between the data. Notably, the U.S. government has recognized that mosaic theory can work against them, as well, resulting in increased assertions of state secrecy in denying FOIA requests. *See* David E. Pozen, *The Mosaic Theory, National Security, and the Freedom of Information Act*, 115 YALE L.J. 628 (2005).

[9] *See* Frank Konkel, *Boston Probe's Big Data Use Hints at the Future*, FCW, Apr. 26, 2013.

All of this information is collected and stored by the service provider, often without our knowledge.

In fact, providers of all kinds of services make it their practice to collect every bit of data we generate—explicitly or implicitly—and store it for some amount of time.[10] Various privacy laws exist at the state and federal level to prevent the collection of personally identifiable information ("PII"), but big data analytics obviates the need for personal information by leveraging the vast amounts of non-PII data we constantly provide.[11]

THE SHRINKING DISTINCTION BETWEEN "RECORD" AND "CONTENT" DATA UNDER THE SCA

The SCA was enacted in 1986, and was intended to extend privacy protections to new forms of telecommunications and computer technology then just emerging, *e.g.*, cell phones and email.[12] The core of the SCA is 18 U.S.C. § 2703, which articulates procedures by which the government may obtain electronic communications and related information. Section 2703 distinguishes between "content" and (non-content) "records," giving greater protection to the content of a communication.

This distinction is based on Congress's original purpose in enacting the SCA. Because Fourth Amendment privacy protections leave gaps when it comes to information sent

[10] Private companies maintain differing data retention policies, which can be based on government regulation, data management best practices, or internal procedures.

[11] Location data alone can make someone's life an open book. "GPS monitoring generates a precise, comprehensive record of a person's public movements that reflects a wealth of detail about her familial, political, professional, religious, and sexual associations. The Government can store such records and efficiently mine them for information years into the future. And because GPS monitoring is cheap in comparison to conventional surveillance techniques and, by design, proceeds surreptitiously, it evades the ordinary checks that constrain abusive law enforcement practices: 'limited police resources and community hostility.'" *United States v. Jones*, 132 S. Ct. 945, 955-956 (2012) (J. Sotomayor concurring) (internal citations omitted).

[12] *See generally* 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 (2004).

to—and stored by—third parties,[13] the SCA was enacted to fill those gaps by providing additional statutory privacy rights against government access to information stored by service providers. It was reasoned that users may have a "reasonable expectation of privacy"[14] in the substance of their stored communications ("content"), but would not enjoy the same expectation in non-content ("record") information shared with their service provider.

Thus, if the government seeks access to non-content subscriber records under the SCA, their agents may get this information without a warrant, using either a subpoena or a "specific or articulable facts" order, and are not required to provide notice of this access to the subscriber.[15] But, armed with the ability to perform real-time analytics over vast amounts of this data, the government can make non-content information more illuminating than content information, thus skirting the original intent of the SCA's content/non-content distinction.

THIRD-PARTY DOCTRINE

Under current doctrine, the Fourth Amendment does not prohibit the government from obtaining information revealed to a third party who then conveys that information to government authorities, even if the information was revealed on the assumption that it will be used only for a limited purpose and the confidence placed in the third party will not be betrayed.[16] This third-party doctrine has been the basis for courts holding that information "voluntarily disclosed" to

[13] A key reason behind these gaps, third party doctrine, is discussed in more detail below.

[14] The "reasonable expectation" Fourth Amendment test was first articulated in *Katz v. United States*, 289 U.S. 347, 360 (1967) (J. Harlan, concurring), and has recently been "added to" in *Jones* and *Florida v. Jardines*, 133 S. Ct. 1409 (2013).

[15] 18 U.S.C. § 2703(d); 18 U.S.C. § 2703(c)(3).

[16] *United States v. Miller*, 425 U.S. 435, 443 (1976). *See also Smith v. Maryland*, 442 U.S. 735, 743-44 (1979).

service providers, including IP addresses, files shared on private peer-to-peer networks, and historical cell phone location records, does not have Fourth Amendment protection.[17]

But courts have begun to question the application of this doctrine as applied to current technologies and use patterns. This nascent recognition of the advent of ubiquitous computing, made possible through Internet-enabled laptops, tablets, and smart phones, and the resulting "voluntary disclosure" by millions of Americans of vast amounts of non-content information to third party service providers, has raised concerns that the aggregation and analysis of these enormous data sets may be more revealing than content information. For example, one court has observed that cell service providers "have records of the geographic location of almost every American at almost every time of the day and night," enabling "mass or wholesale electronic surveillance" by the government, and holding therefore that "an exception to the third-party-disclosure doctrine applies to cell-site-location records."[18]

CONCLUSION

As Judge Garaufis recently observed, "[i]n order to prevent the Fourth Amendment from losing force in the face of changing technology, Fourth Amendment doctrine has evolved . . . and must continue to do so."[19] For most Americans, the use of "always on, always on us" technology has become an indispensable part of everyday life, forcing us to accept

[17] *See In re Application of the United States*, 830 F. Supp. 2d at 135 (IP addresses); *United States v. Brooks*, 2012 U.S. Dist. LEXIS 178453, *6-*7 (E.D.N.Y. 2012) (private peer-to-peer networks); *United States v. Graham*, 846 F. Supp. 2d 384, 389 (D. Md. 2012) (historical cell site location records)

[18] *In re United States for an Order Authorizing the Release of Historical Cell-Site Info*., 809 F. Supp. 2d 113 (E.D.N.Y. 2011). The full reasoning behind the court's decision is beyond the scope of this essay, but it is worth noting the court's closing observation that "the collection of cell-site-location records, without the protections of the Fourth Amendment, puts our country far closer to [Orwell's] Oceania than our Constitution permits." *Id*. at 127.

[19] *In re United States*, 809 F. Supp. 2d at 126.

the fact that private service providers collect the data we constantly generate. Under existing Fourth Amendment doctrine, this non-content data is afforded few protections, even though it may be more revealing than content data. Courts should therefore recognize that our current Fourth Amendment protections must evolve to adapt to the age of big data analytics.