Case No. D073943

IN THE COURT OF APPEAL OF THE STATE OF CALIFORNIA
FOURTH APPELLATE DISTRICT, DIVISION ONE

The People of the State of California,

Plaintiff and Petitioner,

v.

The Superior Court of the State of California, San Diego County,

Respondent.

Florencio Jose Dominguez,

Real Party in Interest and Defendant.

On Appeal from San Diego County Superior Court,
Case No. SCD230596
The Honorable Charles Rogers, Judge

Brief of *Amici Curiae* American Civil Liberties Union and American Civil Liberties Union of San Diego and Imperial Counties In Support of Real Party in Interest Seeking Dismissal

---

Bardis Vakili (SBN 247783)
American Civil Liberties Union Foundation of San Diego and Imperial Counties
2760 Fifth Ave #300
San Diego, CA 92103
T: 619.232.2121
bvakili@aclusandiego.org

Vera Eidelman (SBN 308535)
Andrea Woods
Brett Max Kaufman
Brandon Buskey
Rachel Goodman
American Civil Liberties Union Foundation
125 Broad Street, 18th Floor
New York, New York 10004
T: 212.549.2500
veidelman@aclu.org

Attorneys for Amici Curiae

TABLE OF CONTENTS

TABLE OF AUTHORITIES

Page(s)

Cases

Rules

Other Authorities

INTEREST OF *AMICI CURIAE*[1]

The American Civil Liberties Union ("ACLU") is a nationwide, nonprofit, nonpartisan organization with nearly two million members dedicated to defending the civil liberties guaranteed by the Constitution. The ACLU of San Diego and Imperial Counties is a regional affiliate of the ACLU which serves San Diego and Imperial counties. The ACLU and the ACLU of San Diego and Imperial Counties have appeared in numerous cases, both as direct counsel and as *amici*, before courts in California and throughout the nation in cases involving the meaning and scope of the rights of criminal defendants and the legal limitations on the use of technology by police and prosecutors.

[1] Pursuant to California Rules of Court 8.360(f) and 8.200(c), counsel for *amici curiae* have submitted a motion for leave to file this brief. In addition, counsel for *amici curiae* certify that no party or counsel for a party authored this brief in whole or in part, and no person other than *amici curiae*, their members, or their counsel made a monetary contribution to its preparation or submission.

1. INTRODUCTION AND SUMMARY OF ARGUMENT

A business's private property interests cannot eclipse the constitutional rights of accused individuals and the public. In this case, the State seeks to keep secret the software, source code, user manual, software updates, and internal validation records of STRmix—precisely the materials the prosecution has used to generate DNA evidence that will form the bulk of the government's case at trial. The respondent court correctly recognized that allowing such secrecy would violate the Constitution and properly exercised its broad authority to issue the discovery order now on appeal. Granting the State's writ at this stage would violate the defendant's right to due process, and it would necessarily implicate the rights to confront witnesses and to a fair and public trial if this case proceeds to trial.

STRmix, the technology at the center of this appeal, aims to solve a difficult problem: to analyze "complex" DNA samples, which are so-called because they encompass genetic material from multiple contributors. The precise number of contributors, and how much genetic material came from each one, is typically unknown and, adding to the complexity, the genetic material derived from evidence-gathering is often degraded. To analyze such samples and do what a traditional DNA test cannot do, STRmix relies on an algorithm that combines forensic science, genetics, statistics, and probabilistic programming to produce a seemingly simple score called a "likelihood ratio."

The supposed simplicity and objectivity of the likelihood ratio is belied by the many choices, cognitive biases, and plain-old mistakes that STRmix's programmers have almost certainly embedded within the many lines of computer "source code" that drive the program. Those choices and errors, both known and not, can cause—and in documented cases, have

caused—STRmix to produce wildly different results from programs purporting to calculate the same match statistic for the same suspect and crime scene sample. Moreover, access to source code has exposed serious flaws in other previously accepted probabilistic genotyping programs, as well as other algorithms used as evidence in criminal trials.

Given this, and the centrality of the STRmix test results to the State's case against Mr. Dominguez, the Fourteenth Amendment requires that Mr. Dominguez now be given access to the STRmix source code. The Fourteenth Amendment guarantees him due process and, in tandem with the Sixth, a fundamentally fair trial, including adversarial testing of the State's evidence. If those rights are violated now, Mr. Dominguez's Sixth Amendment rights to confrontation and a public trial will be implicated at trial. Access to the STRmix source code, user manual, software updates, and internal validation records, which Mr. Dominguez sought and the trial court properly granted him, will permit his counsel and experts to inspect the program to uncover its potential flaws and biases, and to meaningfully confront the human choices behind the algorithm at trial.

This adversarial process is necessary to properly inform the trial court of whether or not to admit the STRmix results into evidence and, if so, to inform the jury of what weight to assign them. By granting him access to the source code, the trial court protected Mr. Dominguez's constitutional rights—and this Court should do the same.

This appeal implicates the rights of the public as well. Though not yet at issue in this case, the First Amendment right of access guarantees public oversight of criminal trials to ensure that the State exercises its prosecutorial power fairly and with integrity, and that the public trusts the criminal justice system. The right of access plainly attaches to algorithmic

source code that plays a critical role in establishing a defendant's culpability at trial. The trial court's vindication of Mr. Dominguez's rights, which would allow the source code to become part of the record, is the first necessary step in allowing the public to exercise its constitutionally guaranteed oversight function in this case.

For these reasons and those given below, this Court should dismiss this petition and reinstate the trial court proceedings, including the discovery order.

2. BACKGROUND

STRmix, the technology at issue here, purports to do what traditional DNA testing cannot accomplish. Indeed, in this case, the prosecution tasked STRmix with identifying the perpetrator of a crime after traditional methods repeatedly failed to generate data that was conclusive or convincing to a jury.

Specifically, STRmix claims to be able to identify the perpetrator of a crime from a tiny, degraded DNA sample swimming in a mixture of multiple individuals' DNA. The problem STRmix seeks to solve is difficult. While traditional DNA analysis typically focuses on high-saturation, single-source samples—often, blood or semen collected from a crime scene—STRmix seeks to analyze samples that come from multiple contributors and are often degraded. These samples are typically "touch" samples scraped from an object multiple people have touched—for example, a purse strap, a knife handle, or, as in this case, two gloves. The precise number of contributors to such samples, as well as which specific material belongs to which contributor, is almost always unknown. And because the genetic material is often degraded or low-copy, whether data in a profile accurately reflects a genetic marker or is simply random noise may be

unclear.

This means that, while traditional DNA analysis only looks for a match to a single person's known DNA profile, STRmix must first sketch a series of profiles from the complex DNA mixture based on assumptions about the sample, including factors like how many individuals contributed to the mixture, how much of each person's DNA is present, and how old or degraded the DNA is, before looking for a match. *See* Andrea Roth, *Machine Testimony*, 126 Yale L.J. 1972, 2018–19 (2017). Essentially, traditional DNA analysis is like looking at a photograph, while STRmix's analysis is like starting with an investigator's composite sketch.

To accomplish this feat, STRmix implements an "algorithm" operationalized through "source code" to produce a "likelihood ratio." Each of these quoted terms requires elaboration.

At the most elementary level, an algorithm is a series of steps that transforms inputs into an output. *See* Thomas Cormen et al., *Introduction to Algorithms* 1 (1st ed. 1994). In essence, it is like a formula, a manual, or a recipe: a set of instructions for how to get to an end result from raw materials. "Source code" refers to the human-written instructions that tell a computer how to execute those steps.

In STRmix's case, the output or end result is a single number called a "likelihood ratio," which is computed by dividing (1) the likelihood of the crime scene evidence if the accused individual is included as a contributor, by (2) the likelihood of the evidence if a random person is included instead. *See* William C. Thompson et al., *Forensic DNA Statistics: Still Controversial in Some Cases*, Legal Studies Research Paper Series No. 2013-122, 23 n.17 (Dec. 2012), *available at* https://perma.cc/J6Q6-45R2. In other words, the ratio reflects the likelihood of the evidence if the

prosecution's theory (i.e., the accused individual contributed to the DNA) is correct divided by the likelihood of the evidence if the prosecution's theory is wrong (i.e., he did not).

Unlike its output, STRmix's inputs are not fully known—and this is one of the problems at the crux of this case. Based on the record, the program decides whether something identified in a DNA sample constitutes stutter (i.e., random noise that can be ignored) or an actual allele (i.e., a characteristic that the suspect must match). Pet. Exhibit H at 98. It also appears to offer the ability to test the hypothesis that contributors are related. RPI Exhibit 5 at 429. And STRmix appears to allow analysts to choose the number of contributors to a particular sample. Pet. Exhibit H at 97. Inputs may also include assumptions about the quantity of DNA from each contributor, and the race or ethnicity or other statistical properties of the comparison population.

STRmix is used by the largest number of U.S. crime labs, but it is not the only algorithm that seeks to generate likelihood ratios from complex DNA mixtures. *See* Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 Stan. L. Rev. 1343, 1422 (2018). Other for-profit ventures include TrueAllele, which had been used in approximately 500 criminal cases by late 2016. *See* Edward J. Imwinkelried, *Computer Source Code: A Source of the Growing Controversy Over the Reliability of Automated Forensic Techniques*, 66 DePaul L. Rev. 97, 100–01 (2016). And government actors have also developed such programs, like New York's Forensic Statistical Tool ("FST").

These organizations, including government actors, have asserted a private property interest in their work. *See* Lauren Kirchner, *Traces of*

Crime: How New York's DNA Techniques Became Tainted, N.Y. Times (Sept. 4, 2017), http://nyti.ms/2vJwxze. Most focus their claims on a trade secret privilege, *see generally* Wexler, *supra*, but ESR has extended its arguments to copyright claims in this case. [2]

Not all algorithms aimed at accomplishing the same goal are identical. Indeed, due to differences in their underlying assumptions, they often differ in terms of both inputs and steps. For example, a boiled egg can be made with or without salt, can be cooked for different amounts of time, and can be cooled in running or still water or not at all. Each approach constitutes a boiled-egg-making algorithm—but, as all home cooks know well, the quality of the result may differ. Of course, algorithms used to generate a DNA match statistic are of a different order; their results can put human

---

[2] The State asserts that, pursuant to Cal. Evid. Code § 1060, ESR can claim a statutory trade secret privilege in the materials that the State refuses to disclose. *See* Pet. Br. 56–58. The statute provides that "the owner of a trade secret has a privilege to refuse to disclose the secret . . . if the allowance of the privilege will not tend to conceal fraud *or otherwise work injustice*." Cal. Evid. Code § 1060 (emphasis added). As this brief discusses in detail, refusing to disclose the purported trade secrets in this case would work injustice by robbing the defense, and ultimately the public, of the opportunity to test and confront evidence that is material to the State's case and that is reasonably likely to be necessary to a fair resolution of the case.

Moreover, in this case, the State seeks to keep not only trade secrets, but also copyrighted materials hidden from view. But California's statutory privilege does not extend to copyrighted materials and, unlike a trade secret, a copyright does not require maintaining secrecy. In addition, as discussed further in § 3(D) *supra*, access to copyrighted materials in discovery is protected by fair use. And, perhaps most compellingly, simply introducing the material in court will not give others a right to copy it; ESR can still continue to enforce its copyright if the materials are shared through discovery or even introduced in open court.

beings—including the one on trial in this case—behind bars or even render them eligible for death. But such algorithms, too, differ in their underlying assumptions, inputs, and training datasets—all things the State seeks to keep secret here. And if the underlying pieces differ, so too must the quality of their output.

To an extent, validation studies may reveal these differences. Such studies are meant to test the validity of a program under certain, defined conditions. Internal validation studies, like the ones the State refuses to disclose in this case, may reveal errors and bugs. And external validation studies, like the ones the State has not fully provided to the defense in this case, may offer additional insight because they are conducted by individuals with fresh eyes, who were not involved in building the program. But validation studies alone are not enough for effective defense review because validation studies are constrained by the specific conditions they test. For example, a radar gun that has been validated only against individual automobiles on a test driving range cannot be deemed valid for measuring the speed of a skateboarder on a busy street; it could be accurate, but the only way to know is to specifically test the machine for that use.

To fully confront and put evidence derived from STRmix to the adversarial test, access to its validation studies; underlying model; training data; source code; input parameters and data specific to each case; and any other results from which the final, reported result was chosen is necessary. As explained in further detail below, the algorithm's underlying model reflects the theory and intended process behind the probabilistic analysis, while the source code shows how that intended process has been put into practice. For example, the source code could reveal that concepts not included in the underlying model have somehow been included in the

program; that optimizations meant to, for example, minimize use of the computer's memory inadvertently change output; and that the code includes accidental mistakes. The training data constitutes the dataset on which the algorithm practiced to learn the probabilities it uses; the input parameters and data specific to each case shows the assumptions, human decisions, and raw inputs used to generate a particular likelihood ratio; and any other results calculated offer comparisons for the ultimate result communicated to the prosecution and ultimately the court, including potentially exculpatory evidence.

3. ARGUMENT

3(A) Algorithms are human constructs that include numerous sources for bias and mistake.

Algorithms are not neutral, infallible truth tellers: rather, they are tools designed, built, and employed by humans. Accordingly, they are vulnerable to human bias and mistake—and should therefore be subject to careful adversarial and judicial scrutiny—at each stage.

At the design stage, people make foundational assumptions that undergird the algorithmic model. For probabilistic DNA analysis, this includes the "thresholds for what to count as a true genetic marker versus noise." Roth at 1996–97; *see also* Pet. Exhibit H at 98 ("[STRmix] basically looks at all of the DNA peaks that were detected in the sample, determines whether they could be potential stutter peaks, allelic peaks"). In other words, humans decide at the outset of designing an algorithm what data to ignore and what data matters. Not surprisingly, the line between the two can determine whether or not a defendant is considered a statistical match to a crime scene sample. *Roberts v. United States*, 916 A.2d 922, 933–34 (D.C. 2007); *see also* Roth at 1996. Other assumptions include "the probability of unusual events—such as small amounts of contamination

during testing—that directly affect interpretation." Roth at 1996–97.

On the machine learning side, humans also impact the algorithm's design by, for example, choosing the training data—a decision that can significantly affect the algorithm's output in ways that differ for suspects of different races, ethnicities, or ancestral backgrounds. *See, e.g.*, *People v. Collins*, 15 N.Y.S.3d 564, 580–81 (N.Y. Sup. Ct. 2015) (crediting objection of two defense experts to FST because (1) it was trained on data with only "Asian, European, African, and Latino" categories, which is inadequate for identifying other races or ethnicities, and (2) the training data appeared to include only three Asian individuals, which was insufficient to determine false positive rates for people with Asian ancestry); Roth at 1997 (discussing the importance and difficulty of identifying "the appropriate reference population for generating estimates of the rarity of genetic markers"); *see also* Lauren Kirchner, *ProPublica Seeks Source Code for New York City's Disputed DNA Software*, ProPublica (Sept. 25, 2017), https://perma.cc/3GJ5-ZATJ (noting that "a Hasidic Jew, . . . is now appealing his conviction . . . [because] FST was never tested on a population as insulated as the Hasidic Jews of Williamsburg, who very likely share many of the same ancestors, and therefore much of the same DNA").

Next, at the building stage, people operationalize the algorithm—assumptions and all—through source code. Such code is built from numbers, letters, symbols, and punctuation marks, and it can be materially altered by errors or "bugs" as simple as a misplaced ampersand. *See* Christian Chessman, *A "Source" of Error: Computer Code, Criminal Defendants, and the Constitution*, 105 Cal. L. Rev. 179, 187 (2017); Roth at 1994 (quoting Sergey Bratus et al., *Software on the Witness Stand: What*

Should It Take for Us to Trust It?, *in* Trust and Trustworthy Computing 396, 397 (Alessandro Acquisti et al., eds., 2010)). The risk of bugs only increases with the complexity of the code and the difficulty of the problem it is attempting to solve. Roth at 2024. STRmix's code is likely to be affected by both issues.

To use the algorithm once they have built it, people must make choices about input parameters that can also make the difference between a conclusive and inconclusive match. For example, STRmix allows analysts to set the number of contributors to a DNA sample. Pet. Exhibit H at 97. This decision can change a likelihood ratio by several orders of magnitude. In *People v. Seepersad*, for example, the prosecution used both FST and STRmix to examine a complex DNA mixture; FST assumed that three individuals had contributed to the sample, while STRmix assumed two contributors. The resulting likelihood ratios for the same genetic sample and suspect differed by a factor of nearly 60,000. *People v. Seepersad*, 58 Misc. 3d 1227(A), 2018 WL 1163820, at *1 (N.Y. Sup. Ct. Mar. 5, 2018) (reporting a likelihood ratio of 172 million for FST and 10 trillion for STRmix). *See also* Letter from Mark W. Perlin, Chief Sci. & Exec. Officer, Cybergenetics, to Jerry D. Varnell, Conf. Specialist, U.S. Dep't of Justice, Procurement Sec,, at 3 (Apr. 1, 2015), https://www.cybgen.com/information/newsroom/2015/may/Letter_to_FBI.pdf (asserting that "STRmix can give different answers based on how an analyst sets their input parameters").

Finally, at the output stage, people must interpret the algorithm's result and translate it into terms that others can understand. Crucially, people—and not a computer or other machine—decide which results to communicate to prosecutors and, ultimately, in court. Crime labs

themselves have recognized that "you will get a different likelihood ratio every time you . . . put the same data in." *People v. Bullard-Daniel*, 42 N.Y.S.3d 714, 725 (N.Y. Cty. Ct. 2016).

At each of these stages, people—as they do—will almost certainly make mistakes. For example, with regard to the coding stage, one study found that even highly experienced programmers make a mistake in "almost 1% of all expressions contained in [their] source code." Chessman at 186–87. Mistakes occur even with tasks as simple as inputting "yes" or "no" to match a program's parameters to a particular case. *See* Wexler at 1370–71 (describing how evaluator tasked with calculating an incarcerated individual's risk score mistakenly checked "yes" in response to a question when he should have checked "no"—a mistake that had previously inflated a risk score by a full category); *see also* Erin Murphy, *The New Forensics: Criminal Justice, False Certainty, and the Second Generation of Scientific Evidence*, 95 Cal. L. Rev. 721, 775 (2007) (detailing potential mistakes in traditional DNA analysis—"a manufacturer may contaminate a kit, an analyst may fail to run positive or negative controls, or a technican may erroneously input data into a database"—all of which would also affect the results of a probabilistic genotyping algorithm).

Beyond random mistakes, people hold cognitive biases that can materially affect the variables they include in an algorithm, as well as how they interpret the results—including whether a likelihood ratio is conclusive. *See* Itiel E. Dror & Greg Hampikian, *Subjectivity and Bias in Forensic DNA Mixture Interpretation*, 51 Sci. & Just. 204, 205–07 (2011) (finding that more DNA examiners determined that an individual matched a DNA mixture when they knew that he was a criminal defendant in a gang rape case than when they did not). And, when it comes to an issue as

complex as probabilistic DNA typing embodied in source code, people may simply have conceptual blind spots. The fact that STRmix combines several complex areas of expertise—genetics, forensic science, statistics, and programming—suggests that ESR employees, while expert in one, may make errors due to an incomplete grasp of the other. Chessman at 188.

Moreover, financial incentives may pervert the goals of companies that build probabilistic genotyping algorithms. These dynamics are particularly acute in the field of probabilistic genotyping, where the prosecution, backed by the superior resources of the state, is by far the most frequent and reliable customer. That customer is likely to be most satisfied with an algorithm that delivers a match, and is less likely to question its results. Therefore, private companies may be incentivized to find a match, rather than the truth, in order to attract and retain these customers. *See Melendez-Diaz v. Massachusetts*, 557 U.S. 305, 318 (2009) ("A forensic analyst responding to a request from a law enforcement official may feel pressure—or have an incentive—to alter the evidence in a manner favorable to the prosecution."). Market forces will predictably bias results in this direction, notwithstanding the companies' best intentions. Compounding that problem, private companies are also motivated to push for secrecy—as evidenced by this case—which keeps all of these errors hidden from the public. *See also* Elizabeth E. Joh, *The Undue Influence of Surveillance Technology Companies on Policing*, 92 N.Y.U. L. Rev. Online 101, 106 (2017), http://www.nyulawreview.org/sites/default/files/Joh-FINAL_0.pdf.

Not surprisingly, given these multiple potential sources for error, criminal justice algorithms often fail to meet the needs of a rigorous and fair judicial system. In just the last few years, researchers documented a

coding error in STRmix that had enormous consequences: it produced incorrect results in 60 criminal cases in Australia, altering likelihood ratios by a factor of 10 and forcing prosecutors to replace 24 expert statements in criminal cases. David Murray, *Queensland Authorities Confirm 'Miscode' Affects DNA Evidence in Criminal Cases*, Courier-Mail, Mar. 20, 2015, http://www.couriermail.com.au/news/queensland/queensland-authorities-confirm-miscode-affects-dna-evidence-in-criminal-cases/news-story/833c580d3f1c59039efd1a2ef55af92b. STRmix has documented at least seven additional bugs that affect its reported likelihood ratio, occasionally by more than an order of magnitude. RPI Exhibit 6 at 430–31. In addition, ESR has issued numerous new versions of STRmix—including at least one new version since the SPDP crime lab calculated some of its results in this case—to fix identified bugs.

Access to the source code of other probabilistic genotyping algorithms—precisely what the defendant seeks here—has revealed additional errors. In New York, after a federal court ordered the release of FST's source code to the defense, an expert witness discovered that "the program dropped valuable data from its calculations, in ways that users wouldn't necessarily be aware of, but that could unpredictably affect the likelihood assigned to the defendant's DNA being in the mixture." Kirchner, *Traces of Crime*, *supra*. In response, the prosecution withdrew the DNA evidence against the defendant. *Id*. Earlier this year, the New York State Commission on Forensic Science "shelved" two previously approved probabilistic DNA algorithms for similar reasons. *Id.*

These experiences highlight not only the possibility of error, but also the enormous significance of incorrect or unreliable results. A wrong result is a serious problem—both for criminal defendants, whose lives are put into

jeopardy by faulty coding, and for prosecutors, whose cases can be upended by their introduction of unreliable evidence.

Indeed, notwithstanding the fact that each probabilistic DNA algorithm claims to provide accurate results based on objective scientific principles, competing programs frequently reach different results for the same underlying data. For example, in one case, STRmix and TrueAllele generated vastly different results for the same crime scene sample and suspect: TrueAllele found no statistical support for a match, while STRmix generated a likelihood ratio of 300,000. *See* Roth at 2019–20. As a result, the court excluded the STRmix results from trial. See *New York v. Hillary*, No. 2015-15 (N.Y. Cty. Court Aug. 26, 2016).[3] In another case, as discussed above, FST calculated a likelihood ratio of 172 million, while STRmix calculated a likelihood ratio of 10 trillion. *Seepersad*, 2018 WL 1163820, at *1.

Plainly, algorithms are fallible. While this may surprise laypeople, computer scientists, the creators of algorithms, have long been acutely aware of it. They caution that "the evidence produced by computer programs is no more inherently reliable or truthful than the evidence produced by human witnesses." Chessman at 185.

Yet when these algorithms are introduced in the courtroom, legal experts and prosecutors suggest that they are infallible and that their results are foolproof, "overstat[ing] the probative value of their evidence, going far beyond what the relevant science can justify." President's Council of

---

[3] *Available at* www.northcountrypublicradio.org/assets/files/08-26-16DecisionandOrder-DNAAnalysisAdmissibility.pdf.

Advisors on Science and Technology ("PCAST"), *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods* 29 (Sept. 2016), https://perma.cc/W6S6-GWQG. And juries, when deprived of any countervailing testimony that could expose the algorithm's potential pitfalls, generally do not question the prosecution's results. "The potential prejudicial impact" of such evidence is therefore "unusually high." PCAST at 45 (describing finding that mock jurors heavily underestimated the error rates of qualified, experienced forensic scientists); *see also* Matthew Shaer, *The False Promise of DNA Testing*, Atlantic, June 2016, http://theatln.tc/2xs7XUL (describing finding that sexual-assault cases involving DNA evidence in Australia "were twice as likely to reach trial and 33 times as likely to result in a guilty verdict; homicide cases were 14 times as likely to reach trial and 23 times as likely to end in a guilty verdict"). Indeed, knowing the power of such evidence, most defendants plead guilty when confronted with "unfavorable [probabilistic genotyping algorithm] results," highlighting the importance of granting pre-trial access to information about the algorithm. Kirchner, *Traces of Crime*, *supra*.

Source code reveals the programmers' intent, assumptions, biases, and mistakes in ways that no other form of the program can as easily reveal. Adversarial review of the source code would reveal the set of variables used and underlying assumptions made in the algorithm, as well as any errors or mistakes in the source code. Similarly, internal validation studies would show any problems ESR itself has identified with the program, as well as parameters for which STRmix has not been validated.

Like any other evidence, algorithms are neither inherently good nor inherently bad—they are merely tools to augment or replace human analysis of data, with varying degrees of accuracy. The defendant, and the

public, must be given the opportunity to explore that degree of accuracy if our criminal system is to reach just results.

3(B) <u>Denying an accused individual access to an algorithm that will be used to generate material evidence against him in a criminal trial violates his Fourteenth Amendment right to due process.</u>

The Fourteenth Amendment right to due process guarantees, "in essence, the right to a fair opportunity to defend against the State's accusations." *Chambers v. Mississippi*, 410 U.S. 284, 294 (1973). Due process "speak[s] to the balance of forces between the accused and his accuser" and requires that discovery be a "two-way street." *Wardius v. Oregon*, 412 U.S. 470, 474, 475 (1973). When the State's accusations are premised on the results of computerized algorithms, rather than simpler pieces of evidence, maintaining that due process balance and affording the defense a "fair opportunity to defend against the State's accusations" must include pre-trial access to information about the algorithm.

As state supreme courts have recognized with regard to traditional DNA testing, "fair trial and due process rights are implicated when data relied upon by a laboratory in performing [DNA] tests are not available to the opposing party for review and cross examination" pretrial. *State v. Schwartz*, 447 N.W.2d 422, 427 (Minn. 1989); *Ex parte Perry*, 586 So. 2d 242, 255 (Ala. 1991) (requiring disclosure of full details of DNA analysis methodology and holding "defendant's fair trial and due process rights . . . clearly require that the prosecution allow the defendant access to the DNA evidence"). Given the potential complexity of the DNA tests at issue here, the same must hold true for probabilistic genotyping algorithms.

Due process is concerned with all evidence "material either to guilt or to punishment." *Brady v. Maryland*, 373 U.S. 83, 87 (1963). "It is well settled that the government has the obligation to turn over evidence in its

possession that is both favorable to the accused and material to guilt or punishment," and the assertion of an evidentiary privilege does not end the due process inquiry. *Pennsylvania v. Ritchie*, 480 U.S. 39, 57 (1987) (citations omitted).

Evidence is considered material if "there is a reasonable probability that, had the evidence been disclosed, the result of the proceeding would have been different." *Id*. at 40. The documented errors in *United States v. Johnson*, No. 1:15-cr-00565-VEC (S.D.N.Y. 2016), and the miscodes ESR itself has identified, *see* section 3(A) *supra*, establish such a reasonable probability here. That probability is further supported by the well-established potential for evidence generated by algorithms—even, and especially, those that are poorly understood—to prejudice a jury's evaluation of a case. *See* PCAST at 45 ("The vast majority of jurors have no independent ability to interpret the probative value of results based on the detection, comparison, and frequency of scientific evidence."); Shaer, *The False Promise*, *supra* (showing dramatic increase in guilty verdicts where DNA evidence was introduced in sexual assault cases).

Moreover, evidence based on algorithmic source code is at the very center of the prosecution's case against Mr. Dominguez, and is therefore material and relevant to the question of his guilt. *See Schwartz*, 447 N.W.2d at 427 (recognizing that "DNA test results are material to the issue of guilt and could have an impact on the trial outcome" and ordering pre-trial disclosure of related "data, methodology, and actual results" on that basis).

Indeed, at least one court has ordered the prosecution to produce the type of source code at issue here prior to trial because "the results obtained from the use of [a probabilistic genotyping algorithm] on DNA samples recovered from crime scenes are potentially devastating to a criminal

defendant." Order as to Kevin Johnson at 1, *United States v. Johnson*, No. 1:15-cr-00565-VEC (S.D.N.Y. June 7, 2016), ECF No. 57; *see also* Order on Procedural History and Case Status in Advance of May 25, 2016 Hearing, *United States v. Michaud*, No. 3:15-cr-05351-RJB (W.D. Wash. May 18, 2016), ECF No. 205 (holding that source code underlying technique used to identify defendant was material and defendant therefore has a right to access it before the trial); *see also* Order Denying Dismissal and Excluding Evidence, *Michaud*, No. 3:15-cr-05351-RJB (W.D. Wash. May 25, 2016), ECF No. 212.[4]

The defendant's rights to access this information cannot be satisfied by the constrained access the State and ESR have offered in this case. ESR

---

[4] The State argues that accepting Mr. Dominguez's position in this case would mean that Microsoft Excel's source code would be discoverable in any financial crimes case in which the government uses the Excel program to conduct forensic accounting to make its case. Pet.'s Br. at 10. This analogy misses the mark. In a financial crimes prosecution involving a spreadsheet, the relevant algorithms would likely be the specific formulas used to calculate relevant evidence (e.g., "Cell A4 contains the expression '=SUM:A1-A23'"), not the source code to Excel. Such formulas are commonly understood, and can be extracted from the spreadsheet and verified without access to Excel software (e.g., they can be calculated by hand, or with another spreadsheet tool such as LibreOffice Calc.). Just as a spreadsheet is operationalized by a program like Excel or Calc, a Java program like STRmix is operationalized by the Java Virtual Machine (JVM). The defense is asking for the STRmix source, not for the JVM source. Returning to the spreadsheet analogy: in a case where the spreadsheet's calculation is relevant to the State's case in chief, arguing that its formulas are protected by the trade secrets privilege would be folly. And yet that is what the State and ESR are effectively arguing here, in addition to the argument that any technical explanation of how to use Excel also cannot be disclosed because it is copyrighted.

makes "STRmix . . . available for purchase" by the defense, and also offers defense experts access to the source code of the particular version of STRmix used in the defendant's case—but only after the defense expert signs a confidentiality agreement and agrees to conduct any review under direct supervision by the company in an agreed-upon room and through handwritten notes alone. Pet. Exhibit I at 181–82. The court in *Johnson* refused to approve nearly identical constraints, which it described as "strict" and "draconian." Order Terminating Letter Motion at 2, *Johnson*, No. 1:15-cr-00565-VEC (S.D.N.Y. July 6, 2016), ECF No. 67. Moreover, such constraints do not conform with industry practice. Typically, when someone is hired to analyze or audit software, they are expected to: compile the software, run it with varying inputs, instrument and profile it, build and run test suites that exercise the code, conduct static analysis by running investigatory tools over the source code, and conduct dynamic analysis. Therefore, the limits suggested by the State and ESR are not only strict, but also stand directly in the way of efficacy and meaningful review.

As discussed in section 3(A) *supra*, defense access to the source code proved material in that case. After reviewing the source code, the defense expert concluded that "[t]he correctness of the behavior of the FST software should be seriously questioned" for a number of reasons, including the fact that the algorithm engaged in "aberrant behavior" that "depart[ed] from the published descriptions of [the algorithm's] behavior during its actual operation." Exhibit C to Motion in Limine (Nathan Adams Declaration) at 7, 20, *Johnson*, No. 1:15-cr-00565-VEC (S.D.N.Y Jan. 9, 2017), ECF No. 97-3. Specifically, the expert detected that the algorithm simply removed loci saturated with significant amounts of data from its analysis, rather than examining them. *Id.* at 20.

In the civil context, courts have held that government reliance on secret, proprietary algorithms violates due process. *See K.W. v. Armstrong*, 180 F. Supp. 3d 703, 718 (D. Idaho 2016) (holding that proprietary tool used to allocate Medicaid benefits "arbitrarily deprives participants of their property rights and hence violates due process"); *T. v. Bowling*, No. 2:15-cv-09655, 2016 WL 4870284, at *10 (S.D.W. Va. Sept. 13, 2016) (finding that proprietary algorithm used by government to set Medicaid benefits "present[s] a serious risk of resulting in erroneous determinations and deprivations"). The constitutional stakes are even higher in the criminal context and require the same result here.

Moreover, the trial court has broad authority to manage discovery. "The standard of review for a discovery order is abuse of discretion, because management of discovery lies within the sound discretion of the trial court. Thus, where there is a basis for the trial court's ruling and it is supported by the evidence, a reviewing court will not substitute its opinion for that of the trial court." *Lee v. Superior Court*, 177 Cal. App. 4th 1108, 1124 (Cal. Ct. App. 2009) (alterations omitted). Here, the trial court has correctly determined that Mr. Dominguez must be granted access to STRmix's source code and internal validation studies, and that ruling is supported by the evidence of that information's centrality to and possibility for error in the case.

In addition, overturning that order and delaying the accused's access to evidence until trial will likely put the trial court in the unfortunate position of pitting a defendant's constitutional rights against judicial economy. If access to the evidence is not granted until trial, when the additional rights discussed in section 3(C) *infra*, attach, the defendant may seek a continuance to fully assess it. *See, e.g.*, *People v. Samayoa*, 938 P.2d 2, 32–

33 (Cal. 1997) (trial court maintains broad discretion to continue trial in light of introduction of evidence not disclosed until trial). This is particularly true where, as here, the state's evidence may be complicated and relatively novel, requiring more time to prepare an adequate cross examination.

3(C) If the secret algorithm is not disclosed at this stage, the defendant's Sixth and Fourteenth Amendment rights to confrontation and a fair trial will be implicated at trial.

If this Court were to overturn the trial court's grant of access to STRmix's source code, Mr. Dominguez's Sixth Amendment right to confront "the witnesses against him," U.S. Const. amend. VI, and his Fourteenth and Sixth Amendment rights to a fundamentally fair process would be implicated at trial. While those rights are not at direct issue in this appeal, they will almost certainly come up at trial should the State's writ be granted.

"Whether rooted directly in the Due Process Clause of the Fourteenth Amendment or in the Compulsory Process or Confrontation Clauses of the Sixth Amendment, the Constitution guarantees criminal defendants a meaningful opportunity to present a complete defense." *Holmes v. South Carolina*, 547 U.S. 319, 319 (2006) (quoting *Crane v. Kentucky*, 476 U.S. 683, 690 (1986)) (quotation marks omitted). Broadly speaking, "a fair trial is one in which evidence subject to adversarial testing is presented to an impartial tribunal for resolution of issues defined in advance of the proceeding." *Strickland v. Washington*, 466 U.S. 668, 685 (1984).

In addition to the due process concerns discussed in section 3(B) *supra*, several other strands of the due process doctrine will become relevant at trial. First, with respect to evidence withheld from a defendant, due process asks "whether in its absence [the defendant] received a fair trial, understood

as a trial resulting in a verdict worthy of confidence," *Kyles v. Whitley*, 514 U.S. 419, 434 (1995). In addition, due process requires rejection of asymmetrical evidentiary rules—that is, those that place the prosecution's evidence in a more favorable position than the defendant's. *See Holmes*, 547 U.S. at 331. Finally, due process protects the right to cross-examine witnesses—including adversarial testing of the source code upon which they rely—in part because the jury must be empowered to "judge for itself whether [ ] testimony [is] worthy of belief." *Chambers*, 410 U.S. at 295.

Relatedly, the Confrontation Clause's animating concern is "to ensure the reliability of the evidence . . . by subjecting it to rigorous testing." *Maryland v. Craig*, 497 U.S. 836, 845 (1990). The Supreme Court has recognized that this concern applies with full force to forensic evidence. *Melendez-Diaz*, 557 U.S. at 313 (holding that affidavits reporting the results of a forensic analysis of seized drugs are testimonial and subject to the Confrontation Clause); *Bullcoming v. New Mexico*, 564 U.S. 647, 663–64, 666 (2011) (holding that certification on a forensic laboratory report is testimonial and defendant has a right to confront the specific analyst who made the certification).

The Sixth Amendment also guarantees a defendant the right "to have compulsory process for obtaining witnesses in his favor." U.S. Const. amend. VI. At a minimum, compulsory process means that criminal defendants have "the right to put before a jury evidence that might influence the determination of guilt." *Ritchie*, 480 U.S. at 56.

If this Court grants the State's writ, Mr. Dominguez's confrontation right will almost certainly be violated at trial because his lack of access to STRmix's source code will unduly inhibit his ability to confront any witness testifying about the program's results. Effectively confronting such

testimony necessarily requires that the defense access and confront STRmix's source code.

To be sure, the California Supreme Court held in *People v. Lopez* that mechanical printouts of raw data are not statements, and that "a machine cannot be cross-examined." *People v. Lopez*, 286 P.3d 469, 478 (Cal. 2012). That case held that the results of a blood alcohol analysis performed by a gas chromatography machine were not testimonial under *Crawford v. Washington*, 541 U.S. 36 (2004). *Lopez*, 286 P.3d at 478–79. It did not address the question that will be presented if this writ is granted—whether a court violates a defendant's right to confront an expert by denying him access to the source code used to generate the data underlying an expert's testimony. With this distinction, *Lopez* in fact supports the disclosure of source code. There, defense counsel had access to a printout of the calibrations of the gas chromatography machine taken on the same day as the relevant test—a close parallel to the source code sought by Mr. Dominguez here. *See id.* at 477; *see also People v. Vangelder*, 312 P.3d 1045, 1048 (Cal. 2013).

Any testimony about STRmix's statistics will almost certainly result from the "distributed cognition" among the State's crime lab technicians, ESR, and the software itself. Itiel E. Dror & Jennifer L. Mnookin, *The Use of Technology in Human Expert Domains: Challenges and Risks Arising from the Use of Automated Fingerprint Identification Systems in Forensic Science*, 9 L. Probability & Risk 1, 2 (2010); *see also* Chessman at 220 ("When a forensic report is the output of a computer program, it is thus a joint statement—one composed of the interaction between the statements of the programmer and the input of the program user."). Just as ESR did not gather the DNA samples at issue itself, the company also did not manually

calculate the likelihood ratio; similarly, the crime lab did not generate the methodology by which the calculation was done, nor did it build the program that generated the statistics. *See United States v. Washington*, 498 F.3d 225, 230 (4th Cir. 2007) (noting that technicians who operated a gas chromatograph could not independently verify the results because they only relied on the analysis performed by the machine). Indeed, the entire reason ESR developed STRmix was to replace and improve upon manual calculations. RPI Exhibit 5 at 428. Instead, ESR developed a set of assumptions about probabilistic genotyping and programmed those assumptions into STRmix. Even accepting that STRmix software cannot be a witness under current California law, that software, under ESR's design and direction, performed the only probabilistic calculations of the DNA mixtures in this case. And its analysis produced the inculpatory estimates of the likelihood that Mr. Dominguez was a contributor to the collected DNA samples.

In this light, STRmix's source code will be a critical component of contesting any testimony regarding its likelihood ratios. *See Delaware v. Van Arsdall*, 475 U.S. 673, 680 (1986) ("[A] criminal defendant states a violation of the Confrontation Clause by showing that he was prohibited from engaging in otherwise appropriate cross-examination designed to show a prototypical form of bias on the part of the witness."). Confronting ESR experts or crime lab technicians may reveal some bias or mistakes in their assumptions in formulating an algorithm or performing analysis of its results, but examining the source code is the only way to uncover the software's bias or mistakes. The software's intricate relationship with and dependence upon its human creators means that its operation is not immune from fraud, bias or incompetence. *See* § 3(A), *supra*. To the contrary,

coding errors—both deliberate and benign—are an inherent and significant part of programming. Roth at 1994; *see generally* Chessman at 183–99 (discussing various forms and frequencies of programming errors). As discussed above, consequential coding errors have been discovered in probabilistic genotyping programs once they were subject to outside scrutiny. *See* §§ 3(A) and (B), *supra*. These are the very sorts of evils confrontation is meant to deter. *Melendez-Diaz*, 557 U.S. at 318–19.

The assertion of an evidentiary privilege does not end these constitutional inquiries. *Ritchie*, 480 U.S. at 57. Indeed, the Supreme Court has held that, to preserve the "fundamental fairness of trials," material information covered by an evidentiary privilege should nonetheless have been provided to a criminal defendant, even where it consisted of extremely sensitive information in a state agency's child abuse investigation file. *Id*. at 56–57; *see also Chambers*, 410 U.S. at 302 (cautioning that "[m]echanistic[]" application of hearsay rule to exclude evidence "critical" to a criminal defendant's case can "defeat the ends of justice" and violate due process); *Green v. Georgia*, 442 U.S. 95, 97 (1979).

Nor is it any answer, as the State offers, that the source code and internal validation studies are unnecessary because STRmix's general methodology has been validated. As an initial matter, the State should not be able to rest its argument on validation studies—including STRmix's internal validation and modification studies and the SPDP crime lab's validation studies—that are not disclosed in full to the defense. In addition, as discussed above, defense access to the sort of source code at issue here has proven its worth in circumstances where validation studies were already available. *See* § 3(B), *supra* (discussing *United States v. Johnson*, No. 1:15-cr-00565-VEC (S.D.N.Y. 2016)).

Moreover, the Supreme Court has repeatedly admonished that the right to confrontation is procedural, and cannot be discarded simply because the evidence appears reliable. *Crawford*, 541 U.S. at 62 ("Dispensing with confrontation because testimony is obviously reliable is akin to dispensing with jury trial because a defendant is obviously guilty."); *Melendez-Diaz*, 557 U.S. at 318; *see also Bullcoming*, 564 U.S. at 663 ("If a particular guarantee of the Sixth Amendment is violated, no substitute procedure can cure the violation, and no additional showing of prejudice is required to make the violation complete." (citation and quotation marks omitted)).

Regardless, notwithstanding the useful information that validation studies can offer, validation is far from a panacea for guaranteeing the accuracy of probabilistic genotyping in particular cases. The validation studies themselves are often conducted under conditions far more ideal than the actual circumstances in the field where they are typically deployed, and likelihood ratios are by their very nature more difficult to falsify, since their predictions can rarely be compared to an objectively "correct" result (i.e., whether or not an individual is, in fact, a contributor to a crime-scene DNA sample). Roth at 1982. Thus, a validated program's likelihood ratio still "might be off by orders of magnitude because of a host of human or machine errors." *Id*. This phenomenon introduces the risk that the results of a validated program may still be highly misleading. Examining whether the source code is operating as designed is therefore critical to determining the likelihood ratio's true accuracy.

Second, the prosecution's argument that access to the SPDP Crime Lab's validation and modification studies should suffice is unavailing. Such studies, depending on the parameters they tested, may not be sufficient. In addition, such studies are likely to report merely what the crime lab

believes it did, and it could be mistaken. Evaluation of the source code, in contrast, would allow the defense to verify that the validations were done correctly and reflect the same scientific expectations.

But even if SPDP's studies were superior, the existence of an alternative way to challenge STRmix's results would not change the fact that "the Constitution guarantees one way: confrontation." *Melendez-Diaz*, 557 U.S. at 318. Here, meaningful confrontation will require defense access to the source code. All complex software has errors, and ESR's admission that there have been errors in the code show that STRmix is no exception. RPI Exhibit 6 at 430–31.

At its root, this case reveals the strong parallels between black-box technologies like STRmix and the *ex parte* examinations that motivated the founders to adopt the Confrontation Clause in the first place. Performed at the behest of the state, intentionally cloaked in secrecy, and unduly impressive to the unwitting juror, both render the defendant powerless to test the credibility of the source and undermine the state's case against him. *See generally Crawford*, 541 U.S. at 43–50 (describing history and development of confrontation right). Allowing the defense access to source code is the only reliable means of ensuring that the state cannot place forensic evidence beyond the reach of the Confrontation Clause, simply by automating tasks previously performed by humans. Otherwise, regardless of whether courts consider machines witnesses or their products hearsay, our justice system risks "accept[ing] the product of a computer as the equivalent of Holy Writ." *Perma Research & Dev. v. Singer Co.*, 542 F.2d 111, 121 (2nd Cir. 1976) (Van Graafeiland, J., dissenting).

3(D) **In addition to risking the defendant's rights, rejecting transparency at this stage will ensure that the public's First Amendment right of access is vitiated at trial.**

Dismissing this writ would not only protect Mr. Dominguez's due process rights and ensure protection of his trial rights; it would also benefit the public by positioning it to enforce its First Amendment right of access. In allowing Mr. Dominguez to obtain, examine, and introduce information about the algorithm into the record, the discovery order will also enable the public to exercise its longstanding First Amendment right of access to criminal proceedings. While that right is not at direct issue in this appeal, it provides a different and useful lens for scrutiny of the trial court's decisions below, and offers important context for this Court's consideration of the other constitutional issues in play. Moreover, if the trial court's order is reversed, there will never be a point at which the public's First Amendment right of access is addressed.

If the trial court's discovery order is allowed to stand, the public's qualified First Amendment right of access will attach to any materials about STRmix that are entered into the record or become the subject of substantive litigation. As discussed further below, the simple act of sharing the information through discovery may not suffice to subject it to the right of access, but the right will attach if the parties rely on or incorporate the materials into litigation about substantive rights. Such materials may range from the algorithm's source code to ESR's internal validation studies to any eventual defense expert reports.

Once the right of access attaches, proceedings and records are presumptively open to the public, but they may be closed where there are "specific, on the record findings" that "closure is essential to preserve higher values and is narrowly tailored to serve that interest." *Press-Enter*.

Co. v. Superior Court ("*Press-Enter. II*"), 478 U.S. 1, 13, 14–15 (1986) (quoting *Press-Enter. Co. v. Superior Court* ("*Press-Enter. I*"), 464 U.S. 501, 510 (1984)); *see also Globe Newspaper Co. v. Superior Court*, 457 U.S. 596, 606–07 (1982); *N.Y. Civil Liberties Union v. N.Y.C. Transit Auth.*, 684 F.3d 286, 296 (2d Cir. 2012) (requiring a substantial probability of harm to a compelling government interest, and no alternative that can effectively protect against that harm to overcome presumption of access).

i. The First Amendment right of access exists to allow the public to meaningfully oversee courtroom proceedings.

The First Amendment exists to enable and protect "uninhibited, robust, and wideopen [*sic*]" debate on public issues, *Watts v. United States*, 394 U.S. 705, 708 (1969), and "for the bringing about of political and social changes desired by the people," *Roth v. United States*, 354 U.S. 476, 484 (1957). Neither is possible without public access to judicial proceedings and documents—a principle the Supreme Court recognized almost forty years ago when it held that "the right to attend criminal trials is implicit in the guarantees of the First Amendment," which includes the right to "'receive information and ideas.'" *Richmond Newspapers, Inc. v. Virginia*, 448 U.S. 555, 556, 576 (1980) (quoting *Kleindienst v. Mandel*, 408 U.S. 753, 762 (1972)).

Public access to materials about STRmix will ensure that widespread problems—whether in the algorithm's design, its operationalization through source code, or the algorithm owner's approach to delivering results—can be efficiently audited by independent experts. Allowing the public, including academics and other experts, to examine DNA typing evidence would markedly improve the reliability and fairness of such evidence in criminal trials.

This would achieve one of the main purposes of the First Amendment right of access, which attaches to criminal trials to allow the public to observe and evaluate the workings of the criminal justice system—and to make changes in order to eliminate injustice. See *id*. at 572. As the Supreme Court has explained, "the criminal justice system exists in a larger context of a government ultimately of the people, who wish to be informed about happenings in the criminal justice system, and, if sufficiently informed about those happenings, might wish to make changes in the system." *Gentile v. State Bar of Nev*., 501 U.S. 1030, 1070 (1991). The need for public oversight of government process is strongest in criminal trials, where the state wields its greatest power to affect individual liberty. Public access "enhances the quality and safeguards the integrity" of the judicial process, "heighten[s] public respect" for that process, and "permits the public to participate in and serve as a check upon the judicial process." *Globe Newspaper*, 457 U.S. at 606.[5]

Under the Supreme Court's prevailing "experience and logic" test, the public's First Amendment right of access attaches to judicial proceedings and records where (a) the type of judicial process or record sought has

---

[5] The importance of public access to criminal trials is also embedded in the common law, *see, e.g.*, *Lugosch v. Pyramid Co.*, 435 F.3d 110, 119 (2d Cir. 2006), as well as the Sixth Amendment, which guarantees a criminal defendant the right to a public trial, *see, e.g.*, *In re Oliver*, 333 U.S. 257, 267–68 (1948). Indeed, the Supreme Court has suggested that the demands of the Sixth's Amendment's public-trial right—grounded in the defendant's right to a fair trial—may go even further than the First Amendment right in certain cases. *See Presley v. Georgia*, 558 U.S. 209, 213 (2010); *Waller v. Georgia*, 467 U.S. 39, 46 (1984) ("There can be little doubt that the explicit Sixth Amendment right of the accused is no less protective of a public trial than the implicit First Amendment right of the press and public.").

historically been available to the public, and (b) public access plays a "significant positive role" in the functioning of the process itself. *Press-Enter. II*, 478 U.S. at 9, 11; *see Globe Newspaper*, 457 U.S. at 605–07.

ii. The broad reach of the First Amendment right of access encompasses algorithms used to produce evidence introduced to prove the guilt of a defendant.

Assuming that this case will proceed to trial, there is little question that the right of access will attach to the criminal trial below. Indeed, the Supreme Court grounded the First Amendment "presumption of openness [that] inheres in the very nature of a criminal trial under our system of justice" in the "unbroken, uncontradicted history" of such access, "supported by reasons as valid today as in centuries past." *Richmond Newspapers*, 448 U.S. at 573; *see also Press-Enter. I*, 464 U.S. at 505–08 (discussing history of openness in criminal trials); *Cal. First Amendment Coal. v. Woodford*, 299 F.3d 868, 874 (9th Cir. 2002); *KNSD Channels 7/39 v. Superior Court*, 74 Cal. Rptr. 2d 595, 596–97 (Cal. Ct. App. 1998).

And once the right attaches to a proceeding, the presumption of access applies broadly to all materials essential to that proceeding—including the algorithmic source code in this case. *See Doe v. Pub. Citizen*, 749 F.3d 246, 267 (4th Cir. 2014) ("[T]he First Amendment right of access extends to materials submitted in conjunction with judicial proceedings that themselves would trigger the right to access."); *see also In re Application of WFMJ Broad. Co.*, 566 F. Supp. 1036, 1040 (N.D. Ohio 1983) ("Just as the Supreme Court's reluctance to embrace a 'narrow, literal conception of the [First] Amendment's terms', *Globe Newspaper*[, 457 U.S. at 604], gave rise to a constitutional right of access to criminal trials, the same view could make a constitutional right to evidence an appropriate adjunct to insure that such proceedings are 'open.'").

As the Ninth Circuit recognized in *Woodford*, meaningful access to a proceeding means access to its nuts and bolts. In *Woodford*, a lethal injection case, that meant a right to view "executions from the moment the condemned is escorted into the execution chamber." 299 F.3d at 870–871, 877. The court explained that, for the right of access to accomplish its goals, citizens must have reliable information about the 'initial procedures,' which are invasive, possibly painful and may give rise to serious complications." *Id.* at 876–77. The same must be true for algorithms that produce the prosecution's material evidence in a criminal trial—which also have the potential for serious complications and inaccuracies. Just as without access to the initial procedures of an execution, "the public will be forced to rely on the same prison officials who are responsible for administering the execution to disclose and provide information about any difficulties with the procedure," without access to the algorithms that create material evidence, the public will be forced to rely on the same government officials responsible for introducing the evidence and convincing the judge and jurors that they should trust it. *Id.* at 883. And, much like prison officials, these government officials "do not have the same incentives to describe fully the potential shortcomings of" their evidence. *Id.* at 884. Here, as in *Woodford*, the government cannot artificially cabin the record of a proceeding in order to deny public access to all but the ultimate result.[6]

---

[6] Courts have held that the public's First Amendment right of access attaches to materials in the record of a criminal case for this reason. *See, e.g.*, *In re Globe Newspaper Co.*, 729 F.2d 47 (1st Cir. 1984) (right of access attaches to memorandum, affidavits and transcripts in criminal case); *In re N.Y. Times Co.*, 828 F.2d 110 (2d Cir. 1987) (same for suppression motions and exhibits); *In re Wash. Post Co.*, 807 F.2d 383 (4th Cir. 1986)

Moreover, the work of one legal scholar suggests that limiting access on the basis of a purported trade secret privilege would be ahistorical. Rebecca Wexler has found that "[e]arly historical sources suggest that the [trade secrets] privilege"—precisely the tool companies are now using to keep algorithms out of the record of criminal cases—was historically "unavailable in criminal proceedings." Wexler at 1388–90. Rather, historically, when courts were asked to conceal trade secrets from disclosure in criminal trials, they instead held that the secrets must be disclosed. *See id.* (discussing *R v. Maha Rajah Nundocomar*, 20 Howell State Trials 923, 1057 (1775), and *R v. Webb*, 174 Eng. Rep. 140 (1834)). This suggests that permitting the State to keep source code hidden on the basis of a trade secret privilege would block from view information of a type that would historically have been public. Similarly, an attempt to shield material from disclosure on the assertion of a copyright would contradict precedent allowing for distribution and copying in the context of litigation. *See, e.g.*, *Bond v. Blum*, 317 F.3d 385, 396 (4th Cir. 2003) (abrogated by *Kirtsaeng v. John Wiley & Sons, Inc.*, 136 S. Ct. 1979 (2016), on other grounds) (recognizing that "the societal benefit of having all relevant information," including copyrighted materials, "presented in a judicial proceeding is an important one").

Moreover, openness in the context of algorithms used to produce evidence of guilt would have immense public value. There is a long history of junk science being used under the guise of technological advance in criminal cases in this country—and of public access to and analysis of such

---

(same for plea agreements); *United States v. Peters*, 754 F.2d 753, 763 (7th Cir. 1985) (same for trial exhibits).

evidence establishing its invalidity. "Since a series of high-profile legal challenges in the 1990s increased scrutiny of forensic evidence, a range of long-standing crime-lab methods have been deflated or outright debunked," including bite-mark analysis, ballistics testing, fingerprinting, and microscopic-hair-comparison. Shaer, *The False Promise*, *supra*.

Indeed, the Supreme Court has relied on public scrutiny of forensic processes to inform its interpretation of constitutional protections. *See Melendez-Diaz*, 557 U.S. at 319 ("Serious deficiencies have been found in the forensic evidence used in criminal trials."). And state supreme courts—as well as federal appellate courts—have equally looked to work done by the public, rather than either party or its experts in a criminal case, to determine that evidence based on specific technologies was not sufficiently reliable to be admissible into evidence. *See, e.g.*, *Han Tak Lee v. Houtzdale SCI*, 798 F.3d 159, 166–67 (3d Cir. 2015) (discussing changes in "fire–science"); *People v. Leone*, 255 N.E.2d 696 (N.Y. 1969) (relying on commentary of outside experts to hold that evidence derived from polygraph tests was not fit for admission); *see People v. Davis*, 72 N.W.2d 269, 281–82 (Mich. 1965) (same).

Public scrutiny has had substantial benefits outside of the courtroom as well, leading to important improvements in investigative fields. For example, after a *New Yorker* article exposed a flawed case based on fire-science evidence, Texas not only "reconsider[ed] old cases that had been improperly handled by the original investigators," but also "reinvented itself as a leader in arson science and investigation" by "revamp[ing] the state's training and investigative standards." Jeremy Stahl, *The Trials of Ed Graf*, Slate, Aug. 16, 2015, https://perma.cc/89TJ-4ASK.

And all of this is true of DNA evidence, as well. In the DNA field,

"[b]oth the initial recognition of serious problems and the subsequent development of reliable procedures were aided by the existence of a robust community of molecular biologists" and by "judges who recognized that this powerful forensic method should only be admitted as courtroom evidence once its reliability was properly established." PCAST at 26.

Given the positive effect of public access on the use of arguably simpler technologies in criminal case, public access would plainly enhance the reliability of algorithmic evidence. This is particularly true of technologies that, like the likelihood ratio introduced in this case, have been minimally tested in the field. Most existing validation studies of probabilistic DNA typing have been "conducted under idealized conditions unrepresentative of the challenges of real casework." Roth at 2033; *see also* Christopher D. Steele & David J. Balding, *Statistical Evaluation of Forensic DNA Profile Evidence*, 1 Ann. Rev. Stat. & App. 361, 380 (2014). Moreover, "most of the studies evaluating software packages have been undertaken by the software developers themselves." PCAST at 80. Public access to algorithmic evidence would improve the role such evidence plays in criminal trials—including by preventing the jury from giving it undue weight, where necessary—and increase the public's confidence in the justice system more generally.[7]

---

[7] The government may argue that requiring the release of source code will have a negative effect on the proceedings because it will create additional disputes, but that argument would be misplaced. The government has no interest in unfair proceedings, even if they take longer. Moreover, public vetting of algorithmic source code will surely experience efficiency gains as it becomes a more commonplace check on complex, experimental evidence. As the Supreme Court recently reaffirmed, "the Government's 'interest . . . in a criminal prosecution is not that it shall win a case, but that

Allowing the public, including academics and other experts, to examine DNA typing evidence would markedly improve the reliability and fairness of such evidence in criminal trials. The other checks our judicial system relies upon, like recordation and appeal, "operate rather as cloaks than checks; as cloaks in reality, as checks only in appearance." *In re Oliver*, 333 U.S. at 271. "Without publicity, all other checks are insufficient." *Id.* (quotation marks omitted)).

As one scholar, Erin Murphy, has explained, numerous factors that plague the defense in criminal trials—including "structural asymmetry[,] . . . scarcity of resources, weak discovery practices, and high rate of plea bargaining"—make the "adversarial process an inadequate safeguard of the integrity of forensic science." Murphy at 757. But experts reviewing publicly disclosed information about algorithms, including the source code, should be free of these obstacles and should have the time, resources, and expertise to effectively and efficiently audit the algorithmic programs. Moreover, allowing the public to view at least some of the information would avoid the potential "devastating effect" of overly broad protective orders—and confidentiality agreements like the one used in this case—that prevent expert findings in one case from spreading to others, where they would be equally relevant and useful. Wexler at 1412–13. And independent review of documents across cases may catch errors or mistakes that would not be identifiable in one case alone. Murphy at 773.

Indeed, public review of the sort of source code at issue here has already proven its worth. In a 2008 case, a defense expert's review of Alcotest 7110

---

justice shall be done.'" *Turner v. United States*, 137 S. Ct. 1885, 1893 (2017) (quoting *Berger v. United States*, 295 U.S. 78, 88 (1935)).

source code in one case—which "documented 19,500 errors, nine of which he believed could ultimately affect the breath alcohol reading," Roth at 1995 (internal marks omitted)—led the New Jersey Supreme Court in another case to require modifications to prevent misleadingly high accuracy readings. *State v. Chun*, 943 A.2d 114, 120–21 (N.J. 2008). And, after the court lifted the protective order in *United States v. Johnson*, the state recognized that the secrecy surrounding FST had "exacerbated the substantial misunderstanding of fundamental aspects of the FST source code." Lauren Kirchner, *Federal Judge Unseals New York Crime Lab's Software for Analyzing DNA Evidence* (Oct. 20, 2017, 8:00 A.M.), https://www.propublica.org/article/federal-judge-unseals-new-york-crime-labs-software-for-analyzing-dna-evidence. In other words, such secrecy hurts the criminal justice system on all sides, impeding the process not only for the defense but for the prosecution as well. The expert reports submitted in that case are now publicly available and FST's source code is on GitHub. *See New York City's Forensic Statistical Tool*, GitHub https://perma.cc/348Z-6W6M (last updated Oct. 20, 2017).

Moreover, while some courts have (erroneously) applied a narrower test to determining whether the First Amendment right-of-access attaches—looking to the nature of a particular document rather than proceedings themselves, *see In re Bos. Herald, Inc.*, 321 F.3d 174, 182–84 (1st Cir. 2003) (reviewing case law applying the First Amendment right of access to proceedings and documents)—the right would still attach to information about an algorithm used to produce evidence of guilt in a criminal case under this analysis.

Under the test's "experience" prong, it is not only well established but fundamental that the materials essential to the government's case in chief

enjoy a presumption of openness in the criminal justice system. *See, e.g.*, *In re Application of WFMJ Broad. Co.*, 566 F. Supp. at 1040 (tapes played to jury in open court); *United States v. Posner*, 594 F. Supp. 930, 934–35 (S.D. Fla. 1984) (tax returns admitted into evidence); *United States v. Scott*, 48 M.J. 663 (A. Ct. Crim. App. 1998) (materials entered into evidence at trial); *Valley Broad. Co. v. U.S. Dist. Court*, 798 F.2d 1289, 1292–93 (9th Cir. 1986) (transcripts of exhibits); *In re Times-World Corp*., 488 S.E.2d 677 (Va. 1997) (documents submitted into evidence). And the right also attaches to supporting materials that form a critical component of the record, especially when they pertain to the "adjudicat[ion of] substantive rights," *Rushford v. New Yorker Mag.*, 846 F.2d 249, 252 (4th Cir. 1988).[8]

While courts have held that the "raw fruits" of discovery may not be subject to the right of access, *see, e.g.*, *id.*; *Seattle Times Co. v. Rhinehart*, 467 U.S. 20 (1984), that conclusion is altered where the parties rely on or

---

[8] Moreover, the "experience" prong "is not meant . . . to be construed so narrowly" as to exclude from First Amendment coverage proceedings or documents that are of "relatively recent vintage." *In re Bos. Herald*, 321 F.3d at 184. In such cases, courts look to analogous proceedings and documents of the same "type or kind." *Rivera-Puig v. Garcia-Rosario*, 983 F.2d 311, 323 (1st Cir. 1992) (emphasis omitted); *see El Vocero de P.R. v. Puerto Rico*, 508 U.S. 147, 150–51 (1993) (finding pretrial criminal hearings in Puerto Rico analogous to other pretrial hearings to which First Amendment right applies, despite distinctions noted by Puerto Rico Supreme Court); *Press–Enter. II*, 478 U.S. at 10–11 (evaluating California pre-trial hearings by looking to practices of other states and to other types of hearings, including probable cause hearing in Aaron Burr's 1807 trial for treason); *see also United States v. Chagra*, 701 F.2d 354, 363 (5th Cir. 1983) ("Because the first amendment must be interpreted in the context of current values and conditions, the lack of an historic tradition of open bail reduction hearings does not bar our recognizing a right of access to such hearings." (citations omitted)).

incorporate discovery materials into substantive litigation. Indeed, in the civil context courts have held that under the First Amendment, reports relied upon by parties in the "adjudication stages" of litigation are presumptively "available for public inspection unless exceptional circumstances require confidentiality." *In re Continental Ill. Sec. Litig.*, 732 F.2d 1302, 1314 (7th Cir. 1984); *accord Joy v. North*, 692 F.2d 880, 893 (2d Cir. 1982); *see also Rushford*, 846 F.2d at 253 (documents filed in connection with summary judgment motion); *NBC Subsidiary (KNBC-TV), Inc. v. Superior Court*, 980 P.3d 337, 360 n.28 (Cal 1999) (applying the same principle in a civil context).[9] Those principles apply with even greater force in the criminal context to evidence and its attendant documents, *see, e.g.*, *In re Wash. Post Co.*, 807 F.2d at 389–90 —and, assuming this case proceeds to trial or any proceeding or briefing that adjudicates substantive rights, this would encompass information about an algorithm that produces the evidentiary results at the center of the State's case against Mr. Dominguez. *See Doe*, 749 F.3d at 267.

As discussed above, the "logic" prong also dictates that the First Amendment right of access attaches in this context. Public access to the highly complex algorithmic source code that produced the evidence that will be used against Mr. Dominguez at trial would "enhance[] the quality and safeguard[] the integrity of the factfinding process, with benefits to both the defendant and to society as a whole," *Globe Newspaper Co.*, 457 U.S. at

---

[9] Even courts that have rejected the attachment of a First Amendment right of access in particular contexts have acknowledged that the right may well attach where "the material is important and the decision to which it is relevant amounts to an adjudication of an important substantive right." *Anderson v. Cryovac, Inc.*, 805 F.2d 1, 11 (1st Cir. 1986).

606; *see also, e.g.*, *Grove Fresh Distribs., Inc. v. Everfresh Juice Co.*, 24 F.3d 893, 897 (7th Cir. 1994) (citing *Richmond Newspapers*, 448 U.S. at 555); *Int'l Fed'n of Prof'l & Tech. Eng'rs, Local 21, AFL-CIO v. Superior Court*, 42 Cal. 4th 319, 333 (2007).[10]

Public access to the foundation of the algorithmic evidence introduced to prove Mr. Dominguez's guilt will allow for a thorough public vetting of a new technology, with all its salutary consequences. In particular, in the context of criminal cases in which defendants and their counsel have limited resources, public access to algorithmic evidence would bolster the purpose of the *Kelly-Frye* inquiry at trial to "ensure that any and all scientific testimony or evidence admitted is not only relevant, but reliable," *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 589 (1993), by providing the public with an opportunity to evaluate and test evidentiary material.[11]

---

[10] *See also Leucadia, Inc. v. Applied Extrusion Techs., Inc.*, 998 F.2d 157, 161 (3d Cir. 1993) ("As with other branches of government, the bright light cast upon the judicial process by public observation diminishes the possibilities for injustice, incompetence, perjury, and fraud. Furthermore, the very openness of the process should provide the public with a more complete understanding of the judicial system and a better perception of its fairness." (quoting *Rep. of Phil. v. Westinghouse Elec. Corp.*, 949 F.2d 653, 660 (3d Cir. 1991)); *United States v. Hubbard*, 650 F.2d 293, 315 n.79 (D.C. Cir. 1980) (Like the public trial guarantee of the Sixth Amendment, the First Amendment right of access serves to "safeguard against any attempt to employ our courts as instruments of persecution," to promote the search for truth, and to assure "confidence in . . . judicial remedies."); *Ibrahim v. Dep't of Homeland Sec.*, 62 F. Supp. 3d 909, 934–45 (N.D. Cal. 2014) ("Public oversight of courts and therefore public access to judicial operation is foundational to the functioning of government. Without such oversight, the government can become an instrument for injustice.").

[11] To be clear, a *Kelly-Frye* (or *Daubert*) hearing is plainly an insufficient substitute for scrutiny of algorithmic source code, as it goes only towards

iii. The court may limit the public's access to information about the algorithm, but any limitations must be narrowly tailored to comport with the First Amendment.

Of course, the fact that the First Amendment right of access will *attach* to algorithmic source code properly entered into the record does not dictate that the source code itself will be made public, in part or in its entirety. Because the right is a qualified one, the outcome (in this case or any other) will depend upon the strength of the government's interest in continued secrecy, as well any measures taken to narrowly tailor the denial of the source code to the public, including through a protective order. *See Press-Enter. II*, 478 U.S. at 13–14; *see also Globe Newspaper Co.*, 457 U.S. at 608 (explaining that even a compelling government interest "does not justify a *mandatory* closure rule, for it is clear that the circumstances of the particular case may affect the significance of the interest"); *United States v. Amodeo*, 71 F.3d 1044, 1049 (1995); *Grove Fresh Distribs.*, 24 F.3d at 898. And that process will require the government, and then the court, to make on-the-record findings concerning the reasons justifying full or partial secrecy. *See Press-Enter. II*, 478 U.S. at 13–14.

It is clear, however, that where a criminal case involves algorithmic source code that produces material evidence like that in Mr. Dominguez's case, the strength of the public's right of access should favor some level of disclosure. Indeed, the Supreme Court has explained that the

---

admissibility (a matter decided by the judge), rather than weight (a matter decided by the jury). *See, e.g.*, *People v. Barney*, 8 Cal. App. 4th 798, 817 (1992). Any flaws or errors in source code would tend to undermine the value of state evidence based on it, and would permit the defendant to argue to the jury to disregard the experimental test results introduced into evidence.

"circumstances" in which "the right to an open trial may give way . . . to other rights or interests . . . will be rare." *Waller*, 467 U.S. at 45. Such sufficiently weighty rights and interests might include, for example, "the defendant's right to a fair trial or the government's interest in inhibiting disclosure of sensitive information." *Id.* But the government's interest in this case and those like it does not approach that class of gravity. To the contrary, the defendant's right to a fair trial dovetails—rather than conflicts—with the public's right of access. *See supra* § 3(C).

Here, the government's only interest in secrecy appears to be derivative of a business's intellectual-property interest in purported trade secrets and copyrighted information. This private interest, on its own, will likely fail strict scrutiny. The Supreme Court has "recognized that the First Amendment interests served by the disclosure of purely private information like trade secrets are not as significant as the interests served by the disclosure of information concerning a matter of public importance." *DVD Copy Control Ass'n v. Bunner Inc.*, 31 Cal. 4th 864, 883 (2003) (citing *Bartnicki v. Vopper*, 532 U.S. 514, 533 (2001); *Dun & Bradstreet, Inc. v. Greenmoss Builders, Inc.*, 472 U.S. 749, 759 (1985)). In fact, because the private "makers are under a scientific obligation to release this information for peer review," the validity of the interest is questionable. Jennifer N. Mellon, *Manufacturing Convictions: Why Defendants Are Entitled to the Data Underlying Forensic DNA Kits*, 51 Duke L.J. 1097, 1119 (2001).

. As one commentator, William Thompson, put it, "If scientific evidence is not yet ready for both scientific scrutiny and public re-evaluation by others, it is not yet ready for court." *Id.* (quoting William C. Thompson, *Evaluating the Admissibility of New Genetic Identification Tests: Lessons from the "DNA War,"* 84 J. Crim. L. & Criminology 22, 100

(1993)). As for the purported copyright interest, the introduction of copyrighted material in court will not prevent the business from enforcing its copyright anywhere else.

Moreover, "forc[ing the public] to rely on the same [government] officials who are responsible for [presenting the evidence in court] to disclose and provide information about any difficulties with the [evidence]" does not comport with the First Amendment's requirement of narrow tailoring. *Woodford*, 299 F.3d at 880.

This makes it very likely that the public's oversight role would be realized in one form or another. Regardless, the complete denial of source code used on the public's behalf to seek to convict a criminal defendant would surely be an "exaggerated response" to private-interest concerns. *Woodford*, 299 F.3d at 880. In the context of the First Amendment analysis, the compelling nature of private concerns like trade secrets will be highly suspect when balanced against the momentous and bedrock constitutional rights held by a criminal defendant and the public.

4. CONCLUSION

For the foregoing reasons, this Court should dismiss this petition and reinstate the trial court proceedings, including the court's discovery order.

Respectfully submitted,

Dated: July 2, 2018 */s/ Vera Eidelman*

Bardis Vakili (SBN 247783)
American Civil Liberties Union Foundation of San Diego and Imperial Counties
2760 Fifth Ave #300
San Diego, CA 92103
T: 619.232.2121
bvakili@aclusandiego.org

Vera Eidelman (SBN 308535)
Andrea Woods
Brett Max Kaufman
Brandon Buskey
Rachel Goodman
American Civil Liberties Union Foundation
125 Broad Street, 18th Floor
New York, New York 10004
T: 212.549.2500
veidelman@aclu.org

Attorneys for Amici Curiae

CERTIFICATE OF COMPLIANCE

I certify that the text in the attached Brief contains 12,843 words—as calculated by Microsoft Word, including footnotes but not the caption, the table of contents, the table of authorities, signature blocks, or this certification—and that this document was prepared in a 13-point Times New Roman font. *See* Rule of Court 8.204(c)(1), (3).

Dated: July 2, 2018

BY: */s/ Vera Eidelman*

Vera Eidelman (SBN 308535)
American Civil Liberties Union Foundation
125 Broad Street, 18th Floor
New York, New York 10004
T: 212.549.2500
veidelman@aclu.org