2015

Confronting Big Data: Applying the Confrontation Clause to Government Big Data Collection

Chad Squitieri

NOTE

CONFRONTING BIG DATA: APPLYING THE CONFRONTATION CLAUSE TO GOVERNMENT DATA COLLECTION

Chad Squitieri[*]

[*] J.D. Candidate, 2016, University of Virginia School of Law. I would like to thank Professors George Fisher of Stanford Law School, Darryl Brown of the University of Virginia School of Law, and Richard D. Friedman of the University of Michigan Law School for reading earlier drafts of this Note and providing insightful responses. I would also like to thank those members of the *Virginia Law Review* who similarly provided invaluable feedback, and Professor Michael Collins of the University of Virginia School of Law for indispensable discussions on this topic. All errors are my own. Lastly, I would like to thank my Grandmother, Sharon, whose benevolence in the face of adversity I have found to be inspiring, and my Grandfather, Chester, whose lifelong commitment to learning I hope to imitate.

INTRODUCTION

HOW did you stumble across this Note, and what does that say about you? What words you queried, how quickly you typed them, the websites you recently visited, and your current geographic location are all useful data points that can be aggregated to form an informative picture of who you are and what you have done.[1]

Companies such as Google collect this data because it can be analyzed for patterns that can predict your future acts.[2] This predictive ability is useful to both a salesman predicting when you might purchase your next pair of shoes,[3] as well as an FBI agent predicting when you may perform your next act of terrorism.[4] By collecting vast amounts of data, commonly referred to as "big data," predictions can be exponentially more accurate than ever before.[5] In addition to predicting what you may *do*, analyzing big data allows for a more detailed depiction of what you have already *done*.[6] It is this backwards-looking feature of big data that this Note will address.

When government investigators request data from companies such as Google, they obtain data on targeted individuals with a guarantee that the data has been collected, stored, and analyzed properly. These guarantees constitute a testimonial statement under the Confrontation Clause.[7] Similar to lab analysts who submit test results of cocaine sam-

[1] See Privacy Policy, Google, https://www.google.com/intl/en/policies/privacy (last visited Aug. 23, 2015) (explaining the type of data Google collects).

[2] Phil Simon, Too Big to Ignore: The Business Case for Big Data 101–02 (2013).

[3] See Viktor Mayer-Schönberger & Kenneth Cukier, Big Data: A Revolution That Will Transform How We Live, Work, and Think 51–52 (2013) [hereinafter Mayer-Schönberger & Cukier, Big Data].

[4] See Byron Acohido, Watch Out, Terrorists: Big Data Is on the Case, USA Today (July 29, 2013, 5:04 PM), http://www.usatoday.com/story/cybertruth/2013/07/29/criminals-terrorists-leave-tracks-in-big-data/2596713.

[5] Patrick Tucker, The Naked Future: What Happens in a World that Anticipates Your Every Move?, at xv (2014) ("[S]ystems are developing perceptions that far exceed our own.").

[6] Soumendra Mohanty, Madhu Jagadeesh & Harsha Srivatsa, Big Data Imperatives: Enterprise 'Big Data' Warehouse, 'BI' Implementations and Analytics 15–16 (2013).

[7] See infra Part III.

ples[8] or blood alcohol levels,[9] this Note argues that analysts involved with the collection, storage, and analysis of big data must be available for confrontation under the Sixth Amendment.[10] At least one federal appeals court has adopted a similar view.[11]

In addressing the constitutionality of modern government surveillance, this Note examines a growing problem. Much of the contemporary academic debate regarding the constitutionality of government surveillance focuses on the President's Article II authority and the Fourth Amendment.[12] Missing from this literature is a detailed discussion of the Confrontation Clause. This Note fills that void by examining the usefulness of the Confrontation Clause in addressing mass data collection by the government.

The usefulness of the Confrontation Clause becomes apparent when one considers the finite ability of the Fourth Amendment to address government data collection. Every federal appeals court to address the issue has found that the President possesses the inherent authority to collect data for foreign intelligence purposes without a warrant.[13] The President's authority to collect data, however, does not provide the govern-

[8] See Melendez-Diaz v. Massachusetts, 557 U.S. 305 (2009).

[9] See Bullcoming v. New Mexico, 131 S. Ct. 2705 (2011).

[10] See infra Part III.

[11] See United States v. Cameron, 699 F.3d 621, 651–52 (1st Cir. 2012).

[12] See, e.g., Michael Avery, The Constitutionality of Warrantless Electronic Surveillance of Suspected Foreign Threats to the National Security of the United States, 62 U. Miami L. Rev. 541 (2008); Monu Bedi, Social Networks, Government Surveillance, and the Fourth Amendment Mosaic Theory, 94 B.U. L. Rev. 1809 (2014); Laura K. Donohue, Bulk Metadata Collection: Statutory and Constitutional Considerations, 37 Harv. J.L. & Pub. Pol'y 757, 863–97 (2014); Orin S. Kerr, The Mosaic Theory of the Fourth Amendment, 111 Mich. L. Rev. 311 (2012); Stephen I. Vladeck, Big Data Before and After Snowden, 7 J. Nat'l Security L. & Pol'y 333 (2014) [hereinafter Vladeck, After Snowden]; John Yoo, The Legality of the National Security Agency's Bulk Data Surveillance Programs, 10 J.L. & Pol'y for Info. Soc'y 301, 316–26 (2014) [hereinafter Yoo, Bulk Data]. But see, e.g., Andrew P. Napolitano, A Legal History of National Security Law and Individual Rights in the United States: The Unconstitutional Expansion of Executive Power, 8 N.Y.U. J.L. & Liberty 396, 460–506 (2014) (discussing the First Amendment); Stephen I. Vladeck, The FISA Court and Article III: A Surreply to Orin, Lawfare (Aug. 5, 2014, 9:31 AM), http://www.lawfareblog.com/2014/08/the-fisa-court-and-article-iii (discussing Article III).

[13] See In re Sealed Case, 310 F.3d 717, 742 (FISA Ct. Rev. 2002) (recognizing that "all the other courts to have decided the issue [have] held that the President did have inherent authority to conduct warrantless searches to obtain foreign intelligence information," and thus "tak[ing] for granted that the President does have that authority"); see also United States v. U.S. District Court, 407 U.S. 297, 322 (1972) (leaving open the possibility that warrantless surveillance may be constitutional in the foreign intelligence context).

ment with unfettered authority to use the data in a criminal proceeding against a defendant.[14] When data is presented at trial against a criminal defendant, the Confrontation Clause is implicated, and the clause's ability to act as a check on government surveillance comes into focus.[15] This flexible check on government surveillance can be attained through the application of contemporary Supreme Court Confrontation Clause doctrine.[16]

Many scholars, however, are hesitant to extend the Supreme Court's contemporary Confrontation Clause doctrine.[17] This Note addresses their concerns, and explains why the common objections to extending current

[14] Yoo, Bulk Data, supra note 12, at 324.

[15] While the word "data" is the plural form of the word "datum," this Note uses phrases such as "data is" in an attempt to be less distracting to the reader. Jane Bambauer, Is Data Speech, 66 Stan. L. Rev. 57, 59 n.3 (2014); Mona Chalabi, 'Data Is' vs. 'Data Are', FiveThirtyEight (Mar. 17, 2014, 1:20 PM), http://fivethirtyeight.com/datalab/data-is-vs-data-are.

[16] See Jeffrey L. Fisher, *Crawford v. Washington*: The Next Ten Years, 113 Mich. L. Rev. First Impressions 9, 12–13 (2014) [hereinafter Fisher, The Next Ten] (acknowledging the Confrontation Clause's impact on prosecutors who use testimonial statements in court, as well as its non-impact on investigators who collect such statements).

[17] See, e.g., Brief for Fern L. Nesson & Charles R. Nesson as Amici Curiae Supporting Respondent at 13–17, Ohio v. Clark, 135 S. Ct. 2173 (2015) (No. 13-1352); Jeffrey Bellin, Applying *Crawford*'s Confrontation Right in a Digital Age, 45 Tex. Tech L. Rev. 33, 42 (2012) [hereinafter Bellin, Digital Age]; Jeffrey Bellin, The Incredible Shrinking Confrontation Clause, 92 B.U. L. Rev. 1865 (2012) [hereinafter Bellin, Shrinking]; Craig M. Bradley, *Melendez-Diaz* and the Right to Confrontation, 85 Chi.-Kent L. Rev. 315, 315 (2010); Thomas Y. Davies, What Did the Framers Know, and When Did They Know It? Fictional Originalism in *Crawford v. Washington*, 71 Brook. L. Rev. 105 (2005); Donald A. Dripps, Controlling the Damage Done by *Crawford v. Washington*: Three Constructive Proposals, 7 Ohio St. J. Crim. L. 521, 536, 539 (2010); George Fisher, The *Crawford* Debacle, 113 Mich. L. Rev. First Impressions 17, 19–25 (2014) [hereinafter Fisher, Debacle]; Joëlle Anne Moreno, Finding Nino: Justice Scalia's Confrontation Clause Legacy from Its (Glorious) Beginning to (Bitter) End, 44 Akron L. Rev. 1211, 1248–51, 1255–56 (2011); Robert P. Mosteller, Confrontation as Constitutional Criminal Procedure: *Crawford*'s Birth Did Not Require That *Roberts* Had to Die, 15 J.L. & Pol'y 685 (2007); Deborah Tuerkheimer, Confrontation and the Re-Privatization of Domestic Violence, 113 Mich. L. Rev. First Impressions 32, 32 (2014); Dylan O. Keenan, Note, *Bullcoming* and Cold Cases: Reconciling the Confrontation Clause with DNA Evidence, 30 Yale L. & Pol'y Rev. Inter Alia 13 (2012). But see, e.g., Fisher, The Next Ten, supra note 16, at 13–15 (outlining "a few things the Court should do . . . to clarify and solidify *Crawford*'s exclusionary rule"); Richard D. Friedman, Come Back to the Boat, Justice Breyer!, 113 Mich. L. Rev. First Impressions 1, 5–7 (2014) (criticizing efforts to narrow *Crawford*'s testimonial definition); Richard D. Friedman, Confrontation and Forensic Laboratory Reports, Round Four, 45 Tex. Tech. L. Rev. 51, 53, 57–80 (2012) [hereinafter Friedman, Round Four] (describing "straightforward" applications of *Crawford*).

doctrine do not apply to big data transfers.[18] Moreover, the Supreme Court's recent decision in *Riley v. California*[19] provides additional support for treating big data as unique.[20]

In Part I, this Note will provide an introduction to big data and the legal authority for its collection by government investigators. Part II will explain the Supreme Court's contemporary Confrontation Clause doctrine. Part III will present the argument that the Confrontation Clause of the Sixth Amendment applies to big data transfers under two independent theories: one theory dealing with individual pieces or small collections of data, and another theory dealing with a novel application of the Mosaic Theory. Part IV will describe Google's procedures for answering government requests for data, and will outline the small number of Google employees that would be required for confrontation.

I. INTRODUCTION TO BIG DATA

Oliver Wendell Holmes, Jr. once observed that "the man of the future is the man of statistics."[21] Today society collects data at a level Justice Holmes might never have imagined. Every day, Google processes thousands of times the amount of data contained in all the printed material in the U.S. Library of Congress, and the stock of information in the world doubles about every three years.[22] Though data collection has increased exponentially since the days of Justice Holmes, the number of predictions derived from data is still "not unmanageably large,"[23] thanks in part to the use of predictive analytics that can derive valuable insights from big data.[24]

[18] See infra Subsection III.D.1.
[19] 134 S. Ct. 2473 (2014).
[20] See infra Subsection III.D.2.
[21] Oliver Wendell Holmes, Jr., The Path of the Law, 10 Harv. L. Rev. 457, 469 (1897).
[22] Mayer-Schönberger & Cukier, Big Data, supra note 3, at 9.
[23] Holmes, supra note 21, at 458.
[24] Eric Siegel, Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die 4 (2013).

A. What Is Big Data?

There is no uniform definition for big data.[25] For the purposes of this Note the following definition will suffice: an amount of data so large that traditional analytical tools must give way to statistical models.[26] Traditional survey methodology consists of collecting data from a representative sample in an attempt to discover something about the relevant population as a whole.[27] The statistical models used to analyze big data, by comparison, are "messier" than these traditional models.[28]

An example of traditional survey methodology is a farmer who estimates the yield of an entire apple orchard by manually counting apples on randomly selected trees. Big data works differently By aggregating large amounts of cheaper—though potentially less correlated—data, one can create a more accurate picture than may have existed had fewer pieces of expensive, but more correlated, data been used.[29] For example, the same farmer might estimate her orchard's yield by aggregating data she has already collected for other reasons, such as how much seed was sown, how many sunny days there have been, and how much fruit various parts of the orchard have produced in the past.[30]

Though "messier" data is used, the sheer amount of data compensates for the shortcoming of each piece of data examined individually. Thus big data is useful when analyzed collectively, rather than being split into

[25] See Mayer-Schönberger & Cukier, Big Data, supra note 3, at 6; Sharon D. Nelson & John W. Simek, Big Data: Big Pain or Big Gain for Lawyers?, 39 L. Prac. 24, 24 (2013).

[26] See Mayer-Schönberger & Cukier, Big Data, supra note 3, at 6; see also Neil M. Richards & Jonathan H. King, Three Paradoxes of Big Data, 66 Stan. L. Rev. Online 41, 42 (2013) ("[S]mall data inputs are aggregated to produce large datasets which analytic techniques mine for insight.").

[27] See Mayer-Schönberger & Cukier, Big Data, supra note 3, at 21; see also David A. Freedman, Sampling, *in* 1 SAGE Encyclopedia of Social Science Research Methods 986, 986–87 (Michael S. Lewis-Beck, Alan Bryman & Tim Futing Liao eds., 2004) (explaining sampling and sample designs).

[28] See Mayer-Schönberger & Cukier, Big Data, supra note 3, at 39.

[29] See Greg Satell, Before You Can Manage Big Data, You Must First Understand It, Forbes (June 22, 2013, 8:39 AM), http://www.forbes.com/sites/gregsatell/2013/06/22/before-you-can-manage-big-data-you-must-first-understand-it ("By vastly increasing the data we use, we can incorporate lower quality sources and still be amazingly accurate.").

[30] See Kowligi R. Krishna, Precision Farming: Soil Fertility and Productivity Aspects 29–68 (2013); Dan Charles, Should Farmers Give John Deere and Monsanto Their Data?, NPR (Jan. 22, 2014, 4:45 PM), http://www.npr.org/blogs/thesalt/2014/01/21/264577744/should-farmers-give-john-deere-and-monsanto-their-data.

individual pieces.[31] Analyzing only some hand-selected pieces of data in a collection and deciding to leave other data out can result in a different picture than would have existed had all or different pieces of data been used.[32]

Because of advances in technology, it is cheaper than ever to collect, store, and analyze vast amounts of data.[33] This allows individuals to seek data that they may never have thought worthwhile to seek before.[34] For example, comparing cereal-purchasing habits to voting habits might reveal that purchasing a given brand of cereal has a given correlation with voting for a certain political party. Instead of acquiring data in a traditional way, such as door-to-door surveying, a campaign manager may prefer to purchase the "less accurate" cereal data and aggregate it with other cheap but messy data, such as "Likes" on a candidate's Facebook page.[35]

Similar to the campaign manager, government investigators have found big data useful in combating terrorism and other criminal acts.[36] The government, however, cannot always know in advance which indi-

[31] See Mayer-Schönberger & Cukier, Big Data, supra note 3, at 13–15, 39; see also Satell, supra note 29 ("And that's the beauty of big data, it can be dumb and still be incredibly useful.").

[32] See Problems with Scientific Research: How Science Goes Wrong, Economist (Oct. 19, 2013), http://www.economist.com/news/leaders/21588069-scientific-research-has-changed-world-now-it-needs-change-itself-how-science-goes-wrong (noting that poor research habits, such as excluding relevant data, can skew results).

[33] See Mayer-Schönberger & Cukier, Big Data, supra note 3, at 15, 83–84, 95 (referring to digitization and datafication).

[34] See Nate Silver, The Signal and the Noise: Why So Many Predictions Fail—But Some Don't 253 (2012) ("[I]t is indeed usually valuable to collect more [data].").

[35] See, e.g., Dan Balz, How the Obama Campaign Won the Race for Voter Data, Wash. Post (July 28, 2013) http://www.washingtonpost.com/politics/how-the-obama-campaign-won-the-race-for-voter-data/2013/07/28/ad32c7b4-ee4e-11e2-a1f9-ea873b7e0424_story.html. But see Rasmus Kleis Nielsen & Cristian Vaccari, Do People "Like" Politicians on Facebook? Not Really. Large-Scale Direct Candidate-to-Voter Online Communication as an Outlier Phenomenon, 7 Int'l J. Comm. 2333, 2334 (2013) (suggesting that people pay little attention to politicians on Facebook and other social media platforms).

[36] See, e.g., Jesús Mena, Investigative Data Mining for Security and Criminal Detection 14 (2003) ("The probability of a crime or an attack involves assessing *risk*, which is the objective of data mining."); Theresa M. Payton & Theodore Claypoole, Privacy in the Age of Big Data: Recognizing Threats, Defending Your Rights, and Protecting Your Family 47 (2014) (explaining how big data "could be used to study patterns for any behavior that might be related to terrorism").

viduals in a population are going to be the "bad guys."[37] This inability has resulted in widespread intelligence-gathering programs such as PRISM, a National Security Agency ("NSA") program that collects data from the Internet for foreign intelligence purposes.[38] According to information leaked by Edward Snowden, the PRISM program acquires data from several companies including Facebook, Google, and Microsoft.[39] Section I.B will describe how data can be collected from these companies.

B. Methods of Government Data Collection

Google, the paradigmatic data-gathering entity,[40] gets requests for data about its users in several different forms. (Throughout the remainder of this Note, Google will be used as a representative example of companies similar to Google.) In addition to traditional search warrants and subpoenas, requests can come in the form of a Foreign Intelligence Surveillance Act ("FISA") request or a National Security Letter ("NSL").[41] Government investigators may be restricted in the type of request they can use to collect data about a given suspect. For example, unlike subpoenas "that can be issued in any sort of criminal case, NSLs can only

[37] See Kate Crawford & Jason Schultz, Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms, 55 B.C. L. Rev. 93, 103–05 (2014) (discussing law enforcement use of big data).

[38] Director of Nat'l Intelligence, Facts on the Collection of Intelligence Pursuant to Section 702 of the Foreign Intelligence Surveillance Act 1 (June 8, 2013), available at http://www.dni.gov/files/documents/Facts%20on%20the%20Collection%20of%20Intelligence%20Pursuant%20to%20Section%20702.pdf; see also Yoo, Bulk Data, supra note 12, at 311–13 (discussing Section 702, under which PRISM is justified).

[39] See Napolitano, supra note 12, at 538–40; see also NSA Slides Explain the PRISM Data-Collection Program, Wash. Post (June 6, 2013), http://www.washingtonpost.com/wp-srv/special/politics/prism-collection-documents (explaining how PRISM is used to collect data).

[40] See Frank Pasquale, Beyond Innovation and Competition: The Need for Qualified Transparency in Internet Intermediaries, 104 Nw. U. L. Rev. 105, 112 (2010) (describing how Google has become a "de facto lawmaker" for much of the Internet).

[41] See, e.g., Transparency Report, Google, https://www.google.com/transparencyreport/userdatarequests/legalprocess/ (last visited Aug. 25, 2015); Information for Law Enforcement Authorities, Facebook, https://www.facebook.com/safety/groups/law/guidelines/ (last visited Aug. 25, 2015); Transparency Report, Twitter, https://transparency.twitter.com/country/us (last visited Aug. 25, 2015); Transparency Report, Yahoo!, https://transparency.yahoo.com/government-data-requests/US-JUL-DEC-2013.html (last visited Aug. 25, 2015).

be used during duly authorized national security investigations."[42] Similarly, FISA requests are intended for foreign intelligence gathering.[43] These nuances merely narrow the *type* of criminal conduct that can be investigated under different requests, but do not affect whether the data is sought or provided for the purposes of an ongoing emergency, or to prove a past event for criminal prosecution.

As will be explained in Part II, whether or not data is sought or provided to prove a past event for criminal prosecution is an important consideration under contemporary Confrontation Clause doctrine. It is therefore significant that after the enactment of the USA PATRIOT Act[44]—in which Congress allowed for more communication between law enforcement and foreign intelligence agencies[45]—data can be collected under the authority of a FISA request or NSL even if the primary purpose of the collection is for criminal prosecution.[46] Similarly, search warrants and subpoenas can be used when the primary purpose of the collection is to prove a past event for criminal prosecution.[47]

Much of the contemporary literature regarding the legality of modern government data collection addresses whether agencies such as the NSA should be subject to domestic criminal justice constraints, requiring search warrants and subpoenas, or special wartime foreign intelligence models where surveillance authority is said to derive from the President's constitutional powers.[48] The role the Confrontation Clause plays regarding data transferred from Google to government investigators, however, does not turn on whether a FISA request, NSL, search warrant, or subpoena is used. In each situation, Google is aware that it is transfer-

[42] Michael German et al., National Security Letters: Building Blocks for Investigations or Intrusive Tools?, A.B.A. J. (Sept. 1, 2012, 10:10 AM), http://www.abajournal.com/magazine/article/national_security_letters_building_blocks_for_investigations_or_intrusive_t.

[43] See 50 U.S.C. § 1801 (2012).

[44] Pub. L. No. 107-56, 115 Stat. 272 (2001).

[45] In re Sealed Case, 310 F.3d 717, 733–34 (FISA Ct. Rev. 2002).

[46] See id.; Charles Doyle, Cong. Research Serv., RL32880, Administrative Subpoenas and National Security Letters in Criminal and Foreign Intelligence Investigations: Background and Proposed Adjustments 19 (2005); James G. McAdams III, Foreign Intelligence Surveillance Act (FISA): An Overview, U.S. Dep't of Homeland Security 8 (March 2007), http://www.fletc.gov/training/programs/legal-division/downloads-articles-and-faqs/research-by-subject/miscellaneous/ForeignIntelligenceSurveillanceAct.pdf.

[47] See, e.g., Illinois v. Gates, 462 U.S. 213 (1983); United States v. Camez, No. 2:12-cr-0004-APG-GWF, 2013 WL 6158402 (D. Nev. Nov. 21, 2013); Gonzales v. Google, Inc., 234 F.R.D. 674 (N.D. Cal. 2006).

[48] See, e.g., Yoo, Bulk Data, supra note 12, at 301, 302.

ring data to government investigators.[49] As will be explained in Part II, this awareness is an important consideration in the Supreme Court's contemporary Confrontation Clause doctrine.

II. CONTEMPORARY CONFRONTATION CLAUSE DOCTRINE

The Confrontation Clause of the Sixth Amendment provides, "In all criminal prosecutions, the accused shall enjoy the right . . . to be confronted with the witnesses against him."[50]

A. The Crawford *Framework*

In *Crawford v. Washington*,[51] the Supreme Court set a new course in its Confrontation Clause jurisprudence.[52] Prior to *Crawford*, the Confrontation Clause was subject to the reliability-based doctrine of *Ohio v. Roberts*.[53] Under the *Roberts* doctrine, a statement presented no Confrontation Clause issue if it fell within a "firmly rooted hearsay exception" or contained "particularized guarantees of trustworthiness."[54]

Distancing itself from the *Roberts* doctrine, the Court in *Crawford* wrote that the Founders of the Constitution understood "witnesses" to be defined as those who "bear testimony,"[55] and "testimony" to mean a "solemn declaration or affirmation made for the purpose of establishing or proving some fact."[56] The Court later made clear that the Confrontation Clause plays no role in determining the admissibility of nontestimonial statements.[57] The question of whether an un-confronted, out-of-court statement will be deemed inadmissible under the Confrontation Clause, therefore, turns first on whether the statement is "testimonial."[58]

[49] See infra Section III.C.
[50] U.S. Const. amend. VI.
[51] 541 U.S. 36 (2004).
[52] See Williams v. Illinois, 132 S. Ct. 2221, 2232 (2012).
[53] 448 U.S. 56 (1980).
[54] Id. at 66.
[55] *Crawford*, 541 U.S. at 51 (quoting 2 Noah Webster, An American Dictionary of the English Language 931 (New York, S. Converse 1828)) (internal quotation marks omitted).
[56] Id.
[57] See Whorton v. Bockting, 549 U.S. 406, 420 (2007).
[58] Bellin, Digital Age, supra note 17, at 39.

B. Defining "Testimonial"

Exactly what constitutes "testimonial" under *Crawford* is less than clear.[59] In *Davis v. Washington*,[60] the Court established a "primary purpose" test to determine if statements made during an emergency 911 call constituted a testimonial statement.[61] The Court declared statements nontestimonial when made "under circumstances objectively indicating that the primary purpose of the interrogation is to enable police assistance to meet an ongoing emergency."[62] The Court declared that statements are testimonial when "the circumstances objectively indicate that there is no such ongoing emergency, and that the primary purpose of the interrogation is to establish or prove past events potentially relevant to later criminal prosecution."[63] In *Michigan v. Bryant*,[64] the Court provided structure to the "primary purpose" test when it held that statements made by a dying gunshot victim were nontestimonial because the statements were made during an emergency resulting from the shooter remaining at large.[65]

The Court most recently addressed the primary purpose test in *Ohio v. Clark*.[66] In *Clark*, a preschool teacher noticed signs of physical abuse on a young child and asked the child what had happened.[67] The child responded that "Dee" had hurt him.[68] The teacher then reported these signs of suspected child abuse to authorities, as Ohio law required her to do.[69] The child was later deemed incompetent to testify at trial, though testimony about what he told his teacher was admitted.[70] Highlighting the fact that the child's statements were made to his teachers, and not a law enforcement officer, the Court held that the child's statements to his

[59] See *Crawford*, 541 U.S. at 68 ("We leave for another day any effort to spell out a comprehensive definition of 'testimonial.'").
[60] 547 U.S. 813 (2006).
[61] See id. at 822.
[62] Id.
[63] Id.
[64] 131 S. Ct. 1143 (2011).
[65] Id. at 1163–64.
[66] Ohio v. Clark, No. 13–1352 (U.S. June 18, 2015).
[67] State v. Clark, 999 N.E.2d 592, 594–95 (Ohio 2013).
[68] Id. at 595.
[69] Id. at 594–95.
[70] Id. at 595.

teachers were not created with "the primary purpose of creating an out-of-court substitute for testimony," and therefore were not testimonial.[71]

In its quest to define "testimonial," the Court has also addressed several cases involving forensic lab reports. In *Melendez-Diaz v. Massachusetts*,[72] the prosecution presented affidavits reporting the results of forensic lab analysis that showed the substance seized from the defendant was cocaine.[73] The Court found the lab results, which were sworn before a notary,[74] to be testimonial statements requiring confrontation of the lab analysts.[75] Similarly, the prosecution in *Bullcoming v. New Mexico*[76] presented a blood alcohol report analyzing the defendant's blood alcohol content.[77] The report was formalized and signed, though was not notarized.[78] The prosecution presented testimony of a lab analyst who was "familiar with the testing device used to analyze [the defendant]'s blood and with the laboratory's testing procedures, but had neither participated in nor observed the test on [the defendant]'s blood sample."[79] Because the Court found the report "more than adequate to qualify . . . as testimonial,"[80] the Confrontation Clause required the *specific* analyst whose testimony was incorporated in the report to be made available for confrontation.[81]

The Court addressed a third lab report case in *Williams v. Illinois*.[82] During a bench trial for aggravated criminal sexual assault, the prosecution presented Sandra Lambatos as an expert witness in forensic biology

[71] *Clark*, No. 13–1352, slip op. at 12 (quoting *Bryant*, 562 U.S. at 358 (internal quotation marks omitted)). The Court also stated, in dictum, that "the primary purpose test is a necessary, but not always sufficient, condition for the exclusion of out-of-court statements under the Confrontation Clause." Id. at 7. But see id. at 3 (Scalia, J., concurring in the judgment) (criticizing the "necessary but not always sufficient" language as "dicta" that is "absolutely false"); Richard D. Friedman, Ohio v. Clark: Some Initial Thoughts, The Confrontation Blog (June 19, 2015, 1:09 AM), http://confrontationright.blogspot.com/2015/06/ohio-v-clark-some-initial-thoughts.html (referring to the court's "necessary but not always sufficient" language as dictum).

[72] 557 U.S. 305 (2009).

[73] Id. at 307.

[74] Id. at 308.

[75] Id. at 311.

[76] 131 S. Ct. 2705 (2011).

[77] Id. at 2711.

[78] Id. at 2717.

[79] Id. at 2707.

[80] Id. at 2717.

[81] Id. at 2716.

[82] 132 S. Ct. 2221 (2012).

and forensic DNA analysis.[83] Although Lambatos had performed neither the test analyzing semen obtained from a vaginal swab taken from the victim, nor the test analyzing blood drawn from the defendant, she had examined the records of the DNA experts who had done so, and had testified that she would call the two samples a match.[84] Justice Alito authored the four-Justice plurality that found that Lambatos's testimony did not violate the Confrontation Clause for two independent reasons.[85]

First, the plurality argued that Lambatos did not testify to the truth of the matter asserted in the lab report, but rather the report was merely used as the basis of her expert opinion.[86] Because the Court in *Crawford* ruled that the Confrontation Clause "does not bar the use of testimonial statements for purposes other than establishing the truth of the matter asserted,"[87] the plurality in *Williams* found no Confrontation Clause violation.[88] Both Justice Thomas and the four-Justice dissent disagreed with this rationale presented by the plurality.[89] Second, the plurality argued that, even if the report had been submitted for its truth, the report was nontestimonial.[90] Here the plurality appeared to invoke the primary purpose test in finding that "[t]he report was sought not for the purpose of obtaining evidence to be used against petitioner . . . but for the purpose of finding a rapist who was on the loose."[91]

Justice Thomas, consistent with his position in previous cases,[92] concluded in an opinion concurring in the judgment in *Williams* that there was no Confrontation Clause violation "solely because [the report] lacked the requisite formality and solemnity to be considered testimonial."[93] Justice Breyer, who wrote a separate concurring opinion, noted

[83] Id. at 2229.
[84] Id. at 2230.
[85] Id. at 2244.
[86] Id. at 2239–41.
[87] Crawford v. Washington, 541 U.S. 36, 60 n.9 (citing Tennessee v. Street, 471 U.S. 409, 414 (1985)).
[88] *Williams*, 132 S. Ct. at 2240.
[89] Id. at 2257 (Thomas, J., concurring in the judgment); id. at 2268 (Kagan, J., dissenting).
[90] Id. at 2228, 2242–43 (plurality opinion).
[91] Id. at 2228.
[92] See, e.g., Michigan v. Bryant, 131 S. Ct. 1143, 1167 (2011) (Thomas, J., concurring in the judgment).
[93] *Williams*, 132 S. Ct. at 2255 (Thomas, J., concurring in the judgment) (internal quotation marks omitted).

that "[s]ix to twelve or more technicians could have been involved,"[94] and thought that additional briefing was necessary to answer "what, if any, are the outer limits of the 'testimonial statements' rule set forth in *Crawford v. Washington*?"[95]

III. APPLYING THE CONFRONTATION CLAUSE TO BIG DATA TRANSFERS

This Note argues that every piece of data transferred from Google to government investigators, which is later used against a criminal defendant at trial, contains a testimonial statement under the Confrontation Clause. As explained below, each piece of data has an implicit guarantee that it has been collected and stored correctly. Additionally, an individual piece of data, or a small collection of data, can be enough to be a testimonial statement accusing the defendant of an act. Lastly, aggregating data can create an *additional* testimonial statement under a Mosaic Theory of the Sixth Amendment. Though confrontation in these instances is crucial, only a small number of witnesses are required for confrontation under the theories presented below.

A. Implicit Guarantees

The first theory under which big data transfers implicate the Confrontation Clause results from the implicit guarantee Google provides to government investigators when Google transfers data. When Google receives a government request for data, it must comply with the request or face the possibility of being held in contempt of court.[96] A court order, or a letter from the Director of the FBI, notifies Google of such a penalty for noncompliance,[97] and thus the company is informed of the solemnity of the situation. In describing the process Google uses to transfer data, Google's Chief Legal Officer stated, "We treat [government requests for

[94] Id. at 2247 (Breyer, J., concurring).

[95] Id. at 2244–45.

[96] See, e.g., 18 U.S.C. § 3511(c) (2012); 50 U.S.C. § 1881a(h)(4)(G), (h)(5)(D) (2012); see also In re Under Seal, 749 F.3d 276 (4th Cir. 2014) (affirming sanctions against a company that failed to comply with court orders to turn over particular information to the government relating to the target of a criminal investigation).

[97] Cf. *Under Seal*, 749 F.3d. at 281 (discussing court orders to a private party enjoining it to submit data to the FBI, and outlining penalties for noncompliance).

data] very seriously. We have lawyers review them."[98] Similar to a traditional witness testifying in court, Google is being asked to provide a "solemn declaration or affirmation made for the purpose of establishing or proving some fact."[99] When Google transfers the data to government investigators, it does so with the implicit guarantee that the company has abided by the government request.[100] This guarantee becomes more apparent when Google, similar to the certified statements in *Melendez-Diaz* and *Bullcoming*, sends data "to [i]nvestigators along with a [c]ertificate of [a]uthenticity."[101]

This Note will refer to the Google employees who handle data in these big data transfers as "data analysts." In *Melendez-Diaz*, the Court noted that "forensic analyst[s] responding to a request from a law enforcement official may feel pressure—or have an incentive—to alter the evidence in a manner favorable to the prosecution."[102] The same rationale applies to data analysts.[103] Even if Google is not as cozy with government investigators as state lab analysts might be, the possibility of a fraudulent or mistaken analyst still exists.[104] The Confrontation Clause is thus implicated where data analysts guarantee the reliability of data sent to government investigators.

[98] Interview by Jackie Long with David Drummond, Chief Legal Officer, Google (June 11, 2013), available at http://www.channel4.com/news/google-prism-fsa-attorney-general-david-drummond.

[99] Crawford v. Washington, 541 U.S. 36, 51 (2004) (quoting 2 Webster, supra note 55, at 931) (internal quotation marks omitted).

[100] Google can be said to guarantee this data because they are obligated to deliver data relevant to the government request, or be held in contempt of court. To deliver corrupted data, or data altered by destruction, without informing the government of such corruption or destruction, would mean that Google has fallen short of its obligation.

[101] Google, Way of a Warrant, Youtube (Mar. 27, 2014), http://www.youtube.com/watch?v=MeKKHxcJfh0 [hereinafter Google, Way of a Warrant]. Even if these steps are not followed for every type of government request, or Google has altered their procedure, these steps are useful as an example.

[102] Melendez-Diaz v. Massachusetts, 557 U.S. 305, 318 (2009).

[103] See Clay Helberg, Conference Report, Third International Applied Statistics in Industry Conference, Pitfalls of Data Analysis (or How to Avoid Lies and Damned Lies) (June 5–7, 1995), http://www2.cs.uregina.ca/~rbm/cs100/notes/spreadsheets/lies.htm (discussing "ways people sometimes 'bend the rules' of statistics").

[104] See id.

1. Collection

One way that Google provides an implicit guarantee to government investigators results from the method Google uses to collect data in response to government requests. Google presumably collects data for marketing purposes or other commercial endeavors, and not for the purpose of providing a testimonial statement.[105] The data analysts involved in this original collection are therefore outside the reach of the Confrontation Clause.[106] After initially collecting the data, however, data analysts replying to a government request must sort through collected data to pick out data "relevant" to the request.[107] The analysts who sort through this trove of collected data in order to seek out data "relevant" to the government request play two crucial roles that implicate the Confrontation Clause. First, these analysts decide what data to include in the transfer to the government, and what data to leave out.[108] Second, these analysts guarantee that the data they say was collected from a user's electronic device or account was indeed collected from such device or account.[109]

Regarding the first role, determining what data to include in an aggregation of data, and what data to leave out, constitutes a vitally important step in data collection.[110] Providing the government with data showing that a defendant queried from his computer "how to kill wife," and "how to bury dead body," tells a very different story than is told by providing the same queries, but adding "how to write a fictional novel about deadly lovers."[111] Because Google seeks to narrow a government request for

[105] See United States v. Cameron, 699 F.3d 621, 642 (1st Cir. 2012).

[106] See infra Section IV.B.

[107] See Google, Way of a Warrant, supra note 101 (noting that broad requests are narrowed, and only relevant information is provided); see also *Cameron*, 699 F.3d at 648.

[108] See Google, Way of a Warrant, supra note 101 (noting that the producer determines what data to provide to the government to abide by the request).

[109] See id. (noting that information is sent to investigators with a certificate of authenticity, and a custodian of records is available to appear in court).

[110] See *Cameron*, 699 F.3d at 648 ("[E]mployees removed the images they thought did not depict child pornography, as said images would presumably not be relevant to the prosecution of a child pornography crime."); cf. How Science Goes Wrong, supra note 32, at 1 ("Modern scientists are doing too much trusting and not enough verifying . . . .").

[111] Cf. Terms and Conditions May Apply (Hyrax Films 2013) (discussing a television writer's web searches involving murder mysteries).

data,[112] it could be vital to a defendant's defense to confront an analyst and ask whether an additional exculpating query existed. A witness with the opportunity to determine whether a defendant ends up on the *New York Times* Best Seller list, or on death row, is a witness the Confrontation Clause presumably covers.

Regarding the second role, data analysts take analysis a step further than the analysts in *Melendez-Diaz* and *Bullcoming*. The lab analysts in *Melendez-Diaz* did not guarantee that the cocaine they tested came from the defendant. It was the police, and not the lab analysts, who arrested the defendant and confiscated the cocaine.[113] Similarly, the lab analysts in *Bullcoming* provided no guarantee that the blood sample shipped to them from police actually came from the defendant's body.[114] By comparison, it is the data analysts who tell government investigators that the selected data was collected from the targeted user's device or account.[115] A prosecutor therefore relies on Google's assertion that the data was indeed collected from the defendant when she offers this evidence at trial.[116] If the data analyst responsible for connecting the data to the defendant is not subjected to confrontation, the defendant has no opportunity to confront the very witness pointing the finger.

2. Storage

A second way that Google provides an implicit guarantee resulting in a testimonial statement stems from the unique difficulties big data faces regarding storage.[117] Data, including data stored in the "cloud," is sus-

[112] Transparency Report, Google, http://www.google.com/transparencyreport/userdatarequests/faq/#what_does_fisa_compel (last visited Aug. 25, 2015).

[113] See Melendez-Diaz v. Massachusetts, 557 U.S. 305, 307 (2009).

[114] See Bullcoming v. New Mexico, 131 S. Ct. 2705, 2710 (2011).

[115] See Google, Way of a Warrant, supra note 101 (describing authentication mechanisms such as certificates of authenticity, custodians of record, and corrections of government mistakes such as misspelled names).

[116] Extending this argument to the extreme, all physical evidence could be considered "testimonial." For example, a gun submitted into evidence might contain the police officer's implicit assertion that he found the gun at the defendant's house. These types of underlying statements could often be handled through the use of stipulation and traditional foundation evidence.

[117] See Sean Gallagher, The Great Disk Drive in the Sky: How Web Giants Store Big—and We Mean Big—Data, Ars Technica (Jan 26, 2012, 9:00 PM), http://arstechnica.com/business/2012/01/the-big-disk-drive-in-the-sky-how-the-giants-of-the-web-store-big-data (discussing the complexity of the systems Google uses to store data).

ceptible to corruption while in storage.[118] While corruption can be corrected,[119] an analyst providing the government with data is implying that the data is correct, and that the data has not been altered by unreported corruption.[120] Stored data is also susceptible to destruction.[121] As illustrated above in the "deadly lovers" example, missing data, perhaps missing because of destruction, can drastically alter an analysis.[122]

Analysts involved in transferring requested data to the government are implicitly providing a statement that the data has not been altered by unreported destruction or corruption.[123] Similar to the lab analyst in *Bullcoming*, who was to "not[e] any circumstance or condition which might affect the integrity of the sample or otherwise affect the validity of the analysis,"[124] a data analyst implies that the condition of the data is sufficient to abide by the government's request unless otherwise noted.[125] A criminal defendant should have the right to confront this data analyst in order to inquire whether the data is at risk of containing misinformation, or whether additional data lost to destruction would have painted a different picture.

A critical reader might contend that confrontation is unnecessary to respond to the guarantees of proper storage where, similar to the data in *Williams*, there may be telltale signs that disclose when data is corrupted.[126] This argument is insufficient for three reasons. First, the defendant in *Williams* had the opportunity to confront an expert witness, rather than no witness at all.[127] Second, even if telltale signs existed that could show data *corruption*, that doesn't necessarily provide the defendant with an opportunity to inquire into data lost to *destruction*. Third, even if

[118] See Jordan Tigani & Siddartha Naidu, Google BigQuery Analytics 356 (2014) [hereinafter Tigani & Naidu, BigQuery]; Lakshmi N. Bairavasundaram et al., An Analysis of Data Corruption in the Storage Stack, 6th USENIX Conference on File and Storage Technologies 223 (2008), available at https://www.usenix.org/legacy/event/fast08/tech/full_papers/bairavasundaram/bairavasundaram.pdf.

[119] See Bairavasundaram et al., supra note 118, at 224.

[120] See supra text accompanying note 100.

[121] See Tigani and Naidu, BigQuery, supra note 118, at 25 ("[S]oftware is fallible.").

[122] See supra Subsection III.A.1.

[123] See supra text accompanying note 100.

[124] Bullcoming v. New Mexico, 131 S. Ct. 2705, 2711 (2011) (quoting certificate of analyst).

[125] See supra text accompanying note 100.

[126] Williams v. Illinois, 132 S. Ct. 2221, 2231 (2012) (noting lack of "telltale signs" of defective data).

[127] See id. at 2230.

there was a reliable way to see if these testimonial statements were subject to corruption and destruction worries, the Constitution is left unsatisfied. The Confrontation Clause "commands . . . that reliability be assessed in a particular manner: by testing in the crucible of cross-examination."[128]

3. Analysis of Small Collections of Data

In some instances, a single piece or small collection of data can constitute a testimonial statement in addition to the analyst's guarantee of proper collection and storage. Compare two hypotheticals. In both hypotheticals the defendant, in the course of espionage, stabs a man in an alley. Imagine also that, in both hypotheticals, the police question a witness who can place the defendant in the alley at the time of the stabbing. In addition, the witness tells the police that the defendant asked "how do I wipe fingerprints off a murder weapon?"

The difference between the two hypotheticals is that in one, the witness claims that he was walking by the alley and saw the defendant with his own eyes, and heard the defendant ask about fingerprints with his own ears. In the other, the witness is a Google analyst who tells investigators that the phrase "how do I wipe fingerprints off a murder weapon?" was queried from the defendant's smartphone at the time of the crime, and was queried from the location of the alley.

The witness in the first hypothetical is a straightforward example of someone providing a testimonial statement, and the defendant would have the right to confront this witness.[129] The same rationale should apply to the Google analyst.[130] In both hypotheticals, the witness is telling investigators that the defendant was at a given location at a given time.

Admission of the question that the witness is alleging the defendant to have asked is trickier. The defendant asking "how do I wipe fingerprints off a murder weapon?" is unlikely to be a testimonial statement because

[128] Crawford v. Washington, 541 U.S. 36, 61 (2004); see also Ohio v. Clark, No. 13–1352, slip op. at 2 (U.S. June 18, 2015) (Scalia, J., concurring in the judgment) ("*Crawford* remains the law.").

[129] See *Crawford*, 541 U.S. at 68–69 (holding that a witness statement to the police incriminating the defendant in stabbing was a testimonial statement).

[130] See United States v. Cameron, 699 F.3d 621, 653–54 (1st Cir. 2012) (finding reports containing location data to require confrontation). But see id. at 654 (Howard, J., dissenting) (criticizing the majority for "taking an unjustified step beyond what current Supreme Court precedent dictates").

it is unlikely his own words were provided with the intent to be used against him in a criminal proceeding.[131] The claim made by the *witness* to investigators that the defendant said such a thing, however, is a testimonial statement.[132] Whether the witness claims to have witnessed the defendant ask the question vocally or electronically, the defendant has the right to confrontation.

In many situations, a limited amount of data can be used to accuse the defendant of committing a crime.[133] A computer, however, does not directly say that John Doe was driving at a given speed, bought or sold something illegal, or has a gambling addiction.[134] It takes a person to translate what the computer provides into relevant evidence.[135] By "translate," this Note refers to obvious acts of translation such as converting zeros and ones into English, as well as much more subtle acts of analysis. A zealous prosecutor might argue that data transferred from Google to the government is sufficiently formulaic to be considered computer-produced conclusions rather than human assertions. For the reasons provided in this subsection, such an argument is incorrect.[136] Regardless of whether the translation is subtle or obvious, the Confrontation Clause is implicated.

Regarding the more obvious acts of translation, the importance of an analyst's ability and decisions is clear.[137] If an analyst converts computer-generated data showing that the defendant was in Hollywood, California, into a statement accusing the defendant of being in Hollywood, Florida, the usefulness of confrontation seems straightforward.

---

[131] See James J. Tomkovicz, Constitutional Exclusion: The Rules, Rights, and Remedies That Strike the Balance Between Freedom and Order 397 n.398 (2011) ("The use of a defendant's self-inculpatory hearsay statements at trial raises no Sixth Amendment issue."); Bellin, Digital Age, supra note 17, at 34 ("[F]ew electronic utterances appear to fall within the Court's definition of 'testimonial.'").

[132] See *Crawford*, 541 U.S. at 52 ("Statements taken by police officers in the course of interrogations are also testimonial under even a narrow standard.").

[133] See Riley v. California, 134 S. Ct. 2473, 2492 (2014) (discussing locational data).

[134] See id. at 2490, 2492 (discussing the capabilities of mobile applications).

[135] See Erick J. Poorbaugh, Note, Interfacing Your Accuser: Computerized Evidence and the Confrontation Clause Following *Melendez-Diaz*, 23 Regent U. L. Rev. 213, 220–29 (2010).

[136] See id. for an explanation of when a computer-generated statement transforms into a statement made by a person.

[137] See, e.g., Casen B. Ross, Comment, Clogged Conduits: A Defendant's Right to Confront His Translated Statements, 81 U. Chi. L. Rev. 1931 (2014).

Smaller acts of analysis should also implicate the defendant's right to confrontation. Assuming Google transfers data to the government in its most "raw" form, the Google analyst who handles the data still makes several decisions that represent the analyst's opinion. These opinions are carried forward with the data. This is because there is no such thing as "raw data" that is untouched by the perceptions of those who handle it.[138] Different professions operate under different premises as to what counts as data, and how data should be treated and relied on.[139] Data can therefore never truly be the objective source it is often claimed to be.[140] Some courts have been too quick to assume the opposite,[141] though in one notable instance the Supreme Court has corrected a lower court for doing so.[142] Just like a traditional witness, a data analyst takes in data from the world and translates it into a testimonial statement.

4. Lower Court Application

This Note's first theory under which big data transfers can implicate the Confrontation Clause—where testimonial statements result from the collection, storage, and analysis of data—already finds support in some lower courts. At least one federal district court appears ready to adopt this Note's first theory,[143] and at least one federal appeals court has already done so.[144] In *United States v. Cameron*,[145] the defendant objected on Confrontation Clause grounds to Google, Yahoo!, and CyberTipline

[138] See Lisa Gitelman & Virginia Jackson, Introduction *to* "Raw Data" is an Oxymoron 1 (Lisa Gitelman ed., 2013) [hereinafter Gitelman, Oxymoron]; Laura Kurgan, Close Up at a Distance: Mapping, Technology, and Politics 35 (2013); Tara R. Price, Note, "Bull" Coming from the States: Why the Supreme Court Should Use *Williams v. Illinois* to Close One of *Bullcoming*'s Confrontation Clause Loopholes, 39 Fla. St. U. L. Rev. 533, 550–51 (2012).

[139] See Gitelman, Oxymoron, supra note 138, at 7.

[140] See id.; Kurgan, supra note 138, at 35.

[141] See, e.g., United States v. Washington, 498 F.3d 225, 230–31 (4th Cir. 2007) (referring to the identification of drugs in a defendant's blood as "a conclusion drawn only from the machines' data," whose source was "independent of human observation or reporting"); State v. Bullcoming, 226 P.3d 1, 9 (N.M. 2010) (referring to analyst as a "mere scrivener").

[142] See Bullcoming v. New Mexico, 131 S. Ct. 2705, 2714 (2011) (disagreeing with the "mere scrivener" description of an analyst).

[143] United States v. Muhammad, No. 1:14cr36-HSO-RHW, 2014 WL 6680606, at *3 (S.D. Miss. Nov. 25, 2014) (acknowledging that the court "cannot conclude that the statements of a witness who is necessary to establish the authenticity of records which consist of [electronic] statements attributable to Defendant are clearly not 'testimonial'").

[144] United States v. Cameron, 699 F.3d 621 (1st Cir. 2012).

[145] Id.

records presented at his trial for crimes involving child pornography.[146] In finding some admitted records to have violated the Confrontation Clause, the First Circuit distinguished records that contained only data "collected automatically in order to further . . . business purposes"[147] from records that constituted statements made with the primary purpose of "establishing or proving past events potentially relevant to a later criminal prosecution."[148] As justification for distinguishing these two types of records, the First Circuit cited *Melendez-Diaz* for the proposition that, though business records often do not implicate the Confrontation Clause, some business records may still do so if the primary purpose of the records is to establish or prove past events potentially relevant to later criminal prosecution.[149]

As the above discussion of collection, storage, and analysis demonstrates, this Note agrees with the First Circuit that the Confrontation Clause is implicated when a report "do[es] not merely present pre-existing data . . . [but] convey[s] an *analysis* that was performed using pre-existing data."[150] Accentuating the fact that data analysts removed images that they thought did not depict child pornography before forwarding the data, the First Circuit found the analysts who prepared the data in question to have created a new statement with the primary purpose that was law enforcement related.[151] This Note agrees with this line of reasoning.

This Note will now take the logic in *Cameron* one step further. By distinguishing the situation where data analysts submit *some* of the data that has been collected, the First Circuit seemed to suggest that the Confrontation Clause is not implicated where data analysts submit *all* of the data they have at their disposal.[152] As explained in Section III.B, however, by applying the Mosaic Theory to big data transfers, the Confrontation Clause can still be implicated when Google submits *all* of the data it has collected on an individual.

[146] Id. at 627, 638.
[147] Id. at 641.
[148] Id. at 643 (quoting Bullcoming v. New Mexico, 131 S. Ct. 2705, 2714 n.6 (2011) (internal quotation marks omitted)).
[149] Id. at 640 (citing Melendez-Diaz v. Massachusetts 557 U.S. 305 (2009)).
[150] Id. at 647.
[151] Id. at 648.
[152] Id. at 647.

B. The Mosaic Theory

The second theory under which this Note argues big data transfers can constitute testimonial statements is grounded in a novel application of the Mosaic Theory.[153] The Mosaic Theory is typically discussed in connection with the Fourth Amendment,[154] and can be described as a theory "envision[ing] thousands of bits and pieces of apparently innocuous information, which when properly assembled create a picture."[155] Data analyzed in aggregate can often *create* more value than just the summation of the individual pieces.[156] The Mosaic Theory therefore offers a perfect framework to examine the synergy created by big data.

Under the Mosaic Theory of the Sixth Amendment, a witness who provides "*n*" testimonial statements might be said to have provided "*n + 1*" (or more) testimonial statements when all "*n*" statements are considered in aggregate. Even if the Mosaic Theory is ignored, however, one is still left with an aggregation of testimonial statements to which the Confrontation Clause applies individually.[157]

1. Implicit Conclusions

Under the Mosaic Theory, conclusions can sometimes be drawn through an aggregation of data without an analyst providing an explicit conclusion of their own. For example, a witness who states that she (1) drew four equal lines that were (2) connected by four ninety degree angles, has also stated that (3) she drew a square. A similar example would involve a witness accusing a defendant of committing the separate elements of a crime, but not expressly stating that the defendant committed the crime. This Note refers to such conclusions as "implicit conclusions." Addressing implicit conclusions within big data helps prevent attempts to skirt the requirements of the Constitution when providing a testimonial statement.[158] Consider an example.

---

[153] See Kerr, supra note 12, for an explanation of the Mosaic Theory.

[154] See id. at 320.

[155] Doe v. Gonzales, 449 F.3d 415, 422 (2d Cir. 2006) (Cardamone, J., concurring).

[156] See Mayer-Schönberger & Cukier, Big Data, supra note 3, at 76.

[157] See supra Section III.A (explaining testimonial statements regarding collection, storage, and analysis of small amounts or individual pieces of data).

[158] See Davis v. Washington, 547 U.S. 813, 838 (2006) (Thomas, J., concurring in the judgment in part and dissenting in part) (acknowledging the problem of prosecutorial evasion of the Confrontation Clause).

If the prosecution submitted, for the truth of the matter asserted, a witness affidavit stating "I witnessed Walter verbally assault Henry," Walter would have the right to confront that witness.[159] The Confrontation Clause should not apply differently if the witness's affidavit contained an aggregation of statements that as a whole deliver the same message. Such an affidavit might include statements such as: Walter and Henry were standing next to each other, a human voice verbally assaulted Henry, Walter was moving his lips and staring at Henry, the human voice sounded like Walter's, and there was nobody else within hearing distance of Walter and Henry. Although this collection of statements does not create perfect certainty that the witness saw Walter assault Henry—the witness, for example, might have mistaken his own conscience for an external voice—such a lack of certainty is irrelevant for the purposes of the Confrontation Clause. The statement "I saw Walter verbally assault Henry" is subject to the very same uncertainty as the aggregation of statements describing the same event. The Confrontation Clause should apply the same, even if the statement "I witnessed Walter verbally assault Henry" was deconstructed into a collection of lesser statements that in aggregate deliver the same message.

2. Explicit Conclusions

In addition to implicit conclusions, data analysts may examine an aggregation of data and present a conclusion as to what it means. For example, a data analyst might examine data showing that someone with the defendant's height, weight, fingerprints, gait, and irregular heartbeat was at location "*x*" at time "*t*."[160] This data analyst might therefore conclude that the *defendant* was at location "*x*" at time "*t*." This is an example of what this Note refers to as an "explicit conclusion." An explicit conclusion is nothing more than a data analyst's opinion as to what a collection of data means. Though it may be a very reliable opinion, the Confronta-

[159] See Crawford v. Washington, 541 U.S. 36, 68–69 (2004) (holding that a witness statement to the police incriminating the defendant in stabbing was a testimonial statement implicating the Confrontation Clause).

[160] This data can be collected using "Health & Fitness" apps available on iTunes. See iTunes App Store Health & Fitness, Apple, https://itunes.apple.com/us/genre/ios-health-fitness/id6013?mt=8 (last visited Jan. 17, 2015).

tion Clause cannot be sidestepped simply by extraconstitutional guarantees of reliability.[161]

This Note cannot definitively state whether an explicit conclusion is more likely to be formed by a Google analyst, or a data analyst on the government side of a transfer. Google transfers data to the government both by hand, and through file-transferring technology that allows parties to upload and download files between each other.[162] Although the method Google uses to transfer data to government investigators has been made public, the degree of analysis performed by Google before transferring the data is relatively hard to determine.[163] Regardless of whether a government or Google analyst forms the conclusion, if data is presented at trial in the form of an explicit conclusion, then a criminal defendant has the right to confront the analyst who formed that conclusion.[164]

C. Google's Intent and Targeted Individuals

The intent of the parties is an important consideration under the primary purpose test.[165] In *Ohio v. Clark*, the Court asked whether the statement was "given with the primary purpose of creating an out-of-court substitute for trial testimony."[166] Judged under such a framework, when Google transfers data to government investigators they provide a testimonial statement. Google's ability to designate a "custodian of records" to testify regarding other data analysts' out-of-court statements

[161] See *Crawford*, 541 U.S. at 61.

[162] Claire Cain Miller, Google Offers Some Detail About How It Transfers Data to the Government, N.Y. Times Bits Blog (June 12, 2013, 5:52 PM), http://bits.blogs.nytimes.com/2013/06/12/google-offers-some-detail-about-how-it-transfers-data-to-the-government/?_php=true&_type=blogs&_r=0.

[163] See David Drummond, Asking the U.S. Government to Allow Google to Publish More National Security Request Data, Official Google Blog (June 11, 2013), http://googleblog.blogspot.com/2013/06/asking-us-government-to-allow-google-to.html.

[164] See Bullcoming v. New Mexico, 131 S. Ct. 2705 (2011); Melendez-Diaz v. Massachusetts, 557 U.S. 305 (2009); United States v. Cameron, 699 F.3d 621, 642–43 (1st Cir. 2012).

[165] Michigan v. Bryant, 131 S. Ct. 1143, 1156 (2011).

[166] Ohio v. Clark, No. 13–1352, slip op. at 12 (U.S. June 18, 2015) (quoting *Bryant*, 131 S. Ct. at 1155) (internal quotation marks omitted).

suggests that Google is providing data with the intent to create an out-of-court substitute for trial testimony.[167]

In applying the primary purpose test, the Court in *Clark* highlighted both the formality of the setting in which the statements were made, and whether there was an ongoing emergency.[168] Distinguishing the informal setting in which the teachers questioned the child in *Clark* from formal statements made to law enforcement officers, the Court found the child's statements to be nontestimonial.[169] By comparison, statements made by Google to inquiring government investigators are more like the types of formal statements made to law enforcement officers that the Court in *Clark* was concerned about. In fact, the Court in *Clark* noted that statements made to law enforcement officers are more likely to be found testimonial than statements made to non-law enforcement officers.[170]

Relevant to the ongoing emergency factor of the primary purpose test, Google has a special process for emergency requests.[171] Google has explained that "[t]he government needs legal process—such as a subpoena, court order or search warrant—to force Google to disclose user information. Exceptions can be made in certain emergency cases. . . ."[172] Google describes an example of an emergency as "involving kidnapping or bomb threats," and stated that "[e]mergency requests must contain a description of the emergency and an explanation of how the information requested might prevent the harm."[173] Having a separate procedure for emergency requests suggests that, for nonemergency requests where "[t]he government needs legal process,"[174] Google intends their statement to investigators to establish or prove past events potentially relevant to later criminal prosecution.[175]

[167] See Google, Way of a Warrant, supra note 101; see also United States v. Camez, No. 2:12-cr-0004-APG-GWF, 2013 WL 6158402, at *8 (D. Nev. Nov. 21, 2013) (denying non-party Google, Inc.'s motion to quash trial subpoena).

[168] *Clark*, No. 13–1352, slip op. at 7–9.

[169] Id. at 8–9.

[170] Id. at 7.

[171] Transparency Report, Google, http://www.google.com/transparencyreport/userdatarequests/legalprocess/#does_a_law_enforcement (last visited Feb. 8, 2015).

[172] Id.

[173] Id.

[174] Id.

[175] See Davis v. Washington, 547 U.S. 813, 822 (2006) (recognizing that statements are testimonial when derived from an interrogation whose primary purpose "is to establish or prove past events potentially relevant to later criminal prosecution").

Whether government investigators flag their request as an emergency or nonemergency also seems telling of the investigators' intentions. Even more telling, FISA has a provision where data can be collected without prior court approval if the Attorney General "reasonably determines that an emergency situation exists."[176] This suggests that when the emergency situation provision is not used, the communication between the government and Google under a FISA request would not be considered an ongoing emergency under the "primary purpose" test.

Whether or not government investigators request data on a "targeted" individual is also an important consideration under the primary purpose test. In finding no Confrontation Clause violation, the plurality in *Williams* stressed that the lab report "was not prepared for the primary purpose of accusing a targeted individual," and that when the sample was sent for testing, the defendant was not "under suspicion at that time."[177] Big data transfers, by comparison, satisfy this targeted individual requirement.[178]

The Chief Architect at Google has written that "the only way in which Google reveals information about users are when we receive lawful, specific orders about individuals."[179] Similarly, Microsoft issued a statement that the company "only ever compl[ies] with orders for requests about specific accounts or identifiers."[180] Additionally, Google has made available statistics detailing the extent of government requests.[181] Google filters these statistics in several ways, including through a "Users/Accounts Specified" feature.[182] Google has also stated that they scrutinize government requests and narrow the scope if possible.[183] Unlike the DNA analysis in *Williams*,[184] the data requested by the

[176] 50 U.S.C. § 1805(e)(1)(A) (2012).

[177] Williams v. Illinois, 132 S. Ct. 2221, 2243 (2012).

[178] See Long, supra note 98 ("[W]e get specific orders. They are under the law in the US, targeted orders.").

[179] Yonatan Zunger, Google Plus (June 7, 2013), https://plus.google.com/+YonatanZunger/posts/huwQsphBron.

[180] Brad Smith, Responding to Government Legal Demands for Customer Data, Microsoft on the Issues (July 16, 2013), http://blogs.microsoft.com/on-the-issues/2013/07/16/responding-to-government-legal-demands-for-customer-data.

[181] Transparency Report, Google, http://www.google.com/transparencyreport/userdatarequests/ (last visited Sept. 12, 2014).

[182] Id.

[183] Google, Way of a Warrant, supra note 101.

[184] Williams v. Illinois, 132 S. Ct. 2221, 2228 (2012).

government from Google appears to be tailored toward specifically targeted individuals.

D. Expanding Contemporary Doctrine

Some scholars appear loath to expand contemporary Confrontation Clause doctrine in the area of forensic analysis examined in *Bullcoming* and *Melendez-Diaz*.[185] While expanding contemporary doctrine in this area may at first appear controversial,[186] there are considerable reasons to do so when it comes to big data transfers. First, the common objections raised in present literature to expanding contemporary doctrine do not apply to big data transfers. Second, recent Supreme Court cases have hinted that the Court is prepared to expand doctrine in this area. Lastly, there is a beneficial policy reason that supports such an expansion: the promotion of privacy at reduced cost to security.

1. Objections Do Not Apply to Big Data

One objection to expanding the forensic evidence line of cases is that it would be undesirably costly.[187] Such an unaffordable increase in cost, however, will not materialize for three reasons. First, as argued in Part IV, only individuals who have provided a stand-alone testimonial statement should implicate the Confrontation Clause. One can imagine this group of analysts as a "bottle neck," or top of a pyramid. There are presumably fewer analysts at the top of the pyramid, interacting with gov-

[185] See sources cited supra note 17.

[186] See Tom Jackman & Rosalind S. Helderman, Kaine Calls Legislative Session to Change Laws After Ruling on Trial Testimony, Wash. Post (July 23, 2009), http://www.washingtonpost.com/wp-dyn/content/article/2009/07/22/AR2009072203533.html (describing state legislative action necessitated by *Melendez-Diaz*).

[187] See *Williams*, 132 S. Ct. at 2228 (warning of "economic pressures [which] would encourage prosecutors to forgo DNA testing and rely instead on older forms of evidence . . . that are less reliable"). Amici and subsequent analyses pertaining to *Williams* offered similar warnings. See, e.g., Brief for Fern New York County District Attorney's Office and the New York City Office of the Chief Medical Examiner as Amici Curiae Supporting Respondent at 8–12, *Williams*, 132 S. Ct. 2221 (No. 10-8505); Sean K. Driscoll, "I Messed Up Bad": Lessons on the Confrontation Clause from the Annie Dookhan Scandal, 56 Ariz. L. Rev. 707, 736–37 (2014); Andrew W. Eichner, Note, The Failures of *Melendez-Diaz v. Massachusetts* and the Unstable Confrontation Clause, 38 Am. J. Crim. L. 437, 449–51 (2011). But see *Williams*, 132 S. Ct. at 2275 n.6 (Kagan, J., dissenting) (noting lack of evidence to support plurality's warning in *Williams*); Friedman, Round Four, supra note 17, at 77 ("[M]y response [to the plurality's warning in *Williams*] is, Oh, come on, really.").

ernment investigators and thus implicating the Confrontation Clause, than there are analysts performing more attenuated acts such as originally collecting the data or mopping the front lobby.[188] Second, though mistakes can happen,[189] and an attorney may fish for a mistake by calling every analyst possible, there are internalized costs associated with such a strategy.[190] While rules of evidence cannot trump the Constitution,[191] Federal Rule of Evidence 403[192]—and state equivalents—can limit the incentive of calling repetitive witnesses.[193] Third, notice-and-demand statutes can be used as a cost-limiting tool.[194] Notice-and-demand statutes require prosecutors "to provide notice to the defendant of [their] intent to use an analyst's report as evidence at trial, after which the defendant is given a period of time in which he may object to the admission of the evidence absent the analyst's appearance live at trial."[195] Through the use of notice-and-demand statutes, a defendant may opt to call fewer witnesses than he has the right to, and thereby reduce trial costs.[196]

Another objection to expanding the forensic evidence line of cases is that doing so would hamper the prosecution of sexual assault and do-

[188] See Richard D. Friedman, Is There a Multi-Witness Problem with Respect to Forensic Lab Tests?, The Confrontation Blog (Dec. 7, 2010, 10:11 AM), http://confrontationright.blogspot.com/2010/12/is-there-multi-witness-problem-with.html (finding an average of 1.24 witnesses to present DNA evidence).

[189] See, e.g., *Williams*, 132 S. Ct. at 2264 (Kagan, J., dissenting) (describing an analyst who confused DNA samples collected from the defendant and from the victim).

[190] See Richard A. Posner, Frontiers of Legal Theory 338–40 (2001) (describing a cost-benefit analysis of obtaining and using more evidence); see also Bradley, supra note 17, at 325–27 (suggesting defense counsel must present reasonable explanations for cross-examination).

[191] *Williams*, 132 S. Ct. at 2256 (Thomas, J., concurring in the judgment).

[192] Fed. R. Evid. 403 ("The court may exclude relevant evidence if its probative value is substantially outweighed by a danger of . . . unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence.").

[193] See Old Chief v. United States, 519 U.S. 172, 182–85 (1997) (considering evidentiary alternatives); Fed. R. Evid. 403 advisory committee's note ("The availability of other means of proof may also be an appropriate factor."); Posner, supra note 190, at 349 (recognizing a judge's ability to "limit the amount of search [for evidence] that the lawyers do").

[194] Danae VanSickle Grace, Note, The Sky Is Not Falling: How the Anticlimactic Application of *Melendez-Diaz v. Massachusetts* to Oklahoma's Laboratory Report Procedures Allows Room for Improvement, 63 Okla. L. Rev. 383, 383 (2011); Mark Hansen, Taking Techs to Trial: Two Terms in a Row, Justices Weigh Bringing Lab Analysts into Court, 96 A.B.A. J. 17, 18 (Jan. 2010).

[195] Melendez-Diaz v. Massachusetts, 557 U.S. 305, 326 (2009).

[196] See id. at 314 n.3; Hansen, supra note 194, at 18.

mestic abuse cases where survivors cannot cooperate and forensic evidence is extraordinarily useful.[197] These serious concerns, however, need not necessarily result from requiring confrontation of data analysts.[198] These concerns are motivated by the possibility that increasing the cost of presenting forensic evidence will lead prosecutors to either refrain from prosecuting, or to rely on less reliable evidence.[199] As explained above, however, requiring confrontation of data analysts will not lead to severe increases in cost. Additionally, more fundamental concerns regarding sexual abuse and domestic violence can be addressed through a robust forfeiture doctrine, where a defendant forfeits the right to confrontation when the defendant is responsible for the unavailability of the witness.[200]

Lastly, many scholars have been critical of the historical and textual justifications in *Crawford*, and the resulting applications in *Bullcoming* and *Melendez-Diaz*.[201] Confrontation Clause doctrine, however, had been inconstant before *Crawford*.[202] Considering that other criminal procedure rights have had nearly an additional half-century to settle than

[197] See Bellin, Shrinking, supra note 17, at 1910 (referring to "the difficulties of prosecuting domestic violence and child abuse offenses where victims are unable or unwilling to cooperate with prosecutors"); Sarah M. Buel, *Davis* and *Hammon*: Missed Cues Result in Unrealistic Dichotomy, 85 Tex. L. Rev. 19 (2007) (addressing similar concerns with *Davis v. Washington* and *Hammon v. Indiana*); Dripps, supra note 17 (examining the effect of *Crawford* and its progeny on domestic violence cases); Tuerkheimer, supra note 17 (addressing domestic abuse cases since *Crawford*).; Keenan, supra note 17, at 14 (warning that after *Bullcoming v. New Mexico*, "[p]rosecutors should . . . be concerned that they might lose an essential tool [in the form of DNA database evidence] for prosecuting rape cases and seeking justice for victims").

[198] See Williams v. Illinois, 132 S. Ct. 2221, 2275 n.6 (2012) (Kagan, J., dissenting).

[199] See id. at 2228 (plurality opinion).

[200] See Giles v. California, 554 U.S. 353, 367–68 (2008) (explaining that the right to confrontation can be forfeited when defendant intended to prevent a witness from testifying); see also Deborah Tuerkheimer, *Crawford*'s Triangle: Domestic Violence and the Right of Confrontation, 85 N.C. L. Rev. 1 (2006) (addressing forfeiture and domestic violence cases).

[201] See Bellin, Shrinking, supra note 17, at 1877–78 (criticizing the Court for restricting the right of confrontation to only testimonial statements); Bradley, supra note 17, at 320, 322, 327–28 (same); Davies, supra note 17, at 106–07 (same); Fisher, Debacle, supra note 17, at 19–22 (same). But see Fisher, The Next Ten, supra note 16, at 10 ("*Crawford* is fundamentally sound.").

[202] See Akhil Amar, Sixth Amendment First Principles, 84 Geo. L.J. 641, 641, 690 (1996) (noting, pre-*Crawford*, that "the legal community lacks a good map of [the Six Amendment's] basic contours," with Confrontation Clause case law in particular being "surprisingly muddled in logic and exposition").

contemporary Confrontation Clause doctrine has, such criticism is premature.[203]

Even if the *Crawford* framework were abandoned, statements contained in big data transfers would still require confrontation under alternatively proposed interpretations of the Confrontation Clause. For example, under the alternative regime proposed by Professor George Fisher, the testimonial statements discussed in this Note are unlikely to fall within the group of "rather rare instances when hearsay may be admitted without cross-examination."[204] Outside of Dean John Wigmore's famously narrow view that has never been adopted by the Supreme Court,[205] many scholars find *Crawford* and its progeny too narrow.[206] Broadening the right of confrontation to nontestimonial statements would only encircle the statements in big data described in this Note more comfortably. At the other end of the spectrum, if the Court were to abandon *Crawford* and return to a *Roberts* reliability regime, there does not appear to be a majority of Justices who would remove the types of technical analysis discussed in this Note from the requirements of confrontation.[207]

[203] Cf. Fisher, The Next Ten, supra note 16, at 15–16 (calling attention to the ongoing development of the Fourth Amendment's "reasonable expectation of privacy" and "reasonable person would feel free to leave" case law).

[204] Fisher, Debacle, supra note 17, at 30; see also id. (excepting from cross-examination, inter alia, reports by lab technicians "ignorant of the results prosecutors desire").

[205] Dean Wigmore's view restricted the right of confrontation to live prosecution witnesses at trial. Bellin, Shrinking, supra note 17, at 1872 (citing John Henry Wigmore, A Treatise on the System of Evidence in Trials at Common Law § 1397, at 1755 (1st ed. 1904)).

[206] See, e.g., Bellin, Shrinking, supra note 17, at 1877–78; Randolph N. Jonakait, "Witness" in the Confrontation Clause: *Crawford v. Washington*, Noah Webster, and Compulsory Process, 79 Temp. L. Rev. 155, 164 (2006); Tom Lininger, Yes, Virginia, There Is a Confrontation Clause, 71 Brook. L. Rev. 401, 405–06 (2005); Mosteller, supra note 17, at 709–12; Deborah Tuerkheimer, A Relational Approach to the Right of Confrontation and Its Loss, 15 J.L. & Pol'y 725, 727 (2007).

[207] See Williams v. Illinois, 132 S. Ct. 2221, 2275 (2012) (Kagan, J., dissenting) ("Scientific testing is . . . only as reliable as the people who perform it."); see also id. at 2255 (Thomas, J., concurring in the judgment) (finding a statement nontestimonial because of its lack of formality and solemnity). Justice Thomas makes five Justices. See supra Section III.A, for a discussion on the solemnity of big data transfers.

2. Recent Supreme Court Indicators

Last year, in *Riley v. California*,[208] the Supreme Court unanimously held that the Constitution recognizes the difference between data stored digitally in cell phones and data stored in traditional objects such as a diary or a photo album.[209] The Court stated that "a cell phone collects in one place many distinct types of information . . . that reveal much more in combination than any isolated record."[210] The Court also noted that smartphone applications, when considered in aggregate, "can form a revealing montage of the user's life."[211] Perhaps most relevant to the Confrontation Clause, the Court stated that data collected in a cell phone "can provide valuable incriminating information about dangerous criminals."[212]

The Court's unanimous opinion in *Riley* was consistent with the views of at least five Justices in *United States v. Jones*,[213] where the Court held that the warrantless use of a tracking device on the defendant's car violated the Fourth Amendment.[214] Concurring in *Jones*, Justice Sotomayor explained how location data could be used in aggregate not only to reveal where the person had been, but also to generate a record reflecting the individual's familial, political, professional, religious, and sexual associations.[215] Justice Alito opened his opinion concurring in the judgment by stating that the case called the Court to apply the Constitution to "21st-century surveillance,"[216] and explained how the GPS tracking abilities of smartphones could be aggregated to create useful information.[217]

Although both *Riley* and *Jones* addressed the Fourth Amendment, if the rationale expressed in those opinions is any indication of how the Court will treat the related criminal procedure rights in the Sixth Amendment, then the big data transfers addressed in this Note are likely to implicate the Confrontation Clause. Perhaps because Professor Jeffrey

[208] 134 S. Ct. 2473 (2014).
[209] Id. at 2490.
[210] Id. at 2489.
[211] Id. at 2490.
[212] Id. at 2493.
[213] 132 S. Ct. 945 (2012).
[214] Id. at 949.
[215] Id. at 955 (Sotomayor, J., concurring).
[216] Id. at 957 (Alito, J., concurring in the judgment).
[217] Id. at 963.

Fisher—who successfully argued *Riley*—also argued *Clark*, *Bullcoming*, *Melendez-Diaz*, and *Crawford*,[218] the parallel between the Court's recent Fourth Amendment doctrine and the Confrontation Clause is most notable in the essential rationale of *Riley*. The fundamental question in *Riley* was "whether the police may, without a warrant, search digital information on a cell phone seized from an individual who has been arrested."[219] Distinguishing the physical object of a cell phone, which officers remain free to examine even after *Riley*, the Court held that when it comes to data within the physical object, "officers must generally secure a warrant."[220] The Court came to this holding because digital data "differ[s] in both a quantitative and a qualitative sense from other objects that might be kept on an arrestee's person."[221]

3. Beneficial Policy Reason

Information leaked by former government contractor Edward Snowden renewed a debate regarding the constitutionality of modern government surveillance.[222] Much of the current literature discussing these programs focuses on the President's authority under Article II of the Constitution, and the limitations of the Fourth Amendment.[223] It appears that at least one Supreme Court Justice would agree that addressing these issues with the Fourth Amendment alone is too rigid a strategy.[224] Using the Confrontation Clause as a check on government surveillance provides much of the benefit that relying on the Fourth Amendment provides—namely, the protection of individual privacy[225]—at much less

[218] Resume of Jeffrey L. Fisher, Stanford Law School, law.stanford.edu/wp-content/uploads/2015/06/JFisher-2015-June-resume.pdf (last visited Aug. 27, 2015).

[219] *Riley*, 134 S. Ct. at 2480 (2014).

[220] Id. at 2485.

[221] Id. at 2489.

[222] See, e.g., sources cited supra note 12; Napolitano, supra note 12, at 538–52; Laura K. Donohue, NSA Surveillance May Be Legal—But It's Unconstitutional, Wash. Post (June 21, 2013), http://www.washingtonpost.com/opinions/nsa-surveillance-may-be-legal—but-its-unconstitutional/2013/06/21/b9ddec20-d44d-11e2-a73e-826d299ff459_story.html.

[223] See Vladeck, After Snowden, supra note 12, at 335 (arguing that the debate has shifted away from preventing front-end collection).

[224] See *Riley*, 134 S. Ct. at 2497–98 (Alito, J., concurring in part and concurring in the judgment) (referring to the "blunt instrument of the Fourth Amendment").

[225] See Katz v. United States, 389 U.S. 347, 350–53 (1967); Donohue, supra note 12, at 774–76.

of a cost in terms of security. The reduced cost stems from the minimal burdens of such a check on investigators in the field.[226]

An investigator can acquire as many testimonial statements she deems necessary to prevent crime without implicating the Confrontation Clause.[227] It is only once the evidence is presented at court that the Confrontation Clause is implicated.[228] Though an investigator may decide not to collect evidence she knows will be inadmissible at trial, the primary purpose test alleviates some of this concern as it allows an investigator to collect evidence more freely if the purpose is to meet an ongoing emergency.[229] Additionally, such a hypothetical cost borne by investigators would seem to be much less than the cost created by a rigid Fourth Amendment requirement, under which investigators must delay investigations to seek a warrant.[230] Recognizing the ability of the Confrontation Clause to address big data transfers, rather than relying on a rigid Fourth Amendment doctrine, allows for this beneficial difference in cost to be captured.

IV. WHAT THE CONFRONTATION CLAUSE REQUIRES FROM GOOGLE

This Part responds to the concern Justice Breyer expressed in his concurrence in *Williams*, which called on the Court to address "the outer limits of the 'testimonial statements' rule set forth in *Crawford v. Washington*."[231] This Note will use Google's procedures for dealing with government data requests as a representative example of other companies similar to Google, such as Microsoft, Yahoo!, and Facebook.[232] This

[226] See Brief for Respondent at 31, Ohio v. Clark, 135 S. Ct. 2173 (2015) (No. 13-1352), 2015 WL 106919, at *31 ("The exclusionary rules this Court has created to enforce the Fourth Amendment . . . are all designed to regulate police conduct . . . . 'The Confrontation Clause,' by contrast, 'in no way governs police conduct.'" (quoting Davis v. Washington, 547 U.S. 813, 832 n.6 (2006))).

[227] See id.; Fisher, The Next Ten, supra note 16, at 12–13.

[228] See U.S. Const. amend. VI. (referring to "criminal prosecutions"); see also George Fisher, Evidence 676 (Robert C. Clark et al. eds., 3d ed. 2013) ("[W]e encounter the Confrontation Clause . . . only if the government offers hearsay evidence against a criminal defendant.").

[229] See *Davis*, 547 U.S. at 822 (2006).

[230] See William J. Stuntz, Warrants and Fourth Amendment Remedies, 77 Va. L. Rev. 881, 885–89 (1991).

[231] Williams v. Illinois, 132 S. Ct. 2221, 2244 (2012) (Breyer, J., concurring).

[232] See Pasquale, supra note 40, at 112 (recognizing Google as a "de facto lawmaker for many aspects of life on the Internet").

Part's analysis of Google's procedures under this system demonstrates the small number of witnesses required for confrontation.

A. Attenuation Standard

Whom must a criminal defendant have the opportunity to confront in order to satisfy the Confrontation Clause as applied to big data transfers? This Note proposes the following answer: an individual who provides a "stand-alone" testimonial statement.

"Stand-alone" in this context means a testimonial statement that, by itself, is capable of proving a fact at a criminal trial. This definition uses an "attenuation standard" to determine who has provided a stand-alone testimonial statement, and excludes individuals who either provided no testimonial statement, or who merely played a role in crafting someone else's testimonial statement.

B. Google as an Example

Instead of examining the attenuation standard in the abstract, it is useful to examine how it applies to the procedures Google uses in responding to government requests for data. Google has stated that "[w]hen we receive such a request, our team reviews the request to make sure it satisfies legal requirements and Google's policies. . . . If we believe a request is overly broad, we'll seek to narrow it."[233]

It may be easier to first examine who does *not* provide a testimonial statement. On the outskirts of the attenuation spectrum is a Google employee such as a janitor. A janitor ensures that the office is habitable, which is arguably a prerequisite for other Google employees to prepare the data requested by the government. It seems obvious that the janitor would not implicate the Confrontation Clause. This is because the janitor merely helps to shape someone else's testimonial statement, and does not provide a stand-alone testimonial statement himself.

The results of the attenuation standard are less obvious when applied to the individuals who create and manage the algorithms Google uses to collect data.[234] It is unlikely that these individuals design or manage al-

[233] Transparency Report, Google, https://www.google.com/transparencyreport/userdatarequests/legalprocess (last visited Aug. 27, 2015).

[234] See How Search Works: Algorithms, Google, http://www.google.com/insidesearch/howsearchworks/algorithms.html (last visited Aug. 27, 2015) (explaining algorithms).

gorithms with the intent of providing a testimonial statement.[235] Companies exist, however, that are in the business of aggregating data and selling it to government agencies to assist in criminal investigations.[236] In the case of such companies, it seems more likely that the company's algorithms and procedures are designed and managed with the intent to provide testimonial statements. In fact, employees at these companies have the potential to be very useful witnesses because they can learn what type of testimony is usually admitted or denied and can adapt their algorithms and procedures accordingly.[237] Employees at these companies implicate the Confrontation Clause if they design or manage algorithms with the primary purpose of providing evidence for trial, and those employees would be required to be available for confrontation. Putting those unique companies aside and focusing on Google, the algorithm designers and managers at Google likely originally collect and store data for marketing purposes,[238] and therefore those Google employees do not implicate the Confrontation Clause.

Less attenuated than either the janitor or the employees responsible for designing or managing algorithms is a "screener." A Google screener sorts and prioritizes government requests for data.[239] Although it is tempting to label the screener as providing a testimonial statement, the tasks performed by the screener are too attenuated to require confrontation. By simply screening and prioritizing requests that will be responded to by other employees, the screener, like the janitor, is merely helping to shape someone else's testimonial statement. The fact that the screener is not required for confrontation makes clear that the line separating those employees who must be available for confrontation from those employees who do not cannot always be drawn at the moment the government requests data from a company. It is tempting to draw the line at

[235] See supra text accompanying note 105.

[236] See Chris Jay Hoofnagle, Big Brother's Little Helpers: How ChoicePoint and Other Commercial Data Brokers Collect and Package Your Data for Law Enforcement, 29 N.C. J. Int'l L. & Com. Reg. 595, 595 (2004).

[237] Cf. Genevieve Grant & David M. Studdert, The Injury Brokers: An Empirical Profile of Medical Expert Witnesses in Personal Injury Litigation, 36 Melb. U. L. Rev., 831, 863–64 (2013) (referring to "tried and tested" witnesses).

[238] See Joe Mullin, How Much Do Google and Facebook Profit from Your Data?, Ars Technica (Oct. 9, 2012, 9:38 AM), http://arstechnica.com/tech-policy/2012/10/how-much-do-google-and-facebook-profit-from-your-data.

[239] Google, Way of a Warrant, supra note 101.

that point, because it is then that the company is aware that they will be providing a testimonial statement to the government. If someone tells the janitor, however, that the government has requested data, and the janitor continues to do his job with newfound motivation, he is not now magically transformed into providing a testimonial statement.

It is also tempting to conclude that because the screener "sorts and prioritizes" government requests for information,[240] the screener is therefore determining whether there is an ongoing emergency under the primary purpose test. The primary purpose test, however, is an objective test.[241] The subjective intentions of the screener are therefore not determinative.[242] Even if a Google screener had the authority to subjectively determine what the Sixth Amendment requires, the screener is not necessarily attempting to do so by prioritizing requests. A number of factors unrelated to the emergency status of the statement could go into the screener's ranking, such as which requests would be easier to respond to.

The line designating which Google employees must be available for confrontation begins with the "producer." The producer "determin[es] what information to provide" the government in order to comply with the request.[243] Two aspects of the producer's duty indicate that she is providing a testimonial statement.

First, the producer provides a testimonial statement by determining what data is relevant to the government request, and then producing that data.[244] Here a comparison to the Fifth Amendment's "act of production" analysis is helpful.[245] In *United States v. Hubbell*,[246] the Supreme Court distinguished between the *contents* within material that was produced in response to a government request and the actual *act* of producing the material.[247] Referring to the *act* of producing the material, the Court stated that "[t]he assembly of literally hundreds of pages of material in response to a [government] request . . . is the functional equiva-

[240] Id.
[241] Michigan v. Bryant, 131 S. Ct. 1143, 1156 (2011).
[242] See id.
[243] Google, Way of a Warrant, supra note 101.
[244] Id.
[245] I thank Professor George Fisher of Stanford Law School for bringing this comparison to my attention.
[246] 530 U.S. 27 (2000).
[247] Id. at 41–42.

lent of the preparation of an answer to either a detailed written interrogatory or a series of oral questions . . . ."[248] Just as in *Hubbell*, a Google producer is similarly providing a statement to government investigators when she produces a collection of data. This statement is analytically distinct from the content of the data.[249] The content of the data that the Google producer decides to provide, however, is also significant. As explained in Subsection III.A.1, what a collection of data does not include can be just as important as what it does include. Again, adding "how to write a fictional novel about deadly lovers" to an otherwise incriminating collection of data can paint an entirely different picture than would otherwise be presented.

A second aspect of the producer's duty that indicates that she is providing a testimonial statement results from the solemnity of her statement. When deciding what data to include in a reply to the government, the producer likely considers the legal implications of the data provided.[250] By determining what data is required to comply with a government request, where noncompliance can result in being held in contempt of court,[251] the producer would seem to meet the formality and solemnity requirement sought by Justice Thomas.[252] The producer "gather[s] the information, carefully and accurately."[253] This is evidence that the producer is providing a solemn declaration of fact, with due care taken to ensure that the data is accurate.

After the producer decides what data to provide to the government, the data is then "sent to [i]nvestigators along with a [c]ertificate of [a]uthenticity."[254] This tracks almost exactly with *Melendez-Diaz*, where the lab report was sworn before a notary, and *Bullcoming*, where the report was accompanied with a certificate of analysis.[255] The analyst who

[248] Id.

[249] See supra Subsection III.A.4.

[250] See Long, supra note 98 (stating that Google has "lawyers review" requests).

[251] See 18 U.S.C. § 3511(c) (2012); 50 U.S.C. § 1881a(h)(4)(G), (h)(5)(D) (2012).

[252] See Williams v. Illinois, 132 S. Ct. 2221, 2255 (2012) (Thomas, J., concurring in the judgment).

[253] Google, Way of a Warrant, supra note 101.

[254] See id.; see also United States v. Camez, No. 2:12-cr-0004-APG-GWF, 2013 WL 6158402, at *1–2 (D. Nev. Nov. 21, 2013) ("When the Government subpoenaed Google's custodian of records ('COR') to testify at trial, Google responded with a single-page, five-paragraph 'Certificate of Authenticity' from the COR . . . .").

[255] See Bullcoming v. New Mexico, 131 S. Ct. 2705, 2717 (2011); Melendez-Diaz v. Massachusetts, 557 U.S. 305, 308 (2009).

prepares this certificate, just like the swearing and certifying analysts in *Melendez-Diaz* and *Bullcoming*, is certifying that the data was collected and transmitted correctly. The specific analyst who prepared the certificate must therefore be available for confrontation under the *Bullcoming* framework.[256]

Lastly, Google designates a "custodian of records" who is able to travel and appear in court to verify the data.[257] Because the custodian also certifies the validity of the data,[258] the defendant must also have the ability to confront her. Although Google may want to streamline the process and have a single custodian handle all trial proceedings, the custodian alone is not enough to satisfy the Confrontation Clause. While the custodian can speak to her own certification of the data and may be able to "vouch" for the reliability of the testimonial statements provided by the producer and certifying analyst, *Bullcoming* requires the availability of the *specific* individuals responsible for providing testimonial statements.[259] Because both the certifying analyst and the producer supply testimonial statements, both must be available for confrontation despite any additional testimony that the custodian may provide.

CONCLUSION

The Confrontation Clause of the Sixth Amendment places an important check on the use of big data obtained through government surveillance. The clause is well suited to address the issues of widespread government surveillance, because its flexibility allows data to be used to prevent emergencies, while simultaneously ensuring that incriminating data is handled in accordance with the constitutional safeguards that the Founders provided over two hundred years ago. Though big data was perhaps unimaginable at the time of the Framing, the probativeness of a testimonial statement was well understood.[260]

When companies transfer data to government investigators under the procedures detailed above, they provide testimonial statements. In order to protect the accused from unscrupulous witnesses and human mistakes, the Constitution requires the Sixth Amendment to play a role in the transfer of big data into the hands of government investigators.

---

[256] See *Bullcoming*, 131 S. Ct. at 2716.

[257] See Google, Way of a Warrant, supra note 101; see also *Camez*, 2013 WL 6158402, at *1 (requiring Google's custodian of records to authenticate certain records with live testimony).

[258] See Google, Way of a Warrant, supra note 101.

[259] See *Bullcoming*, 131 S. Ct. at 2716.

[260] Crawford v. Washington, 541 U.S. 36, 53–56, 59 (2004).