No. 18-50440

IN THE

United States Court of Appeals for the Ninth Circuit

UNITED STATES OF AMERICA,
Plaintiff-Appellee,

v.

LUKE WILSON
Defendant-Appellant.

On Appeal from the United States District Court
for the Southern District of California
Case No. 15-cr-02838-GPC
District Judge Gonzalo P. Curiel

BRIEF FOR *AMICI CURIAE* GOOGLE LLC AND FACEBOOK, INC. IN SUPPORT OF PLAINTIFF-APPELLEE AND AFFIRMANCE

RYAN T. MRAZIK
ERIN K. EARL
RACHEL A.S. HANEY
PERKINS COIE LLP
1201 Third Avenue, Suite 4900
Seattle, WA 98101-3099
Telephone: (206) 359-8000
Fax: (206) 359-9000

Counsel for Amici Curiae

June 28, 2019

CORPORATE DISCLOSURE STATEMENT

Google LLC is a wholly owned subsidiary of XXVI Holdings, Inc., which is a wholly owned subsidiary of Alphabet Inc., a publicly traded company. No publicly held company owns 10% or more of Alphabet Inc.'s stock.

Facebook, Inc. has no parent corporation and no publicly held corporation owns 10% or more of its stock.

TABLE OF CONTENTS

Page

TABLE OF AUTHORITIES

Page

CASES

TABLE OF AUTHORITIES
(continued)

Page

STATUTES

OTHER AUTHORITIES

INTEREST OF *AMICI CURIAE*[1]

Amici offer some of the most widely used Internet- and mobile-based communications, sharing, and storage products and services in the world.

Google is a diversified technology company whose mission is to organize the world's information and make it universally accessible and useful. Google offers a variety of web-based products and services—including Search, Gmail, Maps, YouTube, and Chrome—used by people everywhere.

Facebook's mission is to give people the power to build community and bring the world closer together. Through its services, Facebook enables people to stay connected with friends, family, and colleagues; to discover what's going on in the world; and to share and express what matters to them.

Every day, billions of people use amici's services to talk with family and friends, express thoughts and opinions, operate businesses, take and send videos and photos, and discover new content and information from around the world. Unfortunately, a tiny fraction of users abuse amici's services, in violation of their Terms of Service, to offer, store, and transmit child pornography.[2]

---

[1] All parties have consented to the filing of this brief. No counsel for a party authored this brief in whole or in part, and no person other than amici or their counsel has made a monetary contribution intended to fund the preparation or submission of the brief.

[2] Amici and courts sometimes refer to this material using other terms, including "child exploitation material" or "child sexual abuse images." *See, e.g.*, *Paroline v. United States*, 572 U.S. 434, 483 (2014) (Sotomayor, J., dissenting);

For decades, "the exploitive use of children in the production of pornography has [been] a serious national problem." *New York v. Ferber*, 458 U.S. 747, 749 (1982). As use of online communications has increased, the proliferation of child pornography likewise has "grown exponentially." *Paroline*, 572 U.S. at 440 (citation omitted). Amici devote substantial human and technological resources to keeping this material off their services.

One such technological resource is hash matching, an automated computer process that detects duplicates of images previously identified as apparent child pornography. Hash matching enables providers like amici to protect their services and users, independent of any reporting requirement, by reliably and efficiently detecting duplicates of images that were previously identified as apparent child pornography and removing those duplicates from their services. Amici report this material to the National Center for Missing and Exploited Children ("NCMEC") as is their duty under the federal child pornography reporting statute. 18 U.S.C. § 2258A.

ER 188 (quoting Declaration of Cathy A. McGoff). In this brief, amici use the term "child pornography" for clarity and consistency with the parties' briefs. As noted below, providers have a statutory obligation to report any apparent violation of the federal child pornography statutes, so reportable "child pornography" discussed in this brief includes material that appears to satisfy the definitions in Chapter 110 of Title 18, United States Code.

Because of their interests in safeguarding the integrity of their services, protecting their users, and keeping child pornography off of their products and services, amici have a strong interest in the outcome of this case.

SUMMARY OF ARGUMENT

The district court correctly held that the government's review of four child pornography images attached to an email in Wilson's account and identified by hash matching did not violate the Fourth Amendment.

Online service providers like amici share the broad societal interest in combating child pornography and have their own strong business interests in identifying, removing, and reporting child pornography that appears on their services and platforms, and developing and using technology to increase the efficiency, accuracy, and effectiveness of that process. Hash matching is one way that providers like amici pursue those interests. Hash matching involves calculating an alphanumeric value (a "hash value") from a specific file—in this context, an image that has previously been viewed by a human and determined to be apparent child pornography—and then identifying duplicates of that file by comparing its hash value with the hash values of unknown files. This process enables providers like amici to accurately and efficiently identify and remove from their services identical copies of previously-viewed child pornography images. And it relieves

providers' review teams of the need to review, and be exposed to, the same imagery countless times.

When a provider reports copies of such images identified using hash matching, subsequent viewing of those images by the government does not exceed the scope of the provider's initial private review; the high accuracy of hash matching means that the government's examination of the image will not reveal any information not already revealed by the provider's hash match.

ARGUMENT

The district court correctly held that Wilson's motion to suppress should be denied under the private search doctrine, and its decision can and should be affirmed on that basis. This Court should join the Fifth Circuit in affirming that law enforcement does not violate the Fourth Amendment by reviewing images identified by private companies as having hash values corresponding to previously-reviewed apparent child pornography. *See United States v. Reddick*, 900 F.3d 636, 637 (5th Cir. 2018), *cert. denied*, 139 S. Ct. 1617 (2019).[3]

[3] Because the government has not questioned Wilson's reasonable expectation of privacy in this appeal and the private search doctrine applies here in any event, the Court does not need to address the district court's conclusion that the defendant lacked a reasonable expectation of privacy, as discussed below.

A. Hash matching is a reliable, accurate, and efficient technological process for service providers to identify duplicates of child pornography files.

Service providers like amici have strong business interests in enforcing their Terms and ensuring that child pornography is not stored on their platforms. One way that providers advance their private interests in reducing the spread of child pornography online is to use hash matching technology to identify copies of files they have already viewed and reported to NCMEC. Automated technological solutions help counter the spread of child pornography online, as the volume of images grows dramatically. In 2018 alone, for example, NCMEC received more than 18.4 million reports of suspected child sexual exploitation (including apparent child pornography) through the CyberTipline. Nat'l Ctr. for Missing & Exploited Children, http://www.missingkids.org/footer/media/keyfacts.

Some providers therefore use hash matching to identify duplicates of images that a reviewer previously identified as apparent child pornography. In this context, hash matching means calculating an alphanumeric value (a "hash value") from a specific file that a reviewer identifies as apparent child pornography and then identifying duplicates of that file by comparing its hash value with the hash values of unknown files. ER 189-90. Calculating a hash value involves applying a mathematical algorithm to a piece of information. Although there are various

methods and algorithms for doing so,[4] the process, known as "hashing," has been widely used in the technology industry for many years, including to store information in data structures that allow for more efficient searches and to ensure that two files or sets of data are exact matches. *See Hash*, Microsoft Computer Dictionary 214 (4th ed. 1999); Niels Ferguson, Bruce Schneier & Tadayoshi Kohno, *Cryptography Engineering: Design Principles & Practical Applications* 77 (2010).

A hash value is "unique . . . for each offending image" and often referred to as a "digital fingerprint," ER 189, or a "digital signature." ER 189, 192 n.5; Ronald Rivest, *The MD5 Message-Digest Algorithm* (1992), http://tools.ietf.org/html/rfc1321; *see also* Ryan D. Balise & Gretchen Lundgren, *The Fourth Amendment's Governmental Action Requirement: The Weapon of Choice in the War Against Child Exploitation*, 41 New Eng. J. on Crim. & Civ. Confinement 303, 308-09 (2015). Importantly, a hash value is not a mere label or title for a file that might not accurately describe the file's content. Rather, a hash value is specific to that file

[4] For example, some hashing algorithms, such as PhotoDNA, use image-specific functions to identify, with a high degree of accuracy, duplicate and near-duplicate images—i.e., images that have been altered, potentially with the goal of escaping detection by file-based hashing algorithms. *See Reddick*, 900 F.3d at 637-38. Amici therefore disagree with EPIC's assertion that so-called "image hashing" is "fundamentally different" from file hashing—*both* "are good at achieving a near-zero percentage of false positive matches." EPIC Br. 13; *see also* ER 156-57 (agent testimony that after years of reviewing CyberTip reports from Google, none with the "A1" categorization at issue here had been incorrectly reported).

and inextricably linked to the file, bit-for-bit. *See* Richard P. Salgado, *Fourth Amendment Search and the Power of the Hash*, 119 Harv. L. Rev. F. 38, 39 (2005).

Because a hash value can be calculated only for a specific file and not for features in a general category of images (such as images showing sexual activity), providers seeking to identify and remove child pornography from their services can match files on their services only against calculated hash values for images that have already been identified as apparent child pornography. Here, for example, after each offending image "is viewed by at least one Google employee, it is given a digital fingerprint ('hash') that [Google's] computers can automatically recognize and is added to [Google's] repository of hashes of apparent child pornography as defined in 18 U.S.C. § 2256." ER 79.

Then, because the calculated hash value is specific to each image whose hash value was included in the data set, a service provider can use the hash value to identify duplicates of that image. *See* Salgado, 119 Harv. L. Rev. F. at 40 ("[I]f [the] unknown file has a hash value identical to that of [the] known file, then you know that the first file is the same as the second."). For example, the district court found that Google's product abuse detection system recognized four images attached to Wilson's email message as apparent child pornography by calculating the hash value for each image and comparing it to its repository of hash values for apparent child pornography files. ER 191, 204.

Using hash matching to identify duplicates of previously-reviewed child pornography is effective and accurate. Many of the images of child pornography proliferating on the Internet are duplicates of preexisting images. *See, e.g.*, *Paroline*, 572 U.S. at 440-41. Hash matching identifies duplicates of apparent child pornography files more reliably and efficiently than humans, who cannot locate or review content at the rate of an automated computer program and cannot detect duplicates of files as accurately as a computer program. *See* Salgado, 119 Harv. L. Rev. F. at 41.

Using hash matching also relieves providers' review teams of the need to review, and be exposed to, the same imagery countless times. And hash matching provides these benefits without incurring any decrease in accuracy. In its amicus brief supporting Wilson, the Electronic Privacy Information Center claims that hash matching has three sources of potential inaccuracy: (1) human error in the original identification; (2) error in hash matches received from another entity; and (3) false positives. EPIC Br. 11-12. None is persuasive.

First, any potential for human error has nothing to do with hash matching but would be presented equally by *any* form of provider reporting—humans are as likely to make mistakes identifying an image they review personally as they are in identifying an image that is added to a hash database used to automatically identify duplicate images. As such, any potential for human error is immaterial to the legal

issue here, as it has no impact on the scope of the private or governmental search. Further, any risk is low because Google personnel are "trained by counsel on the federal statutory definition of child pornography and how to recognize it on [Google's] products and services." ER 79.

Second, any potential for erroneous matches to hash values received from another entity is nonexistent here, where the record shows that Google relied on its own repository of hashes generated from images its own team had previously reviewed. *See* ER 79-80.

Finally, the risk of false positives is negligible for any industry-standard hashing algorithm. Accuracy in hash matching relies on the uniqueness of the hash value, which depends upon the specific hashing algorithm used. *See* Ferguson, Schneier & Kohno, *supra*, at 78-79; Larry J. Hughes, Jr., *Actually Useful Internet Security Techniques* 54-55 (1995).[5] For any industry-standard algorithms, there is at most a vanishingly small risk of a false positive being reported. *See, e.g.*, Salgado, 119 Harv. L. Rev. F. at 39 n.6; Neth. Forensic Inst. Ministry of Justice &

[5] EPIC complains that "neither Google nor the federal agency has revealed the specific nature of the underlying algorithm" or "established the accuracy, reliability, and validity of this technique." EPIC Br. 2. But providers should not be compelled to provide detailed information about the operation of any proprietary technology they may use to identify and remove duplicates of apparent child pornography from their platforms, at the risk of enabling evasive maneuvers by those who spread such material and its further proliferation, nor should providers be restricted to using hashing algorithms that have been publicly disclosed.

Sec., Technical Supplement - Forensic Use of Hash Values and Associated Hash Algorithms 6 (Jan. 2018), http://www.forensicinstitute.nl/binaries/forensicinstitute/documents/publications/2018/02/13/forensic-use-of-hash-values-and-associated-hash-algorithms/Supplement-hashes-v2018_01a_English.pdf (each of the three hashing functions tested had a false positive risk of "almost zero").

In sum, with billions of users sending tens of billions of communications through amici's services, hash matching is a reliable and accurate automated process for identifying duplicates of previously identified child pornography images, and is the best and most realistic means for service providers to be able to protect their users and services from child pornography.

B. Government review of an image of child pornography that has been identified through hash matching does not violate the Fourth Amendment.

Nearly all other courts to consider the issue have concluded, like the district court here, that government review of a duplicate image of previously-reviewed child pornography identified through hash matching does not violate the Fourth Amendment. *See, e.g.*, *Reddick*, 900 F.3d at 639; *United States v. Ringland*, No. 8:17CR289, 2019 WL 77276, at *6 (D. Neb. Jan. 2, 2019); *United States v. Miller*, No. 16-47-DLB-CJS, 2017 WL 2705963, at *5-6 (E.D. Ky. June 23, 2017), *appeal docketed*, No. 18-5578 (6th Cir. June 5, 2018). *But see United States v. Keith*, 980 F. Supp. 2d 33, 43 (D. Mass. 2013) (reaching the contrary conclusion where,

unlike here, "the provenance of that designation [of the original file as child pornography] is unknown"). These courts' analyses are sound: binding Supreme Court precedent dictates an affirmance in this case.

1. When a private entity conducts a search and informs the government of what it finds, a government agent may repeat the search without violating the Fourth Amendment.

When a private entity conducts a search, it may inform the government of what it has found, and "the Fourth Amendment does not prohibit governmental use of that information."[6] *United States v. Jacobsen*, 466 U.S. 109, 117 (1984). In other words, the actions of a private entity in making "an examination that might have been impermissible for a government agent cannot render otherwise reasonable conduct unreasonable." *Id.* at 114-15. When a government agent reviews or conducts another search based on information provided to it by the private entity, any "additional invasions of . . . privacy by the government agent must be tested by the degree to which they exceed[] the scope of the private search." *Id.* at 115; *see also United States v. Tosti*, 733 F.3d 816, 821-22 (9th Cir. 2013). When a government agent merely repeats an initial private review, no "additional invasion" of privacy occurs, and the government agent does not violate the Fourth Amendment.

[6] Amici assume for purposes of this case that hash matching can constitute a "search" under the Fourth Amendment.

The Supreme Court's decision in *Jacobsen* establishes the standard for determining when a government agent's subsequent search is within the scope of an initial private search. In *Jacobsen*, FedEx employees opened both a package and a tube inside the package to discover plastic bags, the innermost of which contained white powder that the FedEx employees identified as cocaine. *See* 466 U.S. at 111. They turned the package over to the DEA. *Id.* The Court held that the DEA agent's subsequent warrantless search of the package did not violate the Fourth Amendment because the agent did not exceed the scope of FedEx's private search. *Id.* at 125-26. Instead, the agent merely confirmed what the FedEx employees had told him, and there was a "virtual certainty" that he would find contraband and little else within the package. *Id.* at 118-20. The Court reasoned that the agent had not violated the Fourth Amendment by "viewing . . . what a private party had freely made available for his inspection." *Id.* at 119.

2. The district court correctly held that the agent's review of an image of child pornography was within the scope of Google's initial private review.

As the district court explained, not only is this case controlled by *Jacobsen*, but "[t]he facts in this case are even stronger." ER 205. Applying *Jacobsen*, the district court correctly determined that the Fourth Amendment did not prohibit the agent from reviewing the four images that Google reported to NCMEC after it had identified them, using hash matching, as duplicates of child pornography images

Google had previously viewed. ER 204-05. Because the detective did not exceed the scope of Google's review but merely "view[ed] . . . what a private party had freely made available for [its] inspection," the government's review did not implicate the Fourth Amendment. *Jacobsen*, 466 U.S. at 119.

Wilson contends instead that this case is controlled by *Walter v. United States*, 447 U.S. 649 (1980), but his reliance on that case is misplaced. Wilson Br. 37-40. In *Walter*, a private carrier misdelivered a set of packages, which the recipients opened and saw contained film boxes. 447 U.S. at 651-52. The recipients did not view the films, but after seeing "suggestive drawings" and "explicit descriptions of the contents" on the outside of the boxes, they contacted the FBI. *Id.* at 652. The FBI then viewed the films without obtaining a warrant. *Id.* The Supreme Court held that the FBI had violated the Fourth Amendment by exceeding the scope of the initial private search. The controlling opinion emphasized that "the private party had not actually viewed the films" and "[p]rior to the Government screening one could only draw inferences about what was on the films." *Id.* at 657. Therefore, "[t]he projection of the films was a significant expansion of the search that had been conducted previously by a private party." *Id.*

Reading *Walter* and *Jacobsen* together, two "critical measures" determine "whether a governmental search exceeds the scope of the private search that preceded it"—"how certain [the government] is regarding what it will find . . .

when it re-examines the evidence" and "how much information the government stands to gain." *United States v. Lichtenberger*, 786 F.3d 478, 485-86 (6th Cir. 2015). In this case, those factors make clear that the district court was correct to conclude that Special Agent Thompson did not exceed the scope of Google's private review.

First, when Special Agent Thompson viewed the image files reported by Google, there was a virtual certainty that the files would contain nothing other than apparent child pornography. *See Jacobsen*, 466 U.S. at 119-20. As the district court correctly noted, because Google had "previously confirmed that each of the four images in Defendant's email was child pornography," and Google's hash matching process only identifies duplicates of such previously-viewed apparent child pornography files, Special Agent Thompson had "even more of a 'virtual certainty'" that his review would reveal apparent child pornography files that Google personnel had previously reviewed. ER 205.

That extremely high level of certainty distinguishes this case—and providers' use of hash matching in general—from *Walter*. The private employee in *Walter* viewed only the outside of the film boxes, not the films themselves, and the labels and imagery on the film boxes allowed a person only to "draw inferences about what was on the films." 447 U.S. at 657; *see also Tosti*, 733 F.3d at 823

(distinguishing *Walter* because "the content of the films in *Walter* was not apparent from the private inspection").

Here, by contrast, after hash matching the files' *contents*, Google knew what the files were: duplicates of images that a person had previously reviewed and identified as apparent child pornography. A hash match identifying a duplicate is not a mere label on a canister, which can be subjective or inaccurate. Instead, a hash value is a unique, objective, reliable, and accurate identifier for an image file that identifies duplicates, without any need for human inference or interpretation, and without the possibility of human error or misdescription.

Second, because Special Agent Thompson could be virtually certain that the reported images were apparent child pornography, he stood to gain little or no additional information through his review. As a human, Special Agent Thompson had to view the files to confirm their content. But he already knew what he would find: images that Google identified as duplicates of apparent child pornography it had previously viewed. In *Jacobsen*, the DEA agent's search of the box and tube inside was not an additional search under the Fourth Amendment because "a manual inspection of the tube and its contents would not tell him anything more than he already had been told" by FedEx. 466 U.S. at 119. Just so here.

As the district court found, "[a]t least one Google employee had previously viewed each of the four child pornography images Defendant uploaded to his

account." ER 204. That Google's subsequent identification occurred through hash matching does not mean the detective expanded the scope of Google's private review. In *United States v. Tosti*, for example, this Court held that a government agent did not violate the Fourth Amendment when he enlarged images previously identified as child pornography by a private computer technician who had viewed the images only as thumbnails. 733 F.3d 816. This Court explained that the police "did not exceed the scope of [the private] search because" both the police and the private technician "testified that they could tell from viewing the thumbnails that the images contained child pornography," so "the police learned nothing new through their actions." *Id.* at 822. So too here: because "Google had previously confirmed that each of the four images in Defendant's email was child pornography[,] . . . SA Thompson's viewing of the four images allowed SA Thompson to 'learn[] nothing new.'" ER 206 (quoting *Tosti*, 733 F.3d at 822).

Wilson's arguments against application of the private search doctrine—that hash matching technology is akin to a technological dog sniff, and that hash matching involves no human review—are not persuasive. Both arguments fail because hash matching identifies *only* images that are duplicates of images that a Google employee already has personally reviewed. Unlike a dog sniff, Google's hash matching does not identify "the presence of a specific type of material, in this case an image file suspected of being contraband." Wilson Br. 33. Rather, hash

matching can identify *only* duplicates of an image that *a person* previously identified as apparent child pornography—any other apparent child pornography file will not be detected. Accordingly, it is a *human* who reviews the image and a *human* who determines whether the image appears to qualify as child pornography. The scope of the review by a Google reviewer and the government agent are exactly coextensive—a paradigmatic scenario for application of the private search doctrine.[7]

In sum, *Jacobsen*'s "virtual certainty" standard is met here. "Virtual certainty" need not be *absolute* certainty—in *Jacobsen*, the field test could have revealed that the white powder was baking powder and not cocaine. But where, as here, the chances of the images being anything other than child pornography were vanishingly small, and because the agent did not open the email itself or any other images, the government did not exceed the scope of Google's private review.

[7] Wilson also argues that the private search doctrine is inapplicable because he is raising a property-based Fourth Amendment argument under *United States v. Jones*, 565 U.S. 400 (2012). But, as the district court observed, *Jones* did not overrule *Jacobsen*, ER 205 n.8, and the decision was based on "the Government's *physical* trespass of the vehicle." *Carpenter v. United States*, 138 S. Ct. 2206, 2215 (2018) (emphasis added) (citing *Jones*, 565 U.S. at 404-05). *Jones* therefore supplies no basis for declining to apply *Jacobsen*, which remains binding Supreme Court precedent, particularly in the context of reviewing electronic data and not physically trespassing on tangible property.

C. The Court need not and should not reach the reasonable expectation of privacy issue.

This Court need not and should not adopt the district court's conclusion that Wilson lacked a reasonable expectation of privacy in the files he uploaded to his Gmail account. ER 198. Affirmance is appropriate because the district court correctly held that Wilson's motion to suppress should be denied under the private search doctrine. A user's reasonable expectation of privacy in email is not defeated by a provider's ability to access its content or by a service provider's Terms of Service for the reasons explained in the Brief of Amici Curiae Electronic Frontier Foundation & American Civil Liberties Union Foundation. *See* EFF & ACLU Br. 10-12. Rather, the Fourth Amendment generally protects users' reasonable expectations of privacy in the contents of emails held by a third-party service provider from warrantless search and seizure by the government, irrespective of whether the service provider has terminated that user's account or whether the user violated the terms governing his relationship with the service provider. *United States v. Mohamud*, 843 F.3d 420, 442 (9th Cir. 2016), *cert. denied*, 138 S. Ct. 636 (2018); *see also Byrd v. United States*, 138 S. Ct. 1518, 1524 (2018) (drivers have a reasonable expectation of privacy in a rental car even when driving the car in violation of the rental agreement); *United States v. Warshak*, 631 F.3d 266, 286-88 (6th Cir. 2010).

CONCLUSION

The judgment of the district court should be affirmed.

Respectfully submitted.

/s/ Ryan T. Mrazik
Ryan T. Mrazik
Erin K. Earl
Rachel A.S. Haney
Perkins Coie LLP
1201 Third Avenue, Suite 4900
Seattle, WA 98101-3099
Telephone: 206.359.8000

Counsel for Amici Curiae

June 28, 2019

CERTIFICATE OF COMPLIANCE

1. This motion complies with the type-volume limitations of Federal Rule of Appellate Procedure 29(a)(5) because it contains 4,328 words, excluding the parts of the brief exempted by Federal Rule of Appellate Procedure 32(f).

2. This brief complies with the typeface requirements of Federal Rule of Appellate Procedure 32(a)(5) and the typestyle requirements of Federal Rule of Appellate Procedure 32(a)(6) because it has been prepared in a proportionally spaced typeface using Microsoft Office Word 2010 in Times New Roman 14-point font.

/s/ Ryan T. Mrazik
Ryan T. Mrazik

CERTIFICATE OF SERVICE

I certify that I electronically filed the foregoing with the Clerk of the Court for the United States Court of Appeals for the Ninth Circuit by using the appellate CM/ECF system on June 28, 2019. I certify that all participants in the case are registered CM/ECF users and that service will be accomplished by the appellate CM/ECF system.

/s/ Ryan T. Mrazik
Ryan T. Mrazik