DATA SCANNING AND THE FOURTH AMENDMENT

ORIN S. KERR

DATA SCANNING AND THE FOURTH AMENDMENT

ORIN S. KERR*

Abstract: An important question of Fourth Amendment law has recently divided courts: When government agents conduct a digital scan through a massive database, how much of a "search" occurs? The issue comes up in diverse contexts including geofence warrants, reverse keyword searches, tower dumps, and Internet pen registers. When a government agent runs a filter through a massive database, resulting in a list of hits, is the scale of any Fourth Amendment search determined by the size of the database, the filter setting, or the filter output? Fourth Amendment law is closely attuned to the scale of a search. No search means no oversight, small searches ordinarily require warrants, and limitless searches are categorically unconstitutional. But how broad is a data scan?

This Article argues that Fourth Amendment implications of data scans should be measured primarily by filter settings. Whether a search occurs, and how far it extends, should be based on what information is exposed to human observation. This standard demands a contextual analysis of what the output reveals about the dataset based on the filter setting. The proper question is what information is expressly or implicitly exposed, not how much raw data passes through the filter or what the raw data output says expressly. The implications of this approach are then evaluated for a range of important applications, among them geofence warrants, reverse keyword searches, tower dumps, and artificial intelligence queries.

INTRODUCTION

Imagine the police need to search through a massive database to find evidence of crime. The database contains information relating to millions of users. To find the evidence, the police set a filter that scans through the database and collects evidence responsive to the filter.

Now consider a legal question: When this scanning occurs, how much of a Fourth Amendment search has occurred?[1] From a constitutional perspective, is this a massive database search through the entire database, the kind of broad

* Professor, Stanford Law School. Thanks to David Sklansky, Mark Quevedo, Steven Bellovin, Robert Graham, Christina Koningisor, Ron Lee, Catherine Crump, Michelle Mello, Diego Zambrano, Bob Weisberg, Pam Karlan, Bernie Meyler, Lisa Ouellette, and attendees at the Privacy Law Scholars Conference and the Stanford Law School internal workshop for comments on an earlier draft.

[1] *See* U.S. CONST. amend. IV ("The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated . . . .").

search that the Fourth Amendment ordinarily prohibits even with a warrant?[2] Alternatively, is it a narrow search, with the scope defined by the filter setting, of the kind that the Fourth Amendment allows if a valid warrant was obtained?[3] Or is it no search at all, triggering no Fourth Amendment oversight and no need for a warrant, unless the responsive evidence is itself particularly sensitive?[4]

This range of answers seems plausible because our life experience gives us no obvious way to model digital scanning. Our experience with searches is mostly physical. When officers search a room for evidence, they move physical things to bring new items into their line of sight. Searching is exposure, enabling human senses to observe and human minds to recognize.[5] But how do we apply those insights to computer scanning? Put another way, what is the searching "eye" of a digital scan? For Fourth Amendment purposes, is the eye the scanning itself, the positive hit, or the output? And this leaves us with three plausible ways to model the scope of data-scan searches. The search might be measured by the database scanned, the filter setting, or the output that is produced.

You might think that there must be a settled answer to these questions. Until recently, however, the issue rarely arose. Lower courts routinely held that the types of data typically found in large databases were not subject to any Fourth Amendment protection.[6] That made the law of data scanning easy. When information inside a database is categorically outside Fourth Amendment protection, after all, querying a database to extract information is not going to be a search, either.

In the last few years, however, several lower courts have embraced new and expansive understandings of what data within databases retains constitu-

---

[2] *See generally* Andresen v. Maryland, 427 U.S. 463, 480 (1976) ("General warrants of course, are prohibited by the Fourth Amendment. [T]he problem [posed by the general warrant] is not that of intrusion per se, but of a general, exploratory rummaging in a person's belongings . . . ." (alterations in original) (internal quotations omitted) (quoting Coolidge v. New Hampshire, 403 U.S. 443, 467 (1971))).

[3] *See id.* ("[The Fourth Amendment addresses the problem] by requiring a 'particular description' of the things to be seized." (alteration in original) (quoting *Coolidge*, 403 U.S. at 480)).

[4] *See* Smith v. Maryland, 442 U.S. 735, 745–46 (1979) (where no search or seizure was found, there was no Fourth Amendment oversight).

[5] *See infra* Section III.A.

[6] *See, e.g.*, State v. Sloane, 939 A.2d 796, 803 (N.J. 2008) (concluding that a database query of arrests and convictions was not a search because the information in the database was public); United States v. Rhodes, No. 19-CR-73-AT, 2021 WL 1541050, at *2 (N.D. Ga. Apr. 20, 2021) (holding that a tower dump is not a search); United States v. Walker, No. 2:18-CR-37-FL-1, 2020 WL 4065980, at *8 (E.D.N.C. July 20, 2020) (same). A relatively early exception is the Second Circuit's ruling in *United States v. Hasbajrami*, on accessing a database of communications collected and stored under Section 702 of the Foreign Intelligence Surveillance Act. 945 F.3d 641, 647 (2d Cir. 2019). Notably, however, *Hasbajrami* reached no conclusion as to what particular Fourth Amendment standards should apply. *See id.* at 670–72.

tional protection.[7] Courts have held that compelled access to database records of Google search terms,[8] geofencing records,[9] and tower dumps[10] are covered by the Fourth Amendment. It's too early to tell if the new theories will earn broad acceptance.[11] But they prompt a follow-up question: If at least some data in a database is protected, how broad a Fourth Amendment search occurs when government agents direct a scan through it?[12] Although legal scholarship has not yet focused on this question, lower courts have already divided on it.

This Article analyzes the Fourth Amendment implications of data scans. It has three goals. The first goal is to explain how data scanning works and to see why it raises difficult questions for Fourth Amendment law. Appreciating the basics of scanning technology helps explain why there are different constitutional models of scanning. A little bit of technology goes a long way, and it is easy (even for lawyers) to learn. And along the way, you will come to learn common fact patterns in which data scanning comes up, such as geofencing, reverse keyword searches, and Internet pen registers.[13]

The Article's second goal is to expose the current divide in appellate cases on data scanning. In the last two years, a sharp division has emerged. The U.S. Court of Appeals for the Fifth Circuit and two federal trial courts have adopted a standard based on the amount of data scanned, a rule that makes most scanning of massive databases categorically unconstitutional.[14] The Colorado Supreme Court, the Supreme Court of Georgia, and most of the judges on the en banc U.S. Court of Appeals for the Fourth Circuit have focused instead on the

---

[7] *See infra* notes 8–11 and accompanying text (describing lower court decisions); *infra* Part II.

[8] *See, e.g.*, People v. Seymour, 536 P.3d 1260, 1273 (Colo. 2023) (concluding that access to search terms is a Fourth Amendment seizure).

[9] *See, e.g.*, United States v. Smith, 110 F.4th 817, 820 (5th Cir. 2024) (concluding that geofencing is a search), *cert. denied*, No. 24-7237, 2025 WL 3131804 (Nov. 10, 2025).

[10] *See, e.g.*, United States v. Spurlock, 778 F. Supp. 3d 1136, 1145 (D. Nev. 2025) (concluding that tower dumps are searches).

[11] There are cases on both sides of each of these questions. Which views are correct is beyond the scope of this Article, but I have addressed several of them in Chapter 9 of *The Digital Fourth Amendment: Privacy and Policing in Our Online World*. ORIN S. KERR, THE DIGITAL FOURTH AMENDMENT: PRIVACY AND POLICING IN OUR ONLINE WORLD 152–70 (2025).

[12] For purposes of answering this question, it should not matter if the government possesses the database and scans it directly or if government agents direct a private actor to scan a database in the private actor's possession. In both cases, the action is attributed to the state and is covered by the Fourth Amendment. For more on the possible relevance of state action to database scans, see *infra* note 229.

[13] *See infra* Part II (describing different approaches to modeling scanning).

[14] *See Smith*, 110 F.4th at 820 (adopting the data-focused approach and concluding that geofence warrants are categorically unconstitutional); *In re* Four Applications for Search Warrants Seeking Info. Associated with Particular Cellular Towers, No. 25-cr-38-CWR, 2025 WL 603000, at *1 (S.D. Miss. Feb. 21, 2025) [hereinafter *In re* Four Applications for Search Warrants] (extending *Smith* to hold that tower dumps violate the Fourth Amendment); *Spurlock*, 778 F. Supp. 3d at 1145 (agreeing with the reasoning of *Smith* and holding that tower dumps are searches).

filter setting, which allows scanning of databases with a warrant when the filter is appropriately tailored.[15]

The conflict among courts is not just open, but defiant. The Fifth Circuit recently decried the reasoning of the other side of the debate as "breathtaking."[16] Five Fourth Circuit judges soon volleyed back that the Fifth Circuit's view was "nonsensical."[17] In the short run, the division gives other lower courts a menu of options to choose from. In the long run, it sets up eventual review by the U.S. Supreme Court.[18]

The third goal of this Article is to argue for a filter-focused approach to data scanning. When the government uses a filter to scan through data, the scope of any "search" should be determined by what the output reveals in light of the filter setting. The important question is not how much data was searched through, but what facts about that data were explicitly or implicitly revealed. This approach allows for searches through massive databases as long as the filter is appropriately narrowed. This approach is consistent with Fourth Amendment theory and doctrine, avoids arbitrary outcomes, and creates a clear standard by which courts can identify the Fourth Amendment implications of data scans.

In short, courts should reject the Fifth Circuit's approach and adopt a standard resembling that of the Colorado Supreme Court and the Georgia Supreme Court. Data scanning through massive databases is a modern reality. Fourth Amendment law should focus on the substance of what scanning reveals over the form of what data is scanned through or what the output file shows. Under this approach, access to records through geofence warrants is constitutional—assuming the types of records sought are themselves protected—so long as the filters used are appropriately narrow. Searches through the contents of large numbers of accounts should nonetheless remain generally prohibited. The legality of reverse keyword searches remains uncertain, and the answer may depend on the category of search query.[19]

---

[15] *See* People v. Seymour, 536 P.3d 1260, 1268 (Colo. 2023) (adopting the filter-focused approach and concluding that a reverse keyword warrant was constitutional as long as the filter parameters are narrow); Jones v. State, 913 S.E.2d 700, 710 (Ga. 2025) (endorsing the filter-focused approach, but ultimately not reaching the question); United States v. Chatrie, 136 F.4th 100, 101 (4th Cir. 2025) (en banc) (Berner, J., concurring, joined by Gregory, Wynn, Thacker & Benjamin, JJ.) (adopting the provider-as-filter approach and finding that no search occurs when the government gets geofence information in anonymized form), *cert. granted*, No. 25-112, 2026 WL 120676 (U.S. Jan. 16, 2026).

[16] *Smith*, 110 F.4th at 838 n.12.

[17] *See Chatrie*, 136 F.4th at 155 (Berner, J., concurring) ("Though the Fifth Circuit refers to this proposition as 'breathtaking,' any other approach would be nonsensical." (citation omitted)).

[18] Supreme Court review may happen in the short run, too. As this Article was going to print, the Supreme Court agreed to review the Fourth Circuit's ruling in *United States v. Chatrie*, 136 F.4th 100 (4th Cir. 2025) (en banc), *cert. granted*, No. 25-112, 2026 WL 120676 (U.S. Jan. 16, 2026). *Chatrie* may provide the Court with an opportunity to address the questions raised in this Article.

[19] *See infra* notes 91–109 and accompanying text (defining and discussing reverse keyword searches).

The Article proceeds in four parts. Part I covers technology.[20] Part II explores the cases on the different approaches to data scanning.[21] Part III makes the affirmative case for the filter-focused understanding.[22] Part IV applies the approach to common fact patterns that courts are confronting.[23] It also speculates about how the approach may or may not apply to searches enabled by artificial intelligence.

I. The Basics of Scanning Technology

It is helpful to start with the relevant technology. This Part considers what data scanning is, how it works, and how technologists might design databases to be queried. It then considers why that technology raises difficult Fourth Amendment questions.

A. A Simple Search

Imagine you have a Microsoft Word file of a draft law review article. You want to find every time the article uses a particular word, such as "Scalia." Microsoft Word makes it easy with its find function.[24] You click on the find function, enter the word "Scalia," and then additional clicks bring you to each instance in which the word Scalia appears.

This common example is known in the computer science literature as a linear scan of unstructured data.[25] The data is unstructured in the sense that the word you are looking for could be anywhere in the article. To find examples of the word inside the article, the find function just runs through each piece of data in the file and checks whether a particular segment matches the word.[26] The word is known as a "keyword" or a "pattern" that is used to find a match.[27]

A linear scan of unstructured data is a simple kind of search. At a technological level, the computer calls up a string of data and checks for a pattern. Computers use binary, not regular letters. Using what is known as ASCII code, which is used to translate between certain characters and binary numbers, the letter S translates into 01010011. The complete name "Scalia" would be repre-

---

[20] *See infra* notes 24–47 and accompanying text.

[21] *See infra* notes 48–149 and accompanying text.

[22] *See infra* notes 150–207 and accompanying text.

[23] *See infra* notes 208–229 accompanying text.

[24] *See Find and Replace Text in Word*, Microsoft, https://support.microsoft.com/en-us/office/find-and-replace-text-c6728c16-469e-43cd-afe4-7708c6c779b7 [https://perma.cc/QQ3C-U4Y5].

[25] *See* Christopher D. Manning, Prabhakar Raghavan & Hinrich Schütze, Introduction to Information Retrieval 3 (2008) ("The simplest form of document retrieval is for a computer to do this sort of linear scan through documents. This process is commonly referred to as *grepping* through text, after the Unix command grep, which performs this process.").

[26] *See id.*

[27] *See* Donald E. Knuth, James H. Morris, Jr. & Vaughan R. Pratt, *Fast Pattern Matching in Strings*, 6 SIAM J. Comput. 323, 323–50 (1977).

sented by 01010011011000110110000101101100 0110100101100001.[28] With a linear scan, the software can run through the entire law review article (which it understands as a massively long string of zeros and ones), checking the string of data for a match with the pattern 01010011011000110110000101101100 0110100101100001—that is, with the keyword "Scalia."

The keyword of a search acts as a filter. Information that passes through the filter but does not yield a match does not register any output. For example, if an article does not contain the word "Scalia," the search for that keyword would yield nothing. The meaning of a null result—that is, no hits—means that the exact pattern did not appear in the data reviewed. Of course, that gives no insight into what other information was contained there.

It is helpful to add a word on speed and scale. To a human being, the name "Scalia" translated into binary seems like a long string of data. To our human eyes, it comes off as so many zeros and ones that our eyes glaze over it. To a computer, however, this is no challenge at all. Matching patterns is what computer processors do. And to a computer, that one line of zeros and ones is a very tiny segment of data. Modern laptop and desktop computers can process *trillions* of calculations in a *single second*.[29] That (humanly) unimaginable scale allows computers to scan through incredibly large databases for very long strings of data at truly breathtaking speed. It's amazing to us, but it's old hat to a computer. It is just the normal and boring operation of the computer to run through the data and register when that exact pattern occurs.

B. A Search Through Structured Data

With a search through unstructured data, like a law review article, the data searched through is essentially like random things in a big box. In most applications, however, data is structured. Structured data refers to data where different parts of the data are linked in a particular way. A common example is a spreadsheet, in which different boxes have preassigned meanings. The information is no longer just jumbled together. Instead, different kinds of information are put in different places.

Think about an email, where a user might experience something like this:

> From: bob@company.com
> To: sally@company.com
> Date: January 4, 2025, 9:00am
> Going to the meeting?

---

[28] For an ASCII conversion chart, see *ASCII Character Chart with Decimal, Binary and Hexadecimal Conversions*, ESO, https://www.eso.org/~ndelmott/ascii.html [https://perma.cc/N74F-JNWU].

[29] Apple's current M4 chip, widely available in its laptop computers, is capable of thirty-eight trillion calculations in a second. *See Apple Introduces M4 Chip*, APPLE (May 7, 2024), https://www.apple.com/newsroom/2024/05/apple-introduces-m4-chip [https://perma.cc/S4QR-Y6VJ].

This email contains four linked parts of information. There's the *from* email address; the *to* email address; the *time* the email arrived; and the actual contents of the message. They all go together, in that they relate to the same unique message. It is easy to imagine this data together as entries in a single row of a spreadsheet like this:

To	From	Date/Time	Message
Bob	Sally	Jan 4, 9:00am	Going to the meeting?

Of course, databases won't contain just one email. Imagine a company that provides a messaging service to all of its employees. The company stores all of the messages in a single database. The database could have millions of messages. A very tiny excerpt of just a handful of messages might look like this:

To	From	Date/Time	Message
Bob	Sally	Jan 4, 9:00am	Going to the meeting?
Sally	Charles	Jan 4, 9:02am	Sup? Free 4 lunch?
Alan	Sally	Jan 4, 9:03am	Went to a great concern last week.
Alan	Sally	Jan 4, 9:05am	Hey
Charles	Bob	Jan 4, 9:06am	Staff meeting at 11am
Charles	Alan	Jan 4, 9:06am	Staff meeting at 11am
Charles	Bob	Jan 4, 9:07am	Follow up re sales

If the company wanted to search through this large database—here, imagine it has millions of messages—how could it do that? There are a few ways. At the simplest level, it could do the same linear search through the entire database that we saw with the law review article. For example, if the company wanted to collect every email that said the word "meeting," it could simply search through the entire database for that word—much like we saw with the search through the law review article for "Scalia." It could then copy whatever relevant information was associated with the term, such as, say, the entire email entry, and send that to an output file.[30]

At this point, however, we need to start considering efficiency. Remember that the file of messages could have millions of entries. Although computers can calculate incredibly quickly, the process of scanning requires additional steps to work. To scan messages, computers must first fetch them by portions

[30] *See generally* Edgar F. Codd, *A Relational Model of Data for Large Shared Data Banks,* 13 COMMC'NS ACM 377, 377–87 (1970) (discussing the limitations of exact-match keyword searches, including the possibility of null results).

from stable storage and then place them into the temporary memory of a computer. This is something like cooking a meal, where you might need to get the ingredients out of the freezer first and thaw them before preparing items to be eaten. As a technical matter, the process of data retrieval is much slower than the scanning itself.[31] This means that having to scan through a lot of entries can take some time, even if the scanning itself is extremely quick.

To speed things up, we might consider shortcuts. In this case, for example, the word we are looking for is only in the body of the message. If we can isolate the *message* column in the spreadsheet above, and look only through that particular column, we won't need to waste time looking through the *to*, *from*, or *date/time* columns. The querying can be much faster if we limit how much data has to be retrieved from storage and then scanned to carry out the search.

C. The Role of Indexes

The usefulness of speeding up searches brings us to the role of indexes. Lawyers are familiar with indexes in a book. A reader might check out the index in the back of a book to see where a particular topic is covered. In the database setting, indexes work in a similar way. An index is a list of data entries taken from larger data elements together with the locations at which the corresponding larger elements can be found.[32] Indexes can limit the scale of a scan. Instead of requiring a scan through an entire database, a query can merely consult the index and be pointed to a specific part of the database.

Consider the massive database of employee messages above. The company could create an index with entries for each employee's inbox. To scan through messages, the company could go to the index and follow it to the messages for that particular employee. It could look like this:

Alan	*3, 4*
Bob	*1*
Charles	*5, 6, 7*
Sally	*2*

[31] *See* RANDAL E. BRYANT & DAVID R. O'HALLARON, COMPUTER SYSTEMS: A PROGRAMMER'S PERSPECTIVE 647–48 (3d ed. 2016).

[32] *See* Codecademy Team, *What Is a Database Index?*, CODECADEMY, https://www.codecademy.com/article/sql-indexes [https://perma.cc/A3WF-P3YL] ("Indexes are a powerful tool used in the background of a database to speed up querying. Indexes power queries by providing a method to quickly lookup the requested data. Simply put, an index is a pointer to data in a table."). The entries can also represent data, rather than being the data itself, as is the case with data represented by a hash. *See* Codecademy Team, *What Is Hashing, and How Does It Work?*, CODECADEMY (Mar. 27, 2025), https://www.codecademy.com/resources/blog/what-is-hashing [https://perma.cc/6AV4-JTJS] ("Hashing is the process of converting data—text, numbers, files, or anything, really—into a fixed-length string of letters and numbers.").

If the company wanted to see the messages sent to Charles, for example, the computer could consult the index, spot the entry for Charles, and then fetch just the portions of the database (rows 5, 6, and 7) which contain those messages. That would speed up the querying by narrowing it. Instead of having to scan through every employee's messages, the company could limit the scan to those messages in Charles's inbox.

What kind of index is useful depends largely on what kind of scan is needed. If the company expects to query individual-by-individual, then an index with different entries for different employees makes sense. If the company expects to query some other way, such as year-by-year or month-by-month, then a different index would be preferable. The key idea is that indexing allows scans to be faster by limiting what information the computer scans based on criteria that the designers predict will be useful to whoever is running the query.[33]

The takeaway from this overview is that there are different ways to query information in a large database. The basics are the same. Data scanning always has the three basic components: the data that will be scanned, the filter that is set to scan through the data, and the output resulting when a filter match is found. At the same time, a database administrator determines how information is scanned by making choices about how to arrange the data—either leaving it unstructured, creating structure, or using an index. Those database management decisions determine the size of the data scanned as the filter searches for the designated keyword. What casual users might think of as simply scanning through a database really depends on how the database is structured and the tools available to minimize the extent of the scan.

D. Fourth Amendment Implications

Why does all of this matter for Fourth Amendment law? The Fourth Amendment bans unreasonable searches and seizures.[34] Applying the Fourth Amendment ordinarily begins with asking whether a search or seizure occurred. (To avoid confusion, I will refer to a search in the sense of a query through a database as a query or a scan; I will refer to a search in the constitutional sense as a search or a Fourth Amendment search.) When a search or seizure occurs, the next question is its reasonableness.[35] Searches and seizures are reasonable, in general, if the scanning was conducted by a valid warrant or it fits some other reasonableness framework, such as an exception to the warrant requirement.[36]

---

[33] *See* CODECADEMY, *What Is a Database Index?*, *supra* note 32.

[34] *See* U.S. CONST. amend. IV (stating that "[t]he right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated").

[35] NANCY J. KING, ORIN S. KERR & EVE B. PRIMUS, KAMISAR, LAFAVE, AND ISRAEL'S CRIMINAL PROCEDURE: INVESTIGATION 296 (16th ed. 2023).

[36] *See id.*

How we model scanning through a database can implicate each of these questions. Some kinds of data within a database can be protected by the Fourth Amendment, while other data is not.[37] As a result, how we model scanning might determine whether and when a particular scan using a particular filter is a Fourth Amendment search. Second, how scanning is modeled might help determine whether a particular scan recognized as a Fourth Amendment search is a reasonable or an unreasonable search in the absence of a warrant. Third, it might determine how broadly warrants can be drafted to allow a constitutional search through a database to satisfy the requirement that warrants particularly describe the things to be seized.[38] And fourth, it might determine whether it is even possible to allow a scan, even with a warrant, through a particularly large database. Each of these four questions can raise somewhat different issues. But each hinges on how much of a search happens, constitutionally speaking, when a scan through data is made.

This is an abstract point, so an example might help. Consider the Fourth Amendment's ban on general warrants.[39] At the time of the Fourth Amendment's enactment, general warrants were warrants that allowed an official "to search suspected places without evidence of a fact committed, or to seize any person or persons not named, or whose offense is not particularly described and supported by evidence."[40] The text of the Fourth Amendment prohibits general warrants directly: "[N]o Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized."[41]

The textual limit that warrants must particularly describe the place to be searched has been understood to limit the scope of how much can be searched. The search cannot "go too far."[42] The usual unit, in physical space, is a single

---

[37] For example, stored email is protected. *See* United States v. Warshak, 631 F.3d 266, 286 (6th Cir. 2010) ("If we accept that an email is analogous to a letter or a phone call, it is manifest that agents of the government cannot compel a commercial ISP to turn over the contents of an email without triggering the Fourth Amendment."). On the other hand, Internet protocol addresses are generally unprotected. *See* United States v. Hood, 920 F.3d 87, 90–92 (1st Cir. 2019).

[38] *See* U.S. CONST. amend. IV (requiring that searches pursuant to a warrant particularly describe "the place to be searched, and the persons or things to be seized").

[39] *See* Andresen v. Maryland, 427 U.S. 463, 480 (1976) ("General warrants of course, are prohibited by the Fourth Amendment.").

[40] *See* VA. DECLARATION OF RIGHTS art. X (1776) ("That general warrants, whereby an officer or messenger may be commanded to search suspected places without evidence of a fact committed, or to seize any person or persons not named, or whose offense is not particularly described and supported by evidence, are grievous and oppressive and ought not to be granted.").

[41] U.S. CONST. amend. IV.

[42] *See* Steele v. United States, 267 U.S. 498, 503 (1925) (allowing a warrant to search a garage and "any building or rooms connected or used in connection with the garage, or the basement or subcellar beneath the same" because "the elevator of the garage connected it with every floor and room in the building and was intended to be used with it," and therefore holding that the search did not "go too far").

home. The government can get a warrant to search one home for evidence, such as a residence where someone might be suspected of selling drugs. A warrant to search broader spaces is generally prohibited, however. For example, the government cannot get a warrant to search an entire multi-unit apartment building.[43] And agents clearly cannot get warrants to search every house on a city block, or even every house in a small city. Such a warrant would be exactly the kind of general warrant that the Fourth Amendment was enacted to prohibit.

But how do these principles apply to database queries? One wrinkle is that scanning often involves outsourcing the work to companies.[44] When the government obtains a warrant requiring a company to hand over information, and that information is stored by the company inside a massive database, the company will tend to execute its scans based on however the database was previously configured. For the most part, the details will be outside law enforcement's control—or even awareness. A court order will direct the company to hand over the evidence; it won't direct the company on how to do so.[45] As we will see later in the Article, there are different ways under the Fourth Amendment to treat this common outsourcing of scanning authority to private parties.[46] But assume, for now, that outsourcing to a private party makes no constitutional difference; the private party is a government actor when it executes a warrant on the government's behalf, so it is regulated by the Fourth Amendment just like the government is.[47]

---

[43] *See* People v. Avery, 478 P.2d 310, 312 (Colo. 1970) ("The courts are generally agreed that where a structure is divided into several occupancy units, or is a multi-unit dwelling, and there is no common occupancy of the entire structure by all of the tenants, a search warrant which merely describes or identifies the larger multiple-occupancy structure and not the particular subunit to be searched is insufficient to meet the constitutional requirements of particularity of description." (citing W.C. Crais, III, Annotation, *Search Warrant: Sufficiency of Description of Apartment or Room to Be Searched in Multiple-Occupancy Structure*, 11 A.L.R.3d 1330 (1967)).

[44] I have labeled this "indirect surveillance," as compared to "direct surveillance" by the government. Orin S. Kerr, *Internet Surveillance Law After the USA Patriot Act: The Big Brother That Isn't*, 97 NW. U. L. REV. 607, 621–22 (2003) [hereinafter Kerr, *Big Brother*].

[45] In part, this reflects a traditional limit of search warrants. *See* Dalia v. United States, 441 U.S. 238, 257 (1979) ("[I]t is generally left to the discretion of the executing officers to determine the details of how best to proceed with the performance of a search authorized by warrant—subject of course to the general Fourth Amendment protection 'against unreasonable searches and seizures.'").

[46] I return to this question later in the Article. *See infra* note 229 and accompanying text (discussing the state action doctrine).

[47] *See, e.g.*, Commonwealth v. Gumkowski, 167 N.E.3d 803, 812 (Mass. 2021) (holding that a cell phone provider is a state actor for Fourth Amendment purposes when complying with a court order to disclose user data); *id.* ("[I]f law enforcement instigates the search by contacting the cell phone company to request information, there is State action. That Sprint could have refused to provide records in response to [the state trooper's] request does not change the fact that he instigated the search." (citing Commonwealth v. Brandwein, 760 N.E.2d 724, 731 (Mass. 2002)); *see also* Orin S. Kerr, *The Fourth Amendment Limits of Internet Content Preservation*, 65 ST. LOUIS U. L.J. 753, 779–82 (2021) [hereinafter Kerr, *Fourth Amendment Limits*] (arguing that Internet providers are Fourth Amendment state actors when complying with directives from investigators).

Under that assumption, the legality of a data scanning warrant and its execution—and in some cases, whether the scanning requires a warrant at all—will often hinge on how we model scanning. The key issue is what you might call the "eye" problem of scanning. When a scan occurs, what is seen? In the physical world, searches are usually about exposing information to the human eye. Think of what happens when a police officer searches a house for drugs. The officer might break down the door and walk from room to room, shining his flashlight in every nook and cranny. He might use his hands to move items that obscure his view, opening drawers, opening doors, and putting aside items that might be concealing the drugs he is looking for. His hands move items into view, and his human eyes see what is there. In physical space, we naturally understand what is searched based on what the eye sees.

But what is the "eye" of digital scanning? If a government agent sets a filter and scans a database, resulting in an output file responsive to the filter, what is "seen"? On one view, maybe the computer's central processing unit is the relevant eye. On that view, all of the information in the database that is scanned through is seen. If so, scanning a massive database is inherently a massive-scale search. Scanning through a database of data concerning one billion different people may be like rummaging through the homes (or searching through the pockets) of a billion people. Such a warrant may look like a general warrant.

But there are other ways to model the process. Maybe information is seen only when it is revealed to human observation. We can imagine two versions of this model. First, maybe the relevant eye is simply the human eye that sees the output file. On this view, the scope of any search is defined solely by what the output file contains. Alternatively, maybe the seeing occurs by human implication, by a pairing of the eye and the mind. On this view, a human reads the output file from the scanning and reaches conclusions about what the database contains. Perhaps the scope of the search is based on what the human eye and brain can collectively deduce from the output file in light of the filter setting.

As this example suggests, the Fourth Amendment implications of scanning depend in large part on how scanning is modeled. Digital scanning does not fall obviously into a particular framework. Several analogies are at least facially plausible. The next Part shows the range of cases in which the issue has arisen, and the different approaches courts have taken.

II. Existing Approaches to Modeling Scanning and the Fourth Amendment

In the last two years, lower courts have encountered the scanning problem and have disagreed about how to model the Fourth Amendment implications of scanning. Some courts treat scanning like looking through all of the data

scanned. On that view, a scan through data is a massive and unconstitutional search—a view adopted, most prominently, in a Fifth Circuit's ruling on geofence warrants.[48] Other courts suggest that a digital filter only sees what the filter reveals, either directly or by implication.[49] This Part explains the cases, addressing both the technological settings in which courts are encountering the issue and how they use different models to apply the Fourth Amendment to digital scanning.

A. Focusing on the Data Searched

The first approach to modeling data scanning focuses on the data that is scanned. On this view, a data scan should be analogized to an individual going into and personally observing the data scanned. The filter "sees" all, resulting in a search of the entire massive database. The scan is like searching for a needle in a haystack, in which the entire haystack is necessarily searched to find the needle. Although the filter setting changes what information is collected by the search, it does not change what is searched. The key question is what data goes through the filter, and the type and amount of data it contains, rather than the setting of the filter or what is collected. On this view, scans through massive databases are necessarily extraordinarily broad searches—and often unconstitutional ones.

1. Geofence Warrants and *United States v. Smith*

The U.S. Court of Appeals for the Fifth Circuit adopted this reasoning in a 2024 decision on geofence warrants. In *United States v. Smith*, an unknown assailant attacked a postal worker on a rural route.[50] The assailant also stole the mail the postal worker carried, which turned out to be worth more than $60,000.[51] Security camera footage from a nearby farm revealed that a man who was carrying a cell phone was lying in wait for the postal worker. [52] In an effort to identify the man and his potential co-conspirators, investigators served a geofence warrant on Google for location records of any Google users who had enabled the Location History feature and were in the area of the attack within a one-hour period of when it occurred.[53]

---

[48] *See* United States v. Smith, 110 F.4th 817, 838 (5th Cir. 2024) ("Accordingly, we hold that geofence warrants are general warrants categorically prohibited by the Fourth Amendment."), *cert. denied*, No. 24-7237, 2025 WL 3131804 (Nov. 10, 2025).

[49] *See*, *e.g.*, People v. Seymour, 536 P.3d 1260, 1275–76 (Colo. 2023), discussed *infra* Section II.C.

[50] 110 F.4th at 820. *Smith* was written by Judge King and joined by Judges Ho and Englehardt. *Id.*

[51] *See id*. at 820.

[52] *See id*. at 820–21.

[53] *See id*. at 826–27.

Geofence techniques like the one in *Smith* were made possible by Google's since-discontinued practice of allowing users to opt into Google creating and storing a user's location history on Google's own servers.[54] Here's how it worked. If you opted in to the feature "Location History," Google would create a detailed record of your location history and store it internally. Google assembled all the Location History records of every opted-in user worldwide into a single massive database that they called "Sensorvault."[55] At the time of the warrant in *Smith*, the Sensorvault database contained the records of approximately 592 million people around the world—roughly eight percent of the global population.[56]

An affidavit filed in a different geofence warrant case by a Google employee, Sarah Rodriguez, gives us some details of how Google searched its database to comply with a geofence warrant.[57] According to Rodriguez, Google scans the entire Sensorvault database to find any geofence warrant match.[58] "Google must conduct the search across all [Location History] data to identify users with [Location History] data during the relevant timeframe," she explains, "and run a computation against every set of stored [Location History] coordinates to determine which records match the geographic parameters in the warrant."[59]

It is not clear from Rodriguez's affidavit whether Google first created a subset of the Sensorvault database narrowed to the specific time window of a particular geofence—in this case, the one-hour period around the attack—and then searched through only that one-hour period of the database. Alternatively, it is possible that Google simply searched through all the stored Sensorvault data for all time for the records responsive to that one-time window and area.

Either way, the key point is this: To carry out the warrant in *Smith*, Google searched the entire Sensorvault database containing the records of 592 million people—perhaps covering just one hour of records, or perhaps much longer.[60] Google set a filter that matched the Location History coordinates of the crime as indicated in the warrant, and it scanned through the full database

---

[54] Google has since discontinued the practice. *See id.* at 822 n.3.

[55] *See id*. at 822–23.

[56] *See id.* at 823 ("In October 2018, Google estimated that approximately 592 million—or roughly one-third—of Google's users had Location History enabled."). The world population in 2018 was about 7.7 billion people, so this was about 8% of the world's population. *See World Population by Year*, WORLDOMETER, https://www.worldometers.info/world-population/world-population-by-year/ [https://perma.cc/GZ78-QEG5].

[57] Declaration of Sarah Rodriguez at 2, United States v. Chatrie, 590 F. Supp. 3d 901 (E.D. Va. 2022) (No. 19-cr-00130) (emphasis omitted), *aff'd*, 107 F.4th 319, 324 (4th Cir. 2024), *aff'd on reh'g en banc*, 136 F.4th 100 (4th Cir. 2025) (en banc), *cert. granted*, No. 25-112, 2026 WL 120676 (U.S. Jan. 16, 2026).

[58] *Id.*

[59] *Id.*

[60] *See* 110 F.4th at 824.

looking for matches with the location coordinates within the geofence. When a positive "hit" occurred, Google recorded the existence of that phone in the geofence—and, through a complex process it created, eventually disclosed the information to the government.[61]

In *Smith*, the geofence warrant yielded hits on three different phones. Google sent investigators a list of the hits that looked like this:[62]

Device ID	Date	Time	Latitude	Longitude
1091610859	2/5/2018	17:22:45 (-06:00)	34.9044587	-90.2159436
1091610859	2/5/2018	17:24:45 (-06:00)	34.9044587	-90.2159436
1091610859	2/5/2018	17:27:04 (-06:00)	34.9044587	-90.2159436
1091610859	2/5/2018	17:27:35 (-06:00)	34.9044587	-90.2159436
1091610859	2/5/2018	17:28:06 (-06:00)	34.9044587	-90.2159436
1091610859	2/5/2018	17:28:42 (-06:00)	34.9044587	-90.2159436
1091610859	2/5/2018	17:30:56 (-06:00)	34.9044587	-90.2159436
1353630479	2/5/2018	17:58:35 (-06:00)	34.9044587	-90.2159436
1577088768	2/5/2018	17:22:27 (-06:00)	34.9040345	-90.2155529
1577088768	2/5/2018	17:24:04 (-06:00)	34.9042131	-90.2155945
1577088768	2/5/2018	17:25:08 (-06:00)	34.9045528	-90.2151712

Investigators asked Google to de-anonymize the Device IDs for all three phones. It turned out that two of the phones belonged to the wrongdoers: one was Jamarr Smith, the assailant, and the other was Gilbert McThunel, a co-conspirator.[63]

The Fifth Circuit ruled that Google's scanning to execute the geofence warrant in *Smith* violated the Fourth Amendment.[64] The Sensorvault database was too large to search, at least without first naming suspects. The Fourth Amendment requires warrants to specifically describe "the place to be searched, and the persons or things to be seized."[65] That requirement was violated not just in this case, the Fifth Circuit reasoned, but in every case involving a search through the Sensorvault database to identify suspects:

> [The warrant] forces the company to search through its *entire* database to provide a new dataset that is derived from its entire Sensorvault. In other words, law enforcement cannot obtain its requested location data unless Google searches through the entirety of its Sensorvault—all 592 million individual accounts—for *all* of their locations at a given point in time. Moreover, this search is occurring

[61] *See id.* at 825.

[62] This chart appears in *United States v. Smith*, although for the sake of simplicity I have deleted two additional columns that appear in the published chart. *See id.* at 828. They contained the source of the record (Wi-Fi or GPS), and the radius of the phone location in meters. *See id.*

[63] *See id*.

[64] *See id.* at 838 ("[W]e hold that geofence warrants are general warrants categorically prohibited by the Fourth Amendment.").

[65] U.S. CONST. amend. IV.

> while law enforcement officials have *no idea* who they are looking for, or whether the search will even turn up a result. Indeed, the quintessential problem with these warrants is that they *never* include a specific user to be identified, only a temporal and geographic location where any given user *may* turn up post-search. That is constitutionally insufficient.[66]

The Fifth Circuit specifically rejected the government's claim that the scope of the search should be determined by the information obtained rather than the information searched through. According to the Fifth Circuit, however, it was the data searched, rather than the data sought, that mattered: "While the *results* of a geofence warrant may be narrowly tailored, the *search* itself is not."[67] The warrant could not allow "law enforcement to rummage through troves of location data from hundreds of millions of Google users without any description of the particular suspect or suspects to be found."[68] The result was a bright line rule of Fourth Amendment law: "[W]e hold that geofence warrants are general warrants categorically prohibited by the Fourth Amendment."[69]

Note the conceptual understanding of scanning in *Smith*. According to the court, the Fourth Amendment implications are set by the data searched, not the data collected. The *Smith* warrant had ordered Google to gather the records, not to gather them in any particular way. But Google's decision to amass all of its data into a single database effectively determined the legality of a warrant ordering Google to search for the data that the government wanted. The Fourth Amendment implications of scanning were defined by the data searched through, not the data collected—and searching through the data of 592 million people was just too broad.

2. Tower Dump Cases Adopting *Smith*

Although the *Smith* case arose in the context of geofence warrants, two judges have recently applied *Smith* to conclude that tower dump warrants likewise violate the Fourth Amendment.[70] Some background about cell phone surveillance is useful to understand tower dumps. When a cell phone is on, it

---

[66] *Smith,* 110 F.4th at 837.

[67] *Id.*

[68] *Id.* at 837–38.

[69] *Id.* at 838.

[70] *See In re* Four Applications for Search Warrants, No. 25-cr-38-CWR, 2025 WL 603000, at *7 (S.D. Miss. Feb. 21, 2025) ("For the reasons the Fifth Circuit articulated in *Smith*, the Court concludes that the Government's tower-dump warrant applications are not supported by probable cause and particularity."). That ruling has since been appealed by the United States to the district court. *See* Lars Daniel, *DOJ Delays Tower Dumps Appeal as Digital Privacy Battle Intensifies*, FORBES (June 29, 2025), https://www.forbes.com/sites/larsdaniel/2025/06/29/doj-delays-tower-dumps-appeal-as-digital-privacy-battle-intensifies/ [https://perma.cc/RB3J-8BKR].

connects to a nearby cell site to set up a communications channel between the phone and the cell phone network. As shown in *Carpenter v. United States*, this technological feature enables a kind of tracking.[71] Because phones will generally be near the towers that they connect to, cell site location information can create a historical record of where the phone's user was located.[72] If the government has a particular phone in mind, it can track the location of that phone's user by collecting a record of the cell sites to which the phone is connected over time.

Tower dumps are something like a mirror image of this sort of tracking. They focus on one or more cell sites, and they seek records about all the phones that were connected to that cell site at a particular past time. Instead of starting with a known user and seeking to learn that user's location, tower dumps start with a known location and seek to learn who was holding a phone there.[73]

The simplest version of a tower dump focuses on a single tower. The government will obtain an order compelling the provider to disclose a list of all of the phones connected to that cell site during a particular past time window. This can produce a very long list, as thousands of users could be connected to a single site. The term "tower dump" reflects this large-scale information transfer. Cell sites are often bolted to physical towers, and compelling records of every user is a "dump" of information from that tower. Thus, the term: tower dump.[74]

Tower dumps are particularly effective when suspects are thought to have been involved in multiple events at different places and times. A set of different data points enables cross-referencing that typically leads to a small (and therefore useful) list of suspects. If investigators know a few places where criminals were thought to be, but not who the criminals were, they can seek the records of which phones were connected to local-area sites at those times and places and cross-reference the list to find phones in common. For example, if robberies with a particular modus operandi occurred in various places and times, a list of which phones happened to be in each of the different areas may be a very short list that is likely to include the phones used by the robbers.[75]

Although most courts so far have concluded that tower dumps do not implicate the Fourth Amendment at all,[76] two courts recently disagreed.[77] After

---

[71] 585 U.S. 296, 301, 309 (2018).

[72] *See id.* at 301.

[73] *See id.* at 316 (defining tower dumps as "a download of information on all the devices that connected to a particular cell site during a particular interval").

[74] *See* Commonwealth v. Perry, 184 N.E.3d 745, 752–55 (Mass. 2022) (providing background information on tower dumps and their utility for law enforcement).

[75] *See, e.g.*, *id.* at 745 (using tower dumps in this context).

[76] *See, e.g.*, United States v. Rhodes, No. 19-CR-73-AT, 2021 WL 1541050, at *2 (N.D. Ga. Apr. 20, 2021); United States v. Walker, No. 18-CR-37-FL-1, 2020 WL 4065980, at *8 (E.D.N.C. July 20, 2020).

concluding that tower dump data is protected under the Fourth Amendment, the courts next addressed when tower dump warrants are constitutional. Both courts followed the Fifth Circuit's reasoning and ruled that tower dump warrants are categorically unconstitutional under the reasoning of *Smith*.[78]

In the first case, heard by Magistrate Judge Andrew Harris in 2025 in the U.S. District Court for the Southern District of Mississippi, investigators sought to identify members of a violent street gang suspected of various murders and thefts.[79] They applied for a warrant to compel the disclosure of phone numbers and identifiers for phones at nine locations during particular windows of time ranging from ten minutes to an hour at each place.[80] Although the court's opinion does not dwell on the details, it appears likely that these were the times and places of crimes believed to have been committed by the gang members. Cross-referencing the lists would then yield the likely members of the gang.

Magistrate Judge Harris refused to issue the warrant under the Fifth Circuit's *Smith* decision.[81] According to Judge Harris, "the Government is essentially asking the Court to allow it access to an entire haystack because it may contain a needle."[82] This was improper, the court reasoned, because it would constitute a search of "many other flakes of hay in the stack"[83] that "could involve the location data of thousands of cell phone users in various urban and suburban areas."[84] Under *Smith*, scanning through all of the data in the cell site databases would be a massive broad-scale search. Like the geofence warrant in *Smith*, the court reasoned, such a broad-scale search rendered the proposed warrant an unconstitutional general warrant.[85]

Later the same year, the U.S. District Court for the District of Nevada took a similar approach to tower dumps in a murder-for-hire case investigated by local police in a rural county.[86] Because there was only one time and place where the crimes occurred, no cross-referencing was possible. The government instead obtained a warrant compelling Verizon and AT&T to hand over a list of all phone numbers that connected to cell sites in the area of the murder in the

---

[77] *See In re* Four Applications for Search Warrants, No. 25-cr-38-CWR, 2025 WL 603000, at *1 (S.D. Miss. Feb. 21, 2025); United States v. Spurlock, 778 F. Supp. 3d 1136, 1145 (D. Nev. 2025).

[78] *See In re* Four Applications for Search Warrants, 2025 WL 603000, at *7; *Spurlock*, 778 F. Supp. 3d at 1145.

[79] *See In re* Four Applications for Search Warrants, 2025 WL 603000, at *1.

[80] *See id.*

[81] *See id*. at *7 ("For the reasons the Fifth Circuit articulated in *Smith*, the Court concludes that the Government's tower-dump warrant applications are not supported by probable cause and particularity.").

[82] *Id.* at *8.

[83] *Id.*

[84] *Id.*

[85] *Id.*

[86] United States v. Spurlock, 778 F. Supp. 3d 1136, 1139 (D. Nev. 2025).

six-hour window during which the crimes were thought to have occurred.[87] The providers returned Excel spreadsheets listing a total of 1,686 phones.[88] The lead investigator was new to tower dumps, so he manually looked through the list of numbers in the spreadsheet file, later running some through public records databases to get associated identities.[89] Adopting *Smith* and Magistrate Judge Harris's prior tower dump ruling, the court ruled that tower dumps are categorically unconstitutional.[90]

B. Focusing on the Filter Setting

A second way to approach the Fourth Amendment implications of scanning is to focus on the filter setting. On this view, the key question is what information can be directly or indirectly learned from how the filter was configured in light of the information scanned. This approach requires studying the filter and then making a context-sensitive inquiry into what information is revealed by the collection given the filter setting. Whether the scan was a search, and if so, how far it extended, is determined by what facts the scanning reveals in light of the context in which the scanning occurred.

1. Reverse Keyword Searches and *People v. Seymour*

This method was used in two recent state supreme court decisions and in opinions by several judges on the Fourth Circuit. The first case is *People v. Seymour*, a 2023 ruling by the Colorado Supreme Court on reverse keyword warrants.[91] In *Seymour*, several people had died in a home fire that authorities suspected was an arson.[92] At first, the existing leads went cold. But then investigators came up with a theory: maybe the person who torched the home had googled its address in the period before the arson. Investigators wanted to have Google query its database of stored searches to find out who had googled the address.

This is known as a reverse keyword search.[93] The idea is to filter through a search database looking for particular search terms and then go backwards.

---

[87] *See id*. at 1140.

[88] *See id*. at 1141.

[89] *See id.* ("At the time, he did not have access to any software that would have helped him analyze the data. He just manually looked for phone numbers of interest in the Excel sheet and ran those phone numbers through a third-party public records database his department subscribed to.").

[90] *Id.* at 1145. In a footnote, the court recognized the possibility that a tower dump warrant might be constitutional in a case with multiple events that allowed cross-referencing. *Id.* at 1145 n.14. But tower dumps were categorically unconstitutional, at least where they involved "a single cell tower" and crimes that occurred at one place and time. *Id.*

[91] 536 P.3d 1260, 1268 (Colo. 2023).

[92] *See id.*

[93] *See id.* at 1268 n.1.

Granted, a search for "Taylor Swift" or "Donald Trump" would lead to millions of leads and be no reason for suspicion. But a search made before a crime occurred that was idiosyncratic to the planned crime could lead investigators to the account that made the search and then reveal the criminal's identity. In the *Seymour* case specifically, investigators figured that records of a Google search of the targeted home address in the weeks before the arson might provide a lead that could be traced back to the arsonist.[94]

Investigators acquired a warrant directing Google to disclose account information of any accounts that had been used to search one of nine variations of the home address, "5312 N. Truckee St.," in the fifteen days before the house burned down.[95] The warrant included any Google database that might collect query records, including Google Search and Google Maps.[96] Google responded by sending investigators a spreadsheet with sixty-one searches that were made by eight accounts.[97] Below is a relevant portion of the first few entries of that spreadsheet:[98]

Time	Country	Subdivisio	IP	Query	Result
2020-07-23 10:51:03.491848 -0(			2607:fb90	5312 Truckee St	Denv
2020-07-2	US	US-IL	2607:fb90	5312 Truckee St	Denv
2020-07-2	US	US-CO	2607:fb90	5312 Truckee St	Denv
2020-07-23 10:51:21.425946 -0(			2607:fb90	5312 Truckee St	Denv
2020-07-23 10:51:39.787704 -0(			2607:fb90	5312 Truckee St	Denv
2020-07-2	US	US-IL	2607:fb90	5312 Truckee St	Denv
2020-07-2	US	US-CO	2607:fb90	5312 Truckee St	Denv
2020-07-23 10:51:59.500433 -0(			2607:fb90	5312 Truckee St	Denv
2020-07-2	US	US-CO	2607:fb90	5312 Truckee St,	Den

---

[94] *See id.* at 1268.

[95] *Id.* The nine variations were as follows:

"5312 Truckee"
"5312 Truckee St"
"5312 Truckee Street"
"5312 N Truckee St"
"5312 N. Truckee St."
"5312 N. Truckee St"
"5312 N Truckee St."
"5312 North Truckee"
"5312 North Truckee Street"

See Declaration of Legal Investigations Support Analyst at 5–6, ¶ 13, People v. Seymour, 2022 WL 21817804 (Colo. Dist. Ct. Aug. 16, 2022) (No. 2021CR20001).

[96] *See Seymour*, 536 P.3d at 1269.

[97] *See id.*

[98] Rule 21 Exhibit to Motion to Suppress, Exhibit No. 17 at 1, *Seymour*, 2022 WL 21817804 (No. 2021CR20001), available at https://www.nacdl.org/getattachment/9cfa2f30-7422-4c4a-ac70-d2ec743e14c1/seymour-rule-21-exhibits.pdf [https://perma.cc/3D5Z-ZWHG]. I have deleted some columns not useful for our purposes.

Further investigation of the eight accounts indicated from the sixty-one searches led to a suspect, Gavin Seymour, who was charged with and convicted of the arson.[99]

Our understanding of how Google executed the *Seymour* reverse keyword warrant is based largely on a declaration of a Google employee, Nikki Adeli, that was entered into the record.[100] According to Adeli, Google executes reverse keyword searches by taking the search query described in the warrant and "run[ning it] over the records of searches conducted through Google Search and Maps."[101] Adeli's declaration does not make clear whether Google's stored records are limited to users who were logged into Google accounts, or if Google stores all searches. Adeli does clarify, however, that logged-in users can control whether their searches are included in Google's database: "[S]earches conducted from an authenticated user's account can be deleted by the user either manually or automatically based on the user's account settings."[102]

Adeli's declaration is light on important details. It does not say whether Google first limits the database searched to queries in the needed time window. But it does hint that Google may execute reverse keyword searches in a way that resembles how it executes geofence warrants: by a search through data of all users. Google has over one billion users over the course of a month, Adeli's declaration notes, and "Google and its records custodians do not know" before they query the database "which, if any, users may have undertaken searches that would be responsive to the warrant in question."[103] Although Adeli's declaration is oddly vague on the point, it appears that Google searches its database by looking through all searches of the hundreds of millions or even billion-plus users in the relevant time window to see if there is a match.

Put another way, the reverse keyword Google search in *Seymour* was on a similar scale to the geofence Google search in *Smith*. The *Smith* court focused on the scale of the database, concluding that the database was too large to search with a warrant. Was the warrant likewise unconstitutional in *Seymour*?

No, the *Seymour* court ruled. Although warrants had to identify the "place to be searched," *Seymour* reasoned, the warrant satisfied that requirement "when

---

[99] *See Seymour*, 536 P.3d at 1269. Seymour was one of three young men identified in this way as involved in the crime. *Id.* The leader of the group was 16-year-old Kevin Bui, who was seeking revenge for being mugged. *See* Colleen Slevin, *Man Who Sought Revenge for a Stolen Phone Pleads Guilty to Fire That Killed a Senegalese Family of 5*, AP NEWS (May 17, 2024), https://apnews.com/article/denver-fatal-arson-senegalese-family-47486a1112ff525982a03c6ed9f326e3 [https://perma.cc/XB2K-NFWS].

[100] Adeli's Declaration is available as the Rule 21 Exhibit 4 in the National Association of Criminal Defense Lawyers site. *See* Rule 21 Exhibit to Motion to Suppress, *supra* note 98, Exhibit No. 4.

[101] Declaration of Legal Investigations Support Analyst, *supra* note 95, ¶ 4.

[102] *Id.* ¶ 5.

[103] *Id.*

we consider the filter provided by the search parameters."[104] The fact that the filter scanned through an incredible amount of sensitive data was just a necessity. The government had merely, "in some lightning-fast, digital sense, very cursorily examine[d]" [105] the unrelated data in the haystack while looking for the needle. It was the filter that mattered: "Although the database is large, the narrow search terms, the timeframe constraints, and the fact that the initial search was anonymized all served to minimize any invasion of privacy resulting from the search."[106] It could be a different case if the "search parameters" included "more expansive or more intrusive" terms.[107] But at least in *Seymour*, the narrow filter resulted in a narrow kind of search: "There was no rummaging," the court concluded, "at least not of the sort that inspired the founders' creation of the Fourth Amendment."[108]

Note how the reasoning of *Seymour* stands in stark contrast to the reasoning of *Smith*. In *Smith*, the focus was on the size of the database. The constitutional focus was the amount of information scanned through, not what information was revealed. In *Seymour*, by contrast, the emphasis was on the "filter provided by the search parameters."[109] The search covered a limited time of fifteen days, and it was scanning for a match with one of nine specific terms. The scope of any Fourth Amendment search was defined by what the search parameters were looking for, not just the size of the database that was being queried. The Fourth Amendment ballgame was all about the filter, and the narrow search terms made the warrant particular and therefore constitutional.

2. Geofence Warrant Cases Rejecting *Smith*

The prior Section discussed the Fifth Circuit's 2024 ruling in *United States v. Smith* that geofence warrants are categorically unconstitutional.[110] Other courts that addressed geofence warrants have subsequently disagreed with *Smith.* The key cases are the Supreme Court of Georgia's 2025 ruling in *Jones v. State*,[111] and the various concurring opinions of judges on the Fourth Circuit's 2025 ruling in *United States v. Chatrie.*[112]

In *Jones*, a woman was found dead in her home in circumstances that indicated she had been murdered.[113] The government sought to learn what

---

[104] 536 P.3d at 1275.
[105] *Id.* at 1276.
[106] *Id.* (citing United States v. Place, 462 U.S. 696, 705–07 (1983)).
[107] *Id.*
[108] *Id.*
[109] *Id.* at 1275.
[110] *See supra* Part II.A.
[111] 913 S.E.2d 700, 703 (Ga. 2025).
[112] 136 F.4th 100, 101 (4th Cir. 2025) (en banc), *cert. granted*, No. 25-112, 2026 WL 120676 (U.S. Jan. 16, 2026).
[113] *See* 913 S.E.2d at 704.

phones were within 100 meters of the home around the four-hour window when the crime was thought to have happened.[114] Investigators served a geofence warrant on Google, and one of the phones caught in the geofence was particularly suspicious: It "was shown to be in or near the victim's home during the time of the killing, moving around the side of the home"[115] where a suspect had been spotted on surveillance video. That phone belonged to James Christopher Jones, who was later convicted of the murder. On appeal, Jones made a *Smith*-like claim that the geofence warrant was too broad because the Location History database was too large. According to Jones, the scan through the entire database "would 'touch' the data of *all* Google users," such that "Google effectively 'looks at' every Google account in the world."[116] On that view, "the warrants authorized a 'search' of every Google user's location history data"[117] and were overbroad.

The Supreme Court of Georgia did not definitively resolve this claim, reasoning that the defendant lacked standing to complain if others had been searched.[118] Because Fourth Amendment rights are personal, Jones could not challenge the search through other people's data.[119] But the court came about as close as otherwise possible to douse water on Jones's view of scanning. "Color us skeptical,"[120] the court wrote. When an investigator scans through a database, *Jones* reasoned, "he types a set of search terms or criteria into a search box, and then he sees the results, if any, that the query returns."[121] When that happens, he "does not view"[122] all the data in the system—and has no interest in doing do. "As a result, the privacy of the information that is *not* surfaced by the search query is preserved from the investigator."[123] Under those circumstances, the court reasoned, "it is hard to see how one could call what the investigator did a Fourth Amendment 'search' of the information in the database that was not returned as a search result."[124]

---

[114] *See id.*

[115] *Id.*

[116] *Id.* at 707.

[117] *Id.*

[118] *Id.* at 709.

[119] *See id.* ("But we need not resolve this argument definitively because Jones lacks the Fourth Amendment standing to make it . . . . At most, Jones has a Fourth Amendment interest in her own location history data, not the data of any of Google's other users she claims were searched. As a result, she lacks Fourth Amendment standing to challenge the warrant on the basis that it authorizes a search of other users' data." (first citing United States v. Davis, 109 F.4th 1320, 1329–30 (11th Cir. 2024); and then citing Presley v. United States, 895 F.3d 1284, 1289 (11th Cir. 2018))). I will return to this argument *infra* Part IV.

[120] *Id.* at 708.

[121] *Id.*

[122] *Id.*

[123] *Id.*

[124] *Id.*

According to the Supreme Court of Georgia, this filter-focused approach to the geofence warrant in *Jones* helped render the warrant sufficiently narrow to be constitutional. Applying that understanding, "the 'place to be searched' in the warrant was not, in any real sense, Google's entire database, but the part of the database that the police wanted to see."[125] The warrant directing the scan was particular because the filter was sufficiently narrow. The search parameters were "as specific as the circumstances permitted," as "the time range matched the approximate time period when the suspect was seen at and around the victim's home, and the geographic range was reasonably targeted to capture the suspect's movements."[126] In *Jones*, as in *Seymour*, the scope of the search was determined by the filter setting.

In May 2025, concurring opinions in a geofence warrant ruling by the en banc Fourth Circuit echoed the approaches of *Seymour* and *Jones*—but with a twist.[127] *Chatrie* involved a bank robbery in rural Virginia. Just as in *Smith* and *Jones*, the government obtained a geofence warrant compelling Google to reveal the phones near the scene of the crime around the time the crime occurred.[128] The initial Fourth Circuit panel in *Chatrie* concluded 2–1 that access to the Sensorvault database was not a search, and therefore ruled for the government without reaching the lawfulness of the warrant.[129] At the en banc stage, the full court failed to reach a merits ruling beyond a summary one-sentence affirmance. The en banc affirmance was followed by no fewer than eight separate opinions—seven concurrences and one dissent.[130] The en banc court was hopelessly divided, seven to seven, on the threshold question of whether a search occurred.[131] But there was fairly broad agreement on how to approach the scanning question.

In particular, twelve of the Fourth Circuit judges wrote or signed concurrences agreeing that the Fifth Circuit's approach in *Smith* was wrong. That conclusion, expressed in two concurrences, focused on the fact that it was Google that scanned the database, not the government. Judge Richardson, writing for seven judges, addressed that issue only in a footnote. "[T]he right question," according to Judge Richardson, "is what information Google gave to the

---

[125] *Id.* at 710.

[126] *Id.*

[127] United States v. Chatrie, 136 F.4th 100, 101 (4th Cir. 2025) (en banc), *cert. granted*, No. 25-112, 2026 WL 120676 (U.S. Jan. 16, 2026).

[128] United States v. Chatrie, 107 F.4th 319, 324 (4th Cir. 2024), *aff'd on reh'g en banc*, *Chatrie*, 136 F.4th 100.

[129] *Id.* at 339.

[130] For a breakdown, see Orin S. Kerr, *The Fourth Circuit's Geofencing Case Ends Not with a Bang but a Whimper*, VOLOKH CONSPIRACY (May 2, 2025), https://reason.com/volokh/2025/05/02/the-fourth-circuits-geofencing-case-ends-not-with-a-bang-but-a-whimper [https://perma.cc/Z2KC-Q7VR].

[131] Among the fifteen judges, seven held there was a search, seven held there was no search, and one did not answer the question. *See id.*

government, not what data Google perused to find that information."[132] It was a "mistake" to focus on Google's "digging through its own data, most of which it never turned over."[133]

Judge Berner's concurrence for five more judges expanded on the point, flatly rejecting the Fifth Circuit's conclusion in *Smith* that geofence warrants were *per se* unconstitutional.[134] The size of the database that Google searched through was a "red herring," Judge Berner reasoned, because a Fourth Amendment search occurs "only when the government accesses the requested information, not when a company begins looking through its internal database."[135] According to Judge Berner:

> The proper focus of our inquiry is the data the government obtains, not the size of Google's database. Though the Fifth Circuit refers to this proposition as "breathtaking," any other approach would be nonsensical. The scope of a search does not depend on what a company's compliance officer incidentally encounters—but never discloses to law enforcement—while looking through the company's database to fulfill a particular request.[136]

Later in the Article, I will consider whether the Fourth Circuit's focus on the information turned over, rather than what the provider searched through, is persuasive.[137] For now, the key takeaway is the Fourth Circuit's opposition to the Fifth Circuit's reasoning in *Smith*.

C. Focusing on the Output

A third approach to data scanning should be considered. Perhaps the scope of a database search should be based solely on the output. That is, what does the result set report on its face? The results of a scan will either be a null result (that is, no hits) or else some report about positive hits. The third approach would define the scope of a search by the report itself: what information does it expressly reveal?

This is arguably just a minor variation from the second approach focused on the filter setting. After all, that method also looked to the output in the context of the filter setting and the database searched. But the third approach is

---

[132] *Chatrie*, 136 F.4th at 139 n.17 (Richardson, J., concurring) (emphasis omitted).

[133] *Id.* According to Judge Richardson, the contrasting Fifth Circuit's approach was foreclosed by a prior en banc ruling, *Leaders of a Beautiful Struggle v. Baltimore Police Department*. *Id.* (citing 2 F.4th 330, 344 (4th Cir. 2021) (en banc)).

[134] *Chatrie*, 136 F.4th at 155–56 (Berner, J., concurring).

[135] *Id.* at 155.

[136] *Id.* (emphasis omitted) (citation omitted) (quoting United States v. Smith, 110 F.4th 817, 838 n.12 (5th Cir. 2024), *cert. denied*, No. 24-7237, 2025 WL 3131804 (Nov. 10, 2025).

[137] *See infra* Section IV.C.

different in an important way. Instead of looking at what the filter setting and output allow the government to infer, it looks solely at the output in isolation and asks what the output directly states.

Although no cases expressly adopt this approach, it is consistent with the existing practice of allowing warrantless scans of real-time internet traffic for non-content records—so-called Internet pen registers. Some context may be useful. Historically, a "pen register" was a device that the phone company could install at its headquarters to make a permanent record of the numbers dialed from a particular telephone line.[138] The odd term is explained by that history: the machine used a pen to register the number dialed, recording the electrical impulses as the caller dialed the phone.[139] The federal Pen Register Statute has preserved that function but expanded it beyond the telephone setting to include the Internet.[140] Under the modern Pen Register Statute, the government can order an Internet provider to record and preserve the non-content "dialing, routing, addressing, and signaling" information from an Internet account without a warrant.[141] Collecting that information is not a search under prevailing law, the theory runs, so no warrant is needed.[142]

But this theory runs into a potential objection in light of how Internet traffic works. In the traditional telephone setting, protected contents of communication and unprotected non-content records exist at different times. You dial the number first and have the call second, making it easy for the government to surveil one, but not the other. Internet traffic works differently. Internet traffic is packet-based, which means that a typical cable of Internet traffic will jumble together protected and non-protected data. It's like a stream with many objects in it, all going down the river together. Depending on how the Internet pen register is installed, it may require inserting a filter in that stream and collecting only the non-content records responsive to the filter.

The result is a real-time equivalent to a filter search through a stored database. A filter is installed at a particular point, and it scans through everything passing by. The filter "must look at all of the traffic to see if it is the communication sought."[143] For example, if the court order allows the government to record that a message was sent from a particular IP address, without collecting

---

[138] *See* United States v. N.Y. Tel. Co., 434 U.S. 159, 161 n.1 (1977) ("A pen register is a mechanical device that records the numbers dialed on a telephone by monitoring the electrical impulses caused when the dial on the telephone is released. It does not overhear oral communications and does not indicate whether calls are actually completed.").

[139] *See id.*

[140] *See* Kerr, *Big Brother*, *supra* note 44, at 623–44.

[141] 18 U.S.C. § 3121(a)–(c).

[142] *See* Smith v. Maryland, 442 U.S. 735, 745–46 (1979) ("The installation and use of a pen register, consequently, was not a 'search,' and no warrant was required.").

[143] Kerr, *Big Brother*, *supra* note 44, at 649.

the message itself, the filter will scan for that IP address and record that fact without saving the part of the packet that contains the message.[144]

This raises a similar Fourth Amendment challenge to what we have seen previously in the context of stored data. If the filter scans through both protected and non-protected records, but it only registers and outputs non-content records, has a "search" occurred? The filter has "seen" contents, but it was programmed only to register and output non-content records. If the output only contains non-content records, rather than content records protected by the Fourth Amendment, has a search occurred?

So far, courts have not found a search occurring on these facts without actually confronting the specific scanning problem. One example is the U.S. Court of Appeals for the Ninth Circuit's 2008 opinion in *United States v. Forrester*.[145] In *Forrester*, the government obtained a pen register order requiring the Internet provider PacBell to install a "pen register analogue known as a 'mirror port'" on the account of a narcotics suspect.[146] "The mirror port was installed at PacBell's connection facility in San Diego," the Ninth Circuit explained, and it "enabled the government to learn the to/from addresses of [the suspect's] e-mail messages, the IP addresses of the websites that [he] visited and the total volume of information sent to or from his account."[147]

In ruling that this surveillance was not a search, the Ninth Circuit focused on the records that the mirror port generated. The surveillance "reveal[ed] the to/from addresses of e-mail messages, the IP addresses of websites visited and the total amount of data transmitted to or from an account."[148] By looking at what the surveillance revealed—that is, the result set that the filter created—the court found the surveillance analogous to a traditional telephone pen register that the Supreme Court had ruled was not a search.[149] The emphasis was on what information was revealed, not what information was scanned.

To be sure, *Forrester* offers only limited guidance. We don't know the details of how the mirror port was installed. The defense did not specifically argue that the information scanned through the mirror port made the tool's use a search. And it is not entirely clear that the reasoning of *Forrester* reflects the third method as opposed to the second; parts of the opinion touch on what could be inferred from the collection, which is more of a filter issue than just a raw output question. Still, the reasoning of the opinion is at least modestly suggestive of a primary focus on the output. On this view, perhaps we can learn the Fourth Amendment implications of scanning by looking at the output

---

[144] *See id.* at 649–50.
[145] 512 F.3d 500, 505 (9th Cir. 2008).
[146] *Id.*
[147] *Id.*
[148] *Id.* at 510.
[149] *See id.* (citing Smith v. Maryland, 442 U.S. 735, 742, 744 (1979)).

file and seeing what information it contains. If that information is protected by the Fourth Amendment, then a search occurred, with its scale reflecting the amount of information in the output file. If that information is not protected by the Fourth Amendment, then no search occurred.

To sum up, there are three basic ways to model scanning and the Fourth Amendment: focusing on the database, the filter, or the output file. The question is, which is correct?

III. The Case for a Filter-focused Approach to Data Scanning

This Part makes the affirmative case for the filter-focused approach. The Fourth Amendment search test focuses on what information is exposed to outside observation, either explicitly or implicitly. When a filter scans through data, however, the filter does not "see" the data in a Fourth Amendment sense. A digital filter is dumb. It blindly checks for a match with the filter, much like a key to a door checks for a match with a lock. When no match is found, the non-matching information scanned is not revealed. Rather, what is exposed by a scan is answered by a contextual inquiry: what has the scan revealed based on the filter setting, the data scanned, and the output? Put another way, what facts were actually learned about what was in the database by looking at the output of the query in light of the specific filter setting?

Under the view I offer, the Fifth Circuit's approach is wrong. Database scans do not become unconstitutional simply because they require scanning through an extraordinary amount of non-responsive data. What filters reveal in context is what matters, not the absolute questions of what they scan through or what the output looks like.

This Part makes the case for the filter-focused approach in four steps. First, it explains the Supreme Court's relevant caselaw governing what is a search.[150] Second, it explains why the filter-focused approach is consistent with that doctrine.[151] Third, it explains why the database approach is arbitrary, and how it leads to overprotection by imagining searches where none exist.[152] Finally, it explains why the output approach is similarly flawed, and how it leads to underprotection by missing searches that exist by implication from output.[153]

[150] *See infra* notes 154–183 and accompanying text.
[151] *See infra* notes 184–189 and accompanying text.
[152] *See infra* notes 190–199 and accompanying text.
[153] *See infra* notes 200–207 and accompanying text.

A. The Supreme Court's Relevant Search Precedents

There is a massive and often confusing literature on what is a Fourth Amendment search.[154] Fortunately for us, understanding search law in the specific context of scanning needs only a narrow focus. Because data scans do not involve physical intrusion, the correct question is how to apply the expectation of privacy test first stated by Justice Harlan's concurrence in *Katz v. United States*.[155] And because we are already assuming that some data in the database is protected, we can focus on a preliminary question: At exactly what point does a search of information happen?

Fortunately, the Supreme Court has a quartet of cases that answer this inquiry. According to those cases, Fourth Amendment searches are about exposures of information. For a search to occur, information must be exposed to human observation. The constitutional focus is on what information is actually revealed, either expressly or implicitly. The mere potential for exposure is not enough. We can root these principles in four major rulings: *Arizona v. Hicks*,[156] *Kyllo v. United States*,[157] *United States v. Karo*,[158] and *United States v. Jacobsen*.[159]

The Supreme Court endorsed the exposure-based understanding of searches in *Hicks*. Immediately after hearing shots fired from an apartment, officers entered the apartment to search for the shooter. Officer Nelson spotted expensive stereo equipment in what was otherwise a run-down apartment.[160] Suspecting that the equipment might be stolen, Officer Nelson moved the equipment so he could see the serial numbers—numbers that he then reported to headquarters.[161] In an opinion by Justice Scalia, the *Hicks* Court ruled that moving the equipment was a distinct search because it "exposed to view concealed portions of the apartment or its contents" therefore "produc[ing] a new invasion of respondent's privacy unjustified by the exigent circumstance that validated the entry."[162] It did not matter that the search "uncovered nothing of any great personal value to respondent—serial numbers rather than (what

---

[154] *See generally* Orin S. Kerr, *Four Models of Fourth Amendment Protection*, 60 STAN. L. REV. 503, 503–04 (2007) (arguing that the Supreme Court has been unclear in delineating its reasonable expectation of privacy test and offering four possible models of interpretation); *see also* Daniel J. Solove, *Fourth Amendment Pragmatism*, 51 B.C. L. REV. 1511, 1538 (2010) (describing Fourth Amendment law as a "malaise").

[155] 389 U.S. 347, 360–61 (1967) (Harlan, J., concurring). For why the absence of physical intrusion means that the *Jones* trespass approach is not in play, see Orin S. Kerr, *The Two Tests of Search Law: What Is the* Jones *Test, and What Does That Say About* Katz*?*, 103 WASH. U. L. REV. 309, 348–50 (2025).

[156] 480 U.S. 321, 323 (1987).

[157] 533 U.S. 27, 29 (2001).

[158] 468 U.S. 705, 707 (1984).

[159] 466 U.S. 109, 111 (1984).

[160] *Hicks*, 480 U.S. at 323.

[161] *See id.*

[162] *Id.* at 325.

might conceivably have been hidden behind or under the equipment) letters or photographs."[163] As Justice Scalia put it, "[a] search is a search, even if it happens to disclose nothing but the bottom of a turntable."[164]

Note the key move in *Hicks*. A distinct search occurred because moving the equipment "exposed to view" what had been "concealed." A search happened when the government "uncovered" what was previously "hidden" under the equipment. The serial numbers of the stereo equipment had been there the whole time, hidden from view. It was the exposure to the officer's view of the previously hidden away information that triggered the search.

The Supreme Court's exposure standard from *Hicks* covers not just the information that is actually exposed, but also what else that information implicitly reveals. This is clear from *Kyllo v. United States*, in which the government directed an infrared thermal imaging device at a home.[165] The imaging device measured the surface temperature of the exterior of the wall based on the amount of infrared-spectrum light it emitted.[166] Whether this measurement was a search hinged partly on whether it was understood as revealing information about the inside of the home, or merely the outside of the home. Writing for the dissent, Justice Stevens argued that knowing the surface temperature of the exterior wall provided zero information about the interior of the home.[167] In his view, reaching a conclusion about what was going on inside the home from the exterior was merely a matter of inference, which could not be a search.[168]

Writing for the *Kyllo* majority, Justice Scalia dismissed the dissent's approach as "extraordinary."[169] What the dissent described as "making an inference" was better understood as "analysis," Justice Scalia suggested.[170] And it was generally the case that surveillance tools produced output that required some analysis. For example, a tool that could see through walls would not produce "an 8-by-10 Kodak glossy" of what was inside—that is, a clear high-resolution photograph of it.[171] Although this passage is not crystal clear, the *Kyllo* majority seems to be saying that the fact that some information had to be inferred from the output of the surveillance did not mean that the information was not learned from the surveillance. The thermal imaging tool yielded in-

---

[163] *Id.*

[164] *Id.*

[165] 533 U.S. 27, 29–30 (2001).

[166] *See id.*

[167] *See id.* at 41–42 (Stevens, J., dissenting).

[168] *See id.* at 46 ("Since what was involved in this case was nothing more than drawing inferences from off-the-wall surveillance, rather than any 'through-the-wall' surveillance, the officers' conduct did not amount to a search . . . .").

[169] *Id.* at 36 (majority opinion).

[170] *Id.*

[171] *Id.* Kodak was the largest manufacturer of film for photographs, and an "8-by-10 glossy" was an 8-inch by 10-inch image on glossy film—a particularly large photograph.

formation about the inside of the home, even if that was only implicit from the revealed information about the exterior.

The Court has also made clear that merely the potential for an invasion of privacy is not a search. In *United States v. Karo*,[172] officers installed a surveillance device known as a radio beeper inside a can containing chemicals used in the drug trade. To track the suspect, agents sold the can to the suspect with the beeper inside.[173] At first, the radio beeper was turned off. The Supreme Court held that the transfer of the can to the suspect's possession, with the surveillance tool turned off inside, was not a search:

> It conveyed no information that Karo wished to keep private, for it conveyed no information at all. To be sure, it created a *potential* for an invasion of privacy, but we have never held that potential, as opposed to actual, invasions of privacy constitute searches for purposes of the Fourth Amendment.[174]

The Court added: "It is the exploitation of technological advances that implicates the Fourth Amendment, not their mere existence."[175] The surveillance tool existed, and it had the potential to convey information. But because it "conveyed no information at all," it did not trigger a search.[176]

One last Supreme Court ruling to consider is *United States v. Jacobsen*,[177] which asked whether chemical testing to determine if a white substance was cocaine constituted a Fourth Amendment search. In holding that the test was not a search, the Court focused on exactly what the testing revealed. A positive test would only indicate the presence of a substance that was illegal to possess. No search would occur in that case, as "no legitimate interest has been compromised" by revealing only the presence of the drugs in light of Congress deciding "to treat the interest in privately possessing cocaine as illegitimate."[178] A negative test would not trigger a search either, as "merely disclosing that the substance is something other than cocaine . . . reveals nothing of special interest."[179] Because neither possible test outcome would amount to a search, chemical testing was not a search.

What is the takeaway from these cases? The decisions look to what information was "exposed to view," in the words of *Hicks*.[180] They consider not just the raw facts that were exposed, but also include "making an inference"

---

[172] 468 U.S. 705, 708 (1984).
[173] *See id*. at 707.
[174] *Id.* at 712.
[175] *Id.*
[176] *Id*.
[177] 466 U.S. 109, 111 (1984).
[178] *Id.* at 123 (internal quotations omitted).
[179] *Id*.
[180] *See* 480 U.S. 321, 325 (1987).

about those raw facts in context, as *Kyllo* put it.[181] On the other hand, they exclude the mere "potential" to invade privacy, focusing on the "actual" invasion of privacy, as *Karo* explained.[182] And instead of looking at what item was tested, they interpret the results of scientific tests based on what fact the tests reveal, as *Jacobsen* shows.[183]

B. The Case for the Filter-focused Approach

The filter-based understanding is the only one consistent with these principles. When a filter runs through data, creating an output, the filter-based approach looks to what is actually exposed to view. It considers both the raw information revealed as well as the inferences that can be drawn about those facts from the filter setting and the data searched. It looks to the actual invasion of privacy, not just the potential invasion of privacy. And it interprets the results based on what the query through the database reveals, not the scope of the database itself. This is perfectly consistent with the Supreme Court's precedents.

Importantly, the filter-focused approach is not solely about the filter setting. Rather, it is a contextual inquiry. It asks what the scan revealed based on the filter setting, the data scanned, and the output. I call it the filter-focused approach because it focuses primarily on the filter. The filter determines what the scan will trigger, so you can only understand the scan by focusing on the filter. But the filter is not everything. The ultimate question is, what facts did the scan reveal? And the answer is always derived by asking what (if anything) led to a match with that filter, and what output was recorded.

You might break it down like this. The database is the raw item to be scanned, which we might call *A*. The filter says, "if you scan through database *A* and see keyword *B*, do operation *C*, resulting in output *D*." The output says, "here is *D*, the result of all the operations *C* found through the scan, if any." The Fourth Amendment significance of the scan is not based on database *A* being searched through. Nor is it based on the raw output, *D*, appearing. Instead, we need to consider what was revealed by the combination of *A*, *B*, and *C*, to produce *D*. In effect, the result of the scan says: "Whatever was in database *A*, a pattern match for *B* yielded either hits or a null set; and if there were hits, the pattern match for *C* repeated for each hit leads to output *D*." The Fourth Amendment implications of the scan are based on what is revealed. Not just *A*, *B*, *C*, or *D*. But rather, what does *D* reveal based on the specific details of database *A*, the filter *B*, and the output *C*?

We can appreciate this by considering what it means when a scan leads to a null result—that is, no hits at all. If the filter is set so that it cannot return a

[181] *See* 533 U.S. 27, 36 (2001).
[182] *See* 468 U.S. 705, 712 (1984).
[183] *See* 466 U.S. 109, 123 (1984).

hit through the data scanned, the absence of a hit itself says nothing. Like the drug-sniffing dog that does not alert in *Illinois v. Caballes*, the lack of a hit may be meaningless.[184] This does not mean that a null result *always* means nothing. What a null result means depends on the filter setting and the data searched through. But the scope of the search has to be defined by that context, not the amount of raw data scanned through.

Consider two contrasting examples. First, imagine that Google, armed with a warrant to scan geofencing data to find records of phones located in Austin, Texas, decides to search through a database that contains location records only from Europe. Google scans through the database, using the latitude and longitude coordinates in Austin. Of course, there are no hits.

In this example, the null result tells you nothing. Google was scanning for a latitude/longitude range that it knew was not found in that particular database. The null result adds nothing to preexisting knowledge. To treat this as a profoundly invasive search through millions of accounts—akin to a vast search through all of Europe—ignores reality. What matters is what the scan reveals, not how much data it searches through. If the filter is set so that it cannot pick up any data, merely scanning through data means nothing.

In a different case, with a different filter and different database, a null result might be quite meaningful. Imagine a homicide case. The suspect claims an alibi. According to that alibi, he spent the night of the murder in front of his computer doing Internet research on the Battle of the Bulge for a World War II history paper that was due the next morning. If the government served a warrant on Google to retrieve all of the suspect's Google search terms entered that night, and Google responded that the suspect's account had not submitted any search terms at all, that null result—there were no searches by anyone logged into that account—might be quite illuminating.

A null result would not be definitive, of course. It wouldn't necessarily disprove the alibi. Maybe the suspect did his research without using Google. Maybe, for some reason, he later deleted his search results. It's possible. Nonetheless, the null result would provide some reason to be suspicious of the alibi. Because the filter was set to look for all of the suspect's searches, the lack of any searches may be reason to question whether any searches occurred—a conclusion that may be hard to square with the suspect's alibi. The null result did not say nothing, as the suspect's alibi would have predicted that there would be something. It's the context of the claim, the filter, and the database, that determines what the scan result reveals.

In this way, the approach taken by the Colorado Supreme Court in *People v. Seymour* and the Supreme Court of Georgia in *Jones v. State* was on the right

[184] *See* 543 U.S. 405, 406 (2005) (ruling that a dog sniff of a car did not constitute a search under the Fourth Amendment).

track.[185] Both cases focused their attention on what the queries revealed. Although warrants have to identify the "place to be searched," the *Seymour* warrant satisfied that requirement "when we consider the filter provided by the search parameters."[186] The non-responsive parts of the scan were merely "some lightning-fast, digital sense, very cursor[y] examin[ation]."[187] "Although the database [was] large, the narrow search terms, the timeframe constraints, and the fact that the initial search was anonymized all served to minimize any invasion of privacy resulting from the search."[188]

We can make the point more confidently than *Seymour* did. Depending on the filter setting, the scan through the database that does not yield responsive data may be less than a "lightning-fast, digital sense, very cursor[y] examin[ation]."[189] As the *Jones* court recognized, it may not be a search at all.

C. Against the Database Approach

The case for the filter-focused approach largely shows the flaws in the database approach, such as that adopted by the Fifth Circuit in the geofence warrant case of *United States v. Smith.*[190] Digital scanning doesn't search in the same way as a physical search. It merely checks for patterns. If particular parts of data scanned cannot register a hit, there is no invasion of privacy in the *Hicks* sense.[191] There is only a "potential" invasion of privacy, in the words of *Karo*, not an "actual" one.[192] As a result, the database-focused approach imagines searches where none exist. This leads to bizarre results that generally hinge on arbitrary database design over the substance of privacy invasions.

The database approach rests on an erroneous physical analogy that is the root of the problem. This approach treats a digital scan like a person looking through data. Imagine an officer physically rifling through items in a box. As he pulls items out, he sees everything searched through. That observation is important. It opens up what is seen to close inspection and later recollection. It also enables warrantless seizures under the plain view exception.[193] As a result,

---

[185] *See* People v. Seymour, 536 P.3d 1260, 1268 (Colo. 2023); Jones v. State, 913 S.E.2d 700, 710 (Ga. 2025).

[186] 536 P.3d at 1275.

[187] *Id.* at 1276.

[188] *Id.*

[189] *Id*.

[190] *See* 110 F.4th 817, 838 (5th Cir. 2024), *cert. denied*, No. 24-7237, 2025 WL 3131804 (Nov. 10, 2025).

[191] *See* 480 U.S. 321, 325 (1987).

[192] *See* 468 U.S. 705, 707, 712 (1984).

[193] *See Hicks*, 480 U.S. at 323 (noting that "in certain circumstances a warrantless seizure by police of an item that comes within plain view during their lawful search of a private area may be reasonable under the Fourth Amendment" (citing Coolidge v. New Hampshire, 403 U.S. 443, 465–71 (1971))).

what is looked through counts as what is searched. The database approach takes that physical understanding and applies it to digital scanning.

But that is the wrong analogy. Using a filter to scan through data is much more like inserting a key in a lock to check if a door will open. Keys are cut to have particular patterns matched to particular locks. The cut on a key is just a code. It is a specific combination of positions. When you insert a key into a lock and it fits and turns, allowing you to open the door, it does so only when the pattern on the key matches the pattern in the lock. When you try to insert the key and it doesn't fit—or it goes in, but it doesn't turn—you only learn that the key did not work.

Scanning through a database is something like trying locks to different doors to see which (if any) locks match a particular key. The testing does not "see" all of the settings inside all of the locks tested. The key is not an eye that looks inside the lock housing and observes its settings, recording exactly which key would work or even what is behind each door. The key just checks for a pattern. It either works if the pattern matches or does not work if no match is found. It is a tool to find a match, not an eye that sees everything around it.[194]

Digital scans work the same way. If data passes through the filter but no match is found, the data is not "seen" in a Fourth Amendment sense. When a filter scans through a database, reporting when there is a match with a particular binary pattern, the filter does not "see" what is happening any more than the key inserted into a lock. Imagine the filter is 010100110110001101100000 101101100 0110100101100001—which you will recall from Part I is "Scalia" in binary. If a stream of data matches the first few digits but then diverges—say, the initially-matching 010100 is followed by a 1 instead of another 0—the computer will just drop the data and report no match.

It is wrong to conceive of this process as if the computer were a sentient being. The computer doesn't think to itself, "hmmm, 0101001, that sounds mighty suspicious!" The computer won't call up another computer and say, "did you see that 0101001 over in that database? Weird, eh?" Again, consider locks and keys. A metal key is not a sentient being, either. When you insert a key into a lock, it doesn't think, "jeepers, it sure is dark in this lock cylinder," or "wow, look at that metal!" Like keys inserted into locks, scans are just dumb tools. They don't think.[195]

This problem with the database approach is further emphasized by its astonishing arbitrariness. The size of the database searched is usually just a design question for database administrators. It has no practical effect on law

---

[194] As it turns out, lower courts are divided on how to treat the Fourth Amendment implications of inserting a key in a lock. That division is not relevant here.

[195] The implications of artificial intelligence are considered *infra* Section IV.B.

enforcement. And the database administrators often could just take a few simple steps to redesign the scan to dramatically alter how much data is scanned through. If the filter used will not be triggered by this extra data, scanning that data has no bearing on any privacy question. It is bizarre to treat that private design choice as determining Fourth Amendment rights.[196]

Take a case like the Fifth Circuit's geofence ruling in *United States v. Smith*.[197] The government had ordered Google to disclose the records in a particular physical area—certain latitude and longitude bearings representing a zone about 300 meters by 300 meters in rural Mississippi—for one hour.[198] Google had complied with the warrant by searching through a preexisting database that contained the records of approximately 592 million people all around the world. In the Fifth Circuit's view, the fact that Google placed the data of all 592 million Sensorvault users into one database made the database too large to search.[199]

But there was no reason why Google had to conduct the search that way. That was Google's decision, not the government's decision. For example, imagine Google had decided to store Sensorvault using indexes. Imagine Sensorvault divided the world into 300-meter-by-300-meter zones. If you wanted to know where a phone was located, you had to look up that index where the phone was located and then search specifically through the records in that zone. If Google happened to organize its database that way, then it could have executed the warrant by scanning *only* through the 300-meter-by-300-meter zone in rural Mississippi where the crime occurred. The scan would have only affected the handful of people whose phones were reported to the government. No broader search would have been needed.

And here's the thing: this would have made absolutely no difference to the government. The government might never know. It would get the warrant and serve it on Google. Google would respond, searching just the data in the one small zone in rural Mississippi. The government would not know that this was the only database searched, nor would it care. It only asked for evidence from that area, and whether the database was just that zone or the entire world could have made no difference to the investigators. The size of the database searched is a matter of network design, not the scope of any privacy invasion. And the Fourth Amendment should not regulate private company network administration decisions, nor hinge permitted government power on those private choices.

---

[196] *See* Smith v. Maryland, 442 U.S. 735, 744–45 (1979) ("We are not inclined to hold that a different constitutional result is required because the telephone company has decided to automate.").

[197] 110 F.4th 817, 820 (5th Cir. 2024), *cert. denied*, No. 24-7237, 2025 WL 3131804 (Nov. 10, 2025).

[198] *See id.* at 827 (describing the area covered by the geofence as "approximately 98,192 square meters around the Lake Cormorant Post Office" in Lake Cormorant, Mississippi).

[199] *See id.*

D. Against the Output Approach

The output approach suffers from a similar problem. If you look only at the raw output, and if you treat that output as what was "searched," then you ignore *Kyllo*'s warning that inferences from data can be part of the search itself.[200] It's like the database-focused problem, but in reverse. Although the database-based approach ends up seeing searches where there are none, the output approach ends up ignoring searches that do exist. It misinterprets the output of a scan as if it were the only information revealed by scanning, ignoring information revealed by inference. The approach construes the Fourth Amendment too narrowly.

Consider an example. Imagine the government orders Google to disclose a list of IP addresses of every account that searched for the term "how to dispose of a dead body" on a particular day. The government responds with a list of IP addresses, a part of which might look like this:

152.23.23.12
122.22.23.16
102.12.23.14
111.23.23.213

Under the output approach, you might look at that list of IP addresses and conclude that no search occurred. Under prevailing law, collecting IP addresses is not a search.[201] And here, the only information revealed is a list of IP addresses, right?

Wrong. When you consider the filter setting, you realize that the list of IP addresses communicates more than just a list of IP addresses. Rather, it is a list of IP addresses from which a person googled the content message, "how to dispose of a dead body." The fact that the list of IP addresses does not include the contents revealed is just an arbitrary matter of how the order to Google was phrased. When the government asks a company to reveal non-content information that hinges on content, then the government is really asking about content. They're just asking in a way that the content part of the request is hidden in the question, such that the answer alone seems to consist only of non-content information when the context means it contains contents.

This is just a sleight of hand. For example, imagine Google responded to the same query with the following table:

---

[200] *See* 533 U.S. 27, 36 (2001).

[201] *See* United States v. Hood, 920 F.3d 87, 90–92 (1st Cir. 2019).

IP Address	Content Query Made
152.23.23.12	how to dispose of a dead body
122.22.23.16	how to dispose of a dead body
102.12.23.14	how to dispose of a dead body
111.23.23.213	how to dispose of a dead body

The only difference between this chart and the list of raw IP addresses is that the question—what IP addresses were used to submit the search query "how to dispose of a dead body"—is now an explicit part of the output instead of it merely being implicit. Before, the filter setting was in the question; now the filter setting is on the output page.

But this is purely a difference of form. Think of how ordinary language works. If someone asks you, "Did you cheat on your taxes?" and you reply "yes," you have done more than say "yes" in some abstract sense. Your statement "yes" implicitly incorporates the question. In context, "yes" means "I cheated on my taxes." Surely the constitutional rule should not hinge on which way the answer is phrased.

Courts have recognized this dynamic in the related setting of federal surveillance laws. For example, in *District of Columbia v. Meta Platforms, Inc.*,[202] the District of Columbia government served a subpoena on Meta (formerly Facebook) for the identities of Facebook users that Meta had determined violated its content moderation policies for vaccine misinformation. Under the Stored Communications Act, whether a subpoena could be used in that setting depended on whether the subpoena was seeking the contents of communications or merely non-content records.[203] Meta refused to comply with the subpoena, claiming that the government implicitly sought contents, not just metadata.

The Superior Court agreed with Meta that the subpoenas implicitly sought content. "[T]he logical and practical effect of revealing the identities of these users," the court reasoned, "is to reveal the contents of their communications."[204] Because providing a list of identities "necessarily means disclosing that they communicated information about vaccines,"[205] the subpoena disclosed contents. It was true, the court conceded, that providing the list did not

[202] No. 2021 CA 004450 2, 2022 WL 805196, at *2 (D.C. Super. Ct. Mar. 9, 2022).

[203] *Compare* 18 U.S.C. § 2703(c)(2) (allowing a subpoena to be used to compel non-content basic subscriber information), *with id.* § 2703(a) (requiring at least prior notice to the user, and often a warrant, to compel contents).

[204] *Meta Platforms, Inc.*, 2022 WL 805196, at *4.

[205] *Id.*

provide "the specifics of the message" of particular users.[206] The list provided content because some amount of content was revealed.[207]

The ability of records in non-content form to imply contents in substance shows the flaw in the output approach. The facts revealed by a data-scan output are not determined by the output stripped of context. As the filter-focused approach shows, the message communicated by a data scan is a byproduct of the filter, the data set, and the output. In every case, Fourth Amendment implications of the scan should be answered by the meaning communicated by the scan. And the meaning communicated by the scan has to be answered in context: what does that particular output mean in light of that particular filter scanned through that particular dataset?

IV. Applications and Implications

The principles discussed above should answer how the Fourth Amendment properly applies to several important scanning situations. At the same time, general principles do not necessarily resolve concrete cases. This section considers several important applications of the proposed approach.

I want to stress that the answers in this section are tentative. Adopting the filter-focused approach narrows the debate considerably, but it still leaves legitimate disagreement about how those principles apply. Nonetheless, my experience presenting this paper at workshops suggests that more specific guidance may be useful. For some, the applications may just illuminate the theory. For others, the applications may provide practical answers.

This Part has three Sections. It begins with important applications: geofencing, email scans, and reverse keyword searches.[208] It next considers a more speculative question of how the filter-focused approach should apply to analyses of datasets using artificial intelligence.[209] It concludes with a note of caution about relying on the provider's role as a filter suggested by judges of the en banc Fourth Circuit in *United States v. Chatrie*.[210]

A. Applications to Geofencing, Email Contents, and Reverse Keyword Searches

This Section considers possible implications of the filter-focused approach for three important situations: queries through location databases, such as with geofence warrants; content queries, such as with searches through email accounts; and reverse keyword searches through databases of search

---

[206] *Id.* at *5.
[207] *Id.*
[208] *See infra* notes 211–218 and accompanying text.
[209] *See infra* notes 219–224 and accompanying text.
[210] *See infra* notes 225–229 and accompanying text.

terms.[211] These applications show that applying the filter-focused approach requires courts to look closely at what facts can be drawn by the specific output in light of the specific filter setting and the specific database queried. Some cases will be easy. But others will be difficult.

Start with location queries, such as the geofencing techniques at issue in cases like *Chatrie* and *United States v. Smith*. In my view, queries through location databases such as geofencing should be permitted with a warrant as long as that warrant is sufficiently narrowly drawn as to the place and time covered. When a query of a location database is made, the query reveals only the location of the devices believed to be in that place and time. The information revealed by such a location scan is inherently limited. Devices can only be in one place at a time. Location information for a particular time reveals at most one fact: the location of that device at that time. When the government gets a warrant for phones located in a specific place in a specific time window, it obtains only the limited information about what phones were in that place and time.

This is not a novel scale of search, assuming it is a search at all, and it falls well within the traditional warrant power.[212] A physical world analogy might run like this: a warrant to find out what phones were in a particular place and time is something like a warrant that allows officers to break into a private home to see who is inside. After entering the home, the government can see who is within. But the warrant to enter and see who is inside is limited to that place and time. It reveals warrant-protected information about who is physically present there, but it does not allow an actual search of the people there or of their private papers.[213] From this perspective, a warrant for sufficiently narrow location information—limited in both space and time—should be constitutional if based on probable cause. Assuming the geofencing records are held to be protected by the Fourth Amendment, a narrowly-drawn warrant should be constitutional.

Next consider scans through databases containing the contents of user communications. For example, say a warrant orders Google to search its stored email database for a user who is thought to have sent a message containing a particular threatening phrase on a particular day. To get a list of suspects, the government orders Google to provide the names of every account holder who

---

[211] For each example, I assume for the sake of argument that the data is deemed protected under the Fourth Amendment. As noted earlier, that is an issue very much in dispute in the courts today. *See supra* notes 7–11 and accompanying discussion.

[212] *See supra* notes 7–11 and accompanying text (discussing novel understandings of what data is constitutionally protected).

[213] *See* Ybarra v. Illinois, 444 U.S. 85, 96 (1979) (holding that a warrant to search a place does not allow a search of persons in the place absent additional probable cause). Some courts have suggested that geofence warrants actually do search the persons of those whose identities are revealed in the geofence. But this seems plainly incorrect. Revealing the fact that a phone associated with a person has an entry in a database that suggests the person's presence at some place in the past may infringe on that person's Fourth Amendment rights, but it does not actually search the person's physical body.

sent such an email with that particular phrase on that day. Would that be constitutional under the filter-focused approach?

In my view, the Fourth Amendment answer to this question looks very different from the answer for geofencing. Unlike a geofence query, a scan through contents of stored emails belonging to many people is a stunningly massive search. What messages people send is very different from where they may be. A query through a particular person's outbox is a query through their personal papers—their virtual home online. Messages reveal what one person is saying to another in a private setting. A scan through a user's account for a particular phrase reveals whether they shared a particular message with another person or they did not.

If the government has a warrant for the contents of a single user's account, that warrant is particular in the same way that a warrant to search one person's home is particular.[214] But a search through the contents of millions of peoples' messages searches the virtual homes of not just one suspect, but potentially millions or even hundreds of millions of people. A single content scan could reveal what is inside the virtual homes of a large chunk of the planet. This is vastly broader than a geolocation search, and in my view, it falls outside the warrant power because such a warrant would plainly lack the particularity that the Fourth Amendment requires.[215]

Notably, the difference between these two queries is not the amount or type of data searched through, but rather what the filters reveal in light of the filter setting. To appreciate this, imagine that the two queried databases above were combined into a single database before the queries were made. That is, imagine Google combined both the contents of messages database *and* the entries for any stored location information of the user's phone in a single super-duper massive database covering hundreds of millions or billions of people. And now imagine that Google ran the two queries on the combined database rather than the two separate databases. That is, Google first searches the combined database just for location information, setting the filter so that it registers location information at a particular place and time but ignores content records. And then Google searches the combined database for content messages, setting the filter so that it ignores location records.

Under the filter-focused approach, combining the two databases would not alter the Fourth Amendment implications of either distinct query one iota. The location-only query through the combined database would have the exact same Fourth Amendment implications as the query through the earlier location-only database. The content-only query through the combined database

---

[214] *See generally* Orin S. Kerr, *Applying the Fourth Amendment to the Internet: A General Approach*, 62 STAN. L. REV 1005, 1044–47 (2010).

[215] *See id.*

would have the exact same Fourth Amendment implications as the query through the earlier content-only database. The key issue is what the scan reveals in context, not what information is scanned through. Because the location query is set to ignore contents, and the content query is set to ignore location, the added presence of the other data does not change the constitutional picture. But because the content query is vastly more invasive and revealing than the location query, the content query would be out of bounds, while the location query would be permitted with a sufficiently narrow warrant.

A last case to consider is reverse keyword searches, such as the search for the googled address to find the arsonist in *People v. Seymour*.[216] In my view, such searches fall somewhere in between the prior examples of allowed geofencing searches and prohibited content searches. On one hand, web searches are often extremely private. We think therefore we google, allowing search histories to provide a record of our intimate thoughts. In that sense, a web search closely resembles an email message, and scanning through web queries might resemble querying stored email contents.

On the other hand, not all web searches fit that box. Web queries can be varied. Although many are akin to private messages, others can be purely functional. A search query might just be a way to navigate the web. For example, if you wanted to see Cass Sunstein's home page, you might just google "cass sunstein" rather than navigate your way by trial and error to *https://hls.harvard.edu/faculty/cass-r-sunstein*. Although the contents of many web searches are akin to messaging contents, others might be analogized as more like locational web visits.

In my view, this complication suggests a need for caution about weighing in the lawfulness of reverse keyword warrants. Let me float one possibility, however. Perhaps reverse keyword warrants should be deemed constitutional only when they resemble queries for locational web visits instead of queries for private messages. Perhaps the use of reverse keyword searches should be limited to a few categories of locational queries.

The address searched in *Seymour* might be an example. Recall that, in *Seymour*, the government was trying to find an arsonist.[217] The warrant ordered a query of the house address where the crime occurred in the weeks before the crime, on the apparent theory that the arsonist might be looking for directions or to virtually "case" the target location.[218] A reverse keyword search for a physical address is more likely to pick up a narrow and specific kind of locational query, rather than the equivalent of private messages or a person's virtual papers. One approach might be to allow such address-based

[216] 536 P.3d 1260, 1268 (Colo. 2023); *see supra* notes 91–109 and accompanying text (discussing the facts of *Seymour*).

[217] 536 P.3d at 1268.

[218] *See id.*

warrants for search queries as constitutional, while deeming most other warrants for search queries as out of bounds.

B. Applications to Artificial Intelligence

Up to this point, I have analyzed data scans with a mechanistic filter. The filter setting was on or off, looking for a particular sequence of zeros and ones and reporting when an exact match was found. I took this approach because that is the primary scenario in known applications of the scanning question in Fourth Amendment law. Some may wonder: how would the filter-focused approach apply to queries using artificial intelligence?[219]

Consider the following hypothetical. Imagine that a query directs a computer to analyze a large database for patterns.[220] To make the question concrete, say there is a government database of known crime scenes in the last year as well as the known locations of everyone in a particular city in that year based on cell site location records. Using artificial intelligence techniques, a government agent might ask a computer to devise a list of suspected criminal conspiracies, or perhaps to list people to investigate further for potential involvement in crime. What would the Fourth Amendment implications be for such a query?

In general, the filter-focused approach does not yield a distinct insight in this setting. The reason is that machine learning methods generally look for patterns in data in a probabilistic sense.[221] When a computer uses artificial intelligence methods to look for patterns, it usually starts with some training data and develops conclusions about likely associations. It can then look at a new dataset and use lessons from the training data to reach tentative assessments about what relationships might be in the new dataset.[222] Exactly what effect any particular piece of information in the new dataset might have on the program's assessment is generally hard or impossible to know.[223] But the machine

---

[219] *See, e.g.*, Cuyler Dunn, *In Federal Lawsuit, Students Allege Lawrence School District's AI Surveillance Tool Violates Their Rights*, LAWRENCE TIMES (Aug. 1, 2025), https://lawrencekstimes.com/2025/08/01/usd497-gaggle-lawsuit-filed/ [https://perma.cc/98FX-YX6E] (reporting on a lawsuit filed by students challenging use of an "AI tool, called Gaggle, [that] sifts through anything connected to the [school] district's Google Workspace—which includes Gmail, Drive and other products—and flags content it deems a safety risk").

[220] *See, e.g.*, Benjamin Sanchez-Lengeling, Emily Reif, Adam Pearce & Alexander B. Wiltschko, *A Gentle Introduction to Graph Neural Networks*, DISTILL (Sep. 2, 2021), https://distill.pub/2021/gnn-intro/ [https://perma.cc/3A9J-8ZVD].

[221] *Probabilistic Models in Machine Learning*, GEEKSFORGEEKS (July 23, 2025), https://www.geeksforgeeks.org/probabilistic-models-in-machine-learning/ [https://perma.cc/VSS3-NZ4Y].

[222] *See* Danilo Bzdok, Martin Krzywinski & Naomi Altman, *Machine Learning: A Primer*, 14 NATURE METHODS 1119, 1119 (2017) ("The utility of ML algorithms is typically assessed empirically by determining how accurately and reliably extracted patterns generalize to new observations.").

[223] *See* Wojciech Samek, Thomas Wiegand & Klaus-Robert Müller, *Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models*, 1 INT'L TELECOMMS.

learning process is capable of being influenced by all of the data in the new dataset.[224]

From this perspective, a query using artificial intelligence is not using a filter at all. The computer does not check for a match, ignoring data that does not meet the exact criteria. Rather, the computer treats the query as the subject that the data analysis is meant to address. The computer attempts to find query-relevant trends in the data, making informed guesses about what the data might yield when factoring the dataset as a whole. Although it might resemble a filter scan in a superficial sense, the technology works very differently. This Article's approach does not apply.

C. Responding to the Provider-as-Filter Approach

A last question to consider is what I will call the provider-as-filter approach, an alternative solution to the scanning problem that was endorsed recently by the majority of judges on the Fourth Circuit in *United States v. Chatrie*.[225] I am skeptical that the provider-as-filter approach offers the best resolution to the scanning problem. This section explains why.

As the earlier discussion of *Chatrie* suggested, the provider-as-filter approach posits that what the provider passes on to the government, rather than what the provider sees, is what matters for Fourth Amendment purposes. Under this view, "the right question is what information [the provider] gave to the government, not what data [the provider] perused to find that information."[226] It is a "mistake" to focus on the provider's "digging through its own data, most of which it never turned over."[227] A Fourth Amendment search occurs "only when *the government* accesses the requested information, not when a company begins looking through its internal database."[228]

The provider-as-filter approach may lead to similar results as the filter-focused approach, at least when the government orders a provider to run a query through the provider's database. If only the data coming over the transom from the provider is "searched" under the Fourth Amendment, then what or how much data is scanned through becomes irrelevant. At the very least, the

---

UNION J. 39, 39 (2017), https://www.itu.int/dms_pub/itu-s/opb/journal/S-JOURNAL-ICTF.VOL1-2018-1-P05-PDF-E.pdf [https://perma.cc/97DR-STVK] ("[I]t is not clear what information in the input data makes them actually arrive at their decisions.").

[224] *See* Zayd Hammoudeh & Daniel Lowd, *Training Data Influence Analysis and Estimation: A Survey*, 113 MACH. LEARNING 2351, 2352–53 (2024), https://doi.org/10.1007/s10994-023-06495-7 [https://perma.cc/UXV8-FPJE].

[225] 136 F.4th 100, 130 (4th Cir. 2025) (en banc) (Richardson, J., concurring, joined by Wilkinson, Niemeyer, King, Agee, Quattlebaum & Rushing, JJ.), *cert. granted*, No. 25-112, 2026 WL 120676 (U.S. Jan. 16, 2026).

[226] *See id.* at 139 n.17 (emphasis omitted).

[227] *Id.*

[228] *Id.* at 155 (Berner, J., concurring).

approach also leads properly to the rejection of the database approach adopted by the Fifth Circuit in *Smith*.

Despite these similarities, I am skeptical that the provider-as-filter approach offers the best resolution to the scanning problem. This is a complex area, and my views are tentative. But I have two concerns, one formal and the other more substantive. The formal objection is that, under the state action doctrine, there is no neat line between "the government" and "the provider" when complying with court orders. When a provider responds to a court order for records, the provider becomes a state actor for Fourth Amendment purposes.[229] It is tricky to draw a constitutional line between the two when both have the same constitutional status. It may be a useful fiction to treat the provider as outside the Fourth Amendment, in that it will tend to recreate the results of the filter-focused approach. But it seems like a legal fiction.

Second, I worry that the provider-as-filter approach may allow the government to induce Fourth Amendment violations without being held to account for it. Imagine the government orders a provider to read a suspect's incoming and outgoing email without a warrant, and, if a suspect makes incriminating statements, to let investigators know. Every day, an employee at the provider reads the suspect's email for the government and decides whether that day's messages count as incriminating.

Under the Fourth Circuit's provider-as-filter approach, this practice presumably could go on without any Fourth Amendment oversight as long as no incriminating statement has been made, and investigators have not yet been tipped off. After all, no information has been turned over to the government, at least yet. It seems strange to deny the government's constitutional violation in this scenario, however. The government ordered a private party to read a person's private email. The fact that the government hasn't yet benefited from that violation does not mean it didn't happen.

CONCLUSION

Technology poses constant challenges for Fourth Amendment law. As new technologies come, courts must adapt old principles to new facts. The Fourth Amendment rules for querying digital databases is one such challenge. In physical spaces, human senses make observations that are searches. The human eye sees, and exposing private items to human observation counts as the quintessential search. When we shift from human observation to computer

[229] *See* Commonwealth v. Gumkowski, 167 N.E.3d 803, 812 (Mass. 2021) ("[I]f law enforcement instigates the search by contacting the cell phone company to request information, there is State action. That Sprint could have refused to provide records in response to [the state trooper's] request does not change the fact that he instigated the search." (citing Commonwealth v. Brandwein, 760 N.E.2d 724, 731 (Mass. 2002))). I further develop this argument in *The Fourth Amendment Limits of Internet Content Preservation*. *See* Kerr, *Fourth Amendment Limits*, *supra* note 47, at 779–82.

scanning, however, physical analogies are thrown into doubt. Does the computer filter "see," such that all the data scanned is searched? Alternatively, is the output file all that is seen? Or does the human observer determine what is seen, based on what information was actually revealed by the output file in light of the filter and the database?

The path forward is the last of these options. Despite the technological novelty of the question, the fundamentals of the Fourth Amendment remain the same. As always, it is exposure to people that matters. Courts should not be dazzled by the size of databases that are scanned through, but whose secrets are not exposed to human observation. What computers calculate, they do not necessarily see. Whether agents are looking at the screen of a thermal imaging device, the alert of a drug detection dog, or the output file of a database scan, it is the protected information that the surveillance exposes that matters. A search occurs when protected data inside the database is revealed to the outside observer from the output based on the filter setting. Although courts have divided on the question, the filter-focused approach is the correct one. The Colorado Supreme Court and the Supreme Court of Georgia have it basically right, and the Fifth Circuit has it wrong.