epic.org

Electronic Privacy Information Center
1519 New Hampshire Avenue NW
Washington, DC 20036, USA

+1 202 483 1140
+1 202 483 1248
@EPICPrivacy
https://epic.org

December 4, 2020

Honorable Tani Cantil-Sakauye, Chief Justice,
and the Honorable Associate Justices of the
California Supreme Court
350 McAllister Street
San Francisco, CA 94102

Re: Rule 8.500(g) Amicus Curiae Letter in Support of Review
People v. Wilson, Supreme Court No. S265795
Fourth Appellate District, Division One, No. D074992
San Diego Superior Court No. SCD263466

Dear Chief Justice Cantil-Sakauye and Associate Justices:

This letter is submitted pursuant to rule 8.500(g) of the California Rules of Court by the Electronic Privacy Information Center ("EPIC") in support of the petition for review in the above-captioned case. EPIC urges the Court to grant the petition for review. The case implicates an important and novel issue about algorithmic evidence and the Fourth Amendment. The lower court erroneously relied on evidence about a different type of algorithm than the one at issue in this case to conclude that there was a "virtual certainty" that a Google employee had previously viewed an image that the company automatically referred to law enforcement as apparent child pornography. The erroneous ruling could have ramifications for future cases involving searches conducted by proprietary algorithms. Courts cannot simply accept bald assertions from companies and prosecutors that algorithms are infallible.

I. Statement of Interest

EPIC is a public interest research center in Washington, D.C., established in 1994 to focus public attention on emerging civil liberties issues and to protect privacy, the First Amendment, and other constitutional values. EPIC routinely participates as amicus in cases applying the Fourth Amendment to new technologies. Chief Justice Roberts cited EPIC's amicus brief in his majority opinion in *Riley v. California* (2014) 573 U.S. 373. (*Id.* at p. 396, p. 397.) EPIC has participated as amicus in two federal cases presenting a similar Fourth Amendment question as the petition for review, including the federal prosecution of the defendant here. (*See* Brief of Amicus Curiae

EPIC, *United States v. Miller* (6th Cir. Dec. 3, 2020) No. 18-5578, 2020 WL 7074226; Brief of Amicus Curiae EPIC, *United States v. Wilson* (9th Cir. filed Mar. 28, 2019) No. 18-50440.) Our amicus brief in the federal *Wilson* case persuaded Judge Watford, who sat on the Ninth Circuit panel, to question the prosecution during oral argument about the dearth of the evidentiary record on the Google algorithm. The case is still pending.

II. The lower court erred in finding that, on the record before it, there was a "virtual certainty" that Google had previously viewed the images it automatically sent to authorities.

The government may replicate a private search without first obtaining a warrant if there is a "virtual certainty" that the scope of the government search will not exceed that of the private search. (*United States v. Jacobsen* (1984) 466 U.S. 109, at p. 119.) The lower court found that such certainty existed in this case because the method Google uses to match images users upload to the service with images Google employees have previously identified as child pornography is "highly accurate." (Pet.'s Exhibit A ("Opn.") at p. 20 fn. 11). But the evidence the court relied upon does not support this conclusion. Google's declaration does not adequately describe how it matches images to each other, nor does it contain evidence that the method has a low false positive rate. Other authorities cited in the opinion are irrelevant to the type of algorithm Google uses to match images because they describe *file* hashing techniques that Google does not claim to use.

Google uses a proprietary algorithm to scan every file uploaded to its services. (1CT 196.) In line with its proprietary designation, Google has provided hardly any information about how its algorithm works. A Google employee's declaration is the only evidence in the record describing the algorithm. (*Ibid.*) Google says that each image added to its repository of apparent child pornography is assigned a "hash," or hash value, which is a "digital fingerprint." (*Ibid.*) Google also says that it assigns each image uploaded to its service a hash value and then compares the hash value of the uploaded image to the hash values in the repository for any matches. (*Ibid.*) A match indicates the presence of "duplicate images." (*Ibid.*)

Google does not say how its algorithm generates a hash value for an image. This is important because there are many different ways to produce a hash value for a piece of information, and most of these methods do not result in a unique hash value. A hash value is simply a sequence of letters and/or numbers. (Bruce Schneier (1996) *Applied Cryptography* p. 30.) A hash function can use any number of parameters to assign information—like a file, or a sequence of letters and numbers—to a hash value. (*Ibid.*) Programmers use hashing to authenticate information and to make searches

faster. The most robust hash functions are used in cryptography to quickly and securely verify that two files are the same without having to directly compare each bit in each file. Cryptographic file hashing techniques produce a one-to-one correspondence between a file and a hash value; each file is assigned its own, unique hash value. (*Ibid.*) Changing one bit in a file will change the hash value for a file hashed using a cryptographic technique, which makes these hash values akin to a "fingerprint" for the file. (*Id.* at p. 30–31.) But not all hash functions assign unique hash values to files. Depending on the needs of the program, a programmer may design a hash function to assign many different files the same hash value. (*Id.* at p. 30.) In fact, hash functions are typically many-to-one, not one-to-one. (*Ibid.*)

Google does not claim to use any known cryptographic hashing method, like MD5[1]—and for good reason. Using a cryptographic hash function to hash apparent child pornography files would make it exceedingly easy for criminals to avoid detection—all they would have to do is change one bit of the file to prevent a match to the same exploitative content in the repository. Aware of this problem, Microsoft developed PhotoDNA to match *similar* images *even if the files are not identical*. (Microsoft, Digital Crimes Unit, *PhotoDNA* ("PhotoDNA Slides") at p. 4.[2]) Microsoft, unlike Google, has publicly released information about how its image matching algorithm works. PhotoDNA does more than just assign hash values to files; the algorithm converts the image to grayscale, resizes the image, divides it into squares, and then assigns the image a numerical value—the hash value—based on the shading in each square. Microsoft, *PhotoDNA & PhotoDNA Cloud Service*.[3] As a result, PhotoDNA does not produce a one-to-one correspondence between image files and hash values; instead, many image files have the same hash value. (PhotoDNA Slides at p. 4.) The two files could look essentially the same to the human eye, or they could be cropped differently (and thus reveal more or less information), or one could be tinted blue and the other not. (*See Id*. at p. 5.)

Like PhotoDNA, Google says that its algorithm matches images, not whole files. (1CT 196.) To match images contained in files with different cryptographic hash values, Google's algorithm must extract certain information from the file, possibly manipulate the information, and then assign a hash value to that information using some set of criteria. The Government has not provided evidence about any step of this

---

[1] *See* Ron Rivest, *The MD5 Message-Digest Algorithm RFC 1321* (Apr. 1992).

[2] https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016802f249e (last accessed Dec. 5, 2020).

[3] https://news.microsoft.com/uploads/2016/03/photoDNACloudServiceFS.pdf (last accessed Dec. 5, 2020).

process. Depending on the methods Google uses to assign images a hash value, and how different an image file can be from another image file and still have the same hash value, the image a Google employee previously viewed could look different than the image detected and automatically reported by its algorithm. Only after obtaining further information about the Google algorithm could the court determine whether the files are similar enough that a Google employee could be fairly said to have previously viewed the matched file. For instance, if Google crops images before calculating their hash values, two images with matching hash values could have very different backgrounds that reveal very different information.

Because Google does not use a cryptographic method to assign each file a unique hash value, it was improper for the lower court to rely on evidence about file hashing to make conclusions about the accuracy of Google's algorithm. (Opn. at p. 6 (citing the Salgado article for the assertion that Google's hashes are "unique to the computer file"); *Id.* at p. 20 fn. 11 (citing the Salgado article for the statement that "no two files will have matching values 'except a file that is identical, bit-for-bit.'")) The Salgado law review article describes cryptographic methods to hash *files* and nowhere mentions image matching techniques like the one Google and Microsoft use that assign many different files to the same hash value. (Richard P. Salgado (2005) *Fourth Amendment Search and the Power of the Hash*, 119 Harv. L. Rev. 38.) Several other courts have erroneously relied on this article to conclude that image matching techniques that use hashing are the same as cryptographic file hashing techniques. (*See* Opn. at p. 20 fn. 11 (citing *United States v. Reddick* (5th Cir. 2018) 900 F.3d 636 for the conclusion that courts describe hash matching as "a highly accurate technology.") But these courts were mistaken, and the lower court erred in relying on them. Similarly, the lower court's reliance on a Department of Homeland Security agent's description of cryptographic file hashing techniques in general, and the MD5 method in particular, are irrelevant because Google does not use the MD5 hash function or any other cryptographic hash function. (See Opn. at p. 10 (quoting the agent for various assertions about the accuracy of *file* hashing using cryptographic techniques, including a "commonly-used hash value algorithm," MD5)). The Government must present evidence about *Google's* algorithm, not other, unrelated hashing methods.

The only evidence the Government has presented about the Google algorithm's accuracy is Google's assertion that their method produces a "digital fingerprint" for each image. (1CT 196.) But an analogy to another forensic method does not prove an algorithm's accuracy. The Government must present evidence about the algorithm's error rate and tolerance for differences in images to demonstrate that an employee had previously viewed the image with a "virtual certainty."

Without evidence of the algorithm's accuracy, the hash value is little more than a label. (*See Walter v. United States* (1980) 447 U.S. 649.) A label is not a reliable indicator of the contents of a container because anyone can label a container whatever they like. A hash function, in its most basic form, simply assigns a label to information using whatever criteria a programmer likes. The labels in *Walter* were not enough to justify a search of the contents of the film canisters because the private party had not previously viewed the films themselves and the labels only allowed the Government to draw inferences about the content of the films. (*Id.* at 657). Similarly, the reliability of a hash value match depends entirely on the accuracy and tolerance for difference of the underlying algorithm. If the algorithm has a significant false positive rate or matches images that have different background content, the match may be no better an indication of the contents of the file than the film canister labels in *Walter*. Without knowing more about how Google's algorithm works, the only conclusion that can be drawn from a hash value match between two images is that Google *may* have previously viewed the image and identified it as apparent child pornography. To conclude that Google *has* viewed the image with a "virtual certainty," the Government must provide more information about Google's algorithm than is currently in the record.

Accordingly, we ask this Court to grant review and decide the Fourth Amendment question presented in the petition.

Respectfully submitted,

/s/ Alan Butler
Alan Butler (State Bar No. 281291)
Interim Executive Director
Electronic Privacy Information Center
1519 New Hampshire Avenue NW
Washington, D.C. 20003
(202) 483-1140
butler@epic.org

PROOF OF SERVICE

I, Alan Butler, declare under penalty of perjury that I am a member of the State Bar of California and over 18 years of age. My business address is 1519 New Hampshire Avenue NW, Washington, D.C. 20003, and my email is butler@epic.org.

On December 4, 2020, I served true and correct copies of the attached **Rule 8.500(g) Amicus Curiae Letter in Support of Review—*People v. Wilson*, Supreme Court No. S265795** by TrueFiling as follows:

Court of Appeal, Fourth Appellate, Division One

Office of the Attorney General
sdag.docketing@doj.ca.gov

San Diego District Attorney's Office
DA.Appellate@sdcda.org

San Diego Superior Court
Appeals.Central@SDCourt.ca.gov

Counsel for Petitioner
chuck@charlessevilla.com

/s/ Alan Butler
Alan Butler (State Bar No. 281291)
Interim Executive Director
Electronic Privacy Information Center
1519 New Hampshire Avenue NW
Washington, D.C. 20003
(202) 483-1140
butler@epic.org