DECODING PROBABILISTIC GENOTYPING SOFTWARE

Questioning Forensics 2020
22 and You:
Fighting for Privacy & Justice in an Age of Genetic Surveillance
Jan. 14 & 15, 2020 Brooklyn Law School

Jeanna Matthews, Clarkson University
Jessica Goldthwaite, Legal Aid Society DNA Unit
Nathan Adams, Forensic Bioinformatics

Why, when, and by how much do the results of probabilistic genotyping software programs differ from one another?

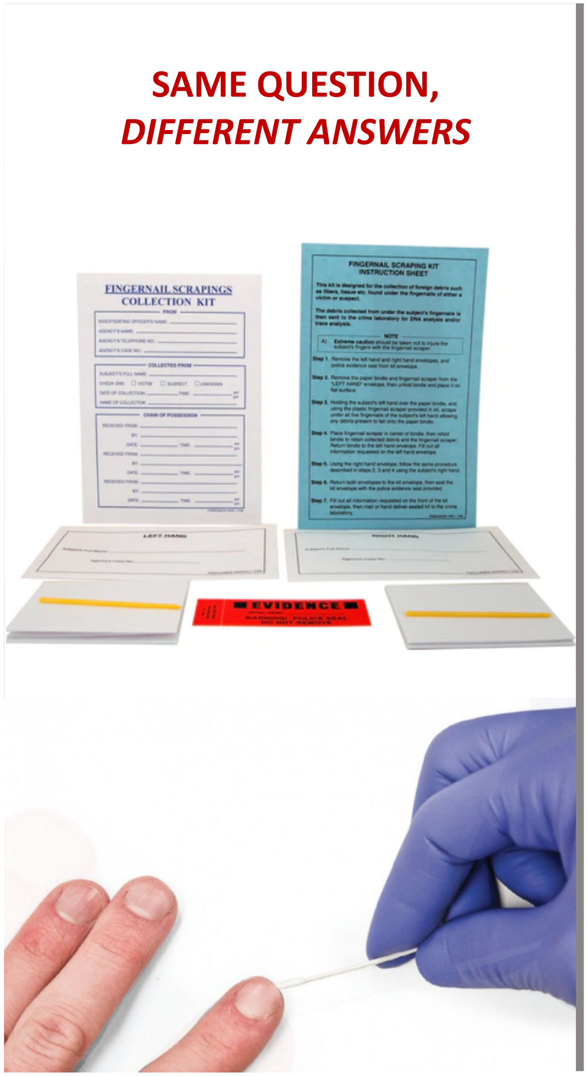
The New York Times

Oral Nicholas Hillary Acquitted in Potsdam Boy's Killing

Oral Nicholas Hillary embraced his lawyer Norman Siegel outside court on Wednesday after he was found not guilty of the 2011 murder of Garrett Phillips. Nathaniel Brooks for The New York Times

By Jesse McKinley

SAME QUESTION, *DIFFERENT ANSWERS*



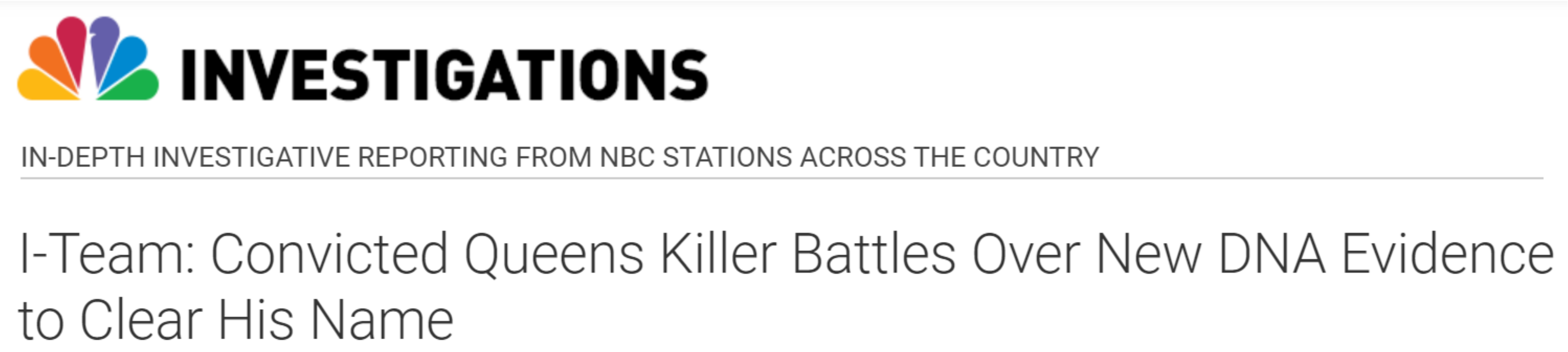
INCONCLUSIVE

Cybergenetics

NO STATISTICAL SUPPORT FOR A MATCH

STRmix™

10,000	Possible artifacts - removed
10,000,000	Possible artifacts - treated as real alleles
300,000	Possible artifacts – newer version models for them



INVESTIGATIONS

IN-DEPTH INVESTIGATIVE REPORTING FROM NBC STATIONS ACROSS THE COUNTRY

I-Team: Convicted Queens Killer Battles Over New DNA Evidence to Clear His Name

"After being told the [medical examiner's] lab found inconclusive results from the DNA samples, Legal Aid enlisted the analysis of Cybergenetics, a forensic firm NYC hired in 2005 to identify victims of the 911 World Trade Center terror attacks.

"In court on Sept. 16 the lead scientist testified that the patented TrueAllele method had also been used by prosecutors in four different New York jurisdictions and by State Police.

SAME QUESTION, *DIFFERENT MAGNITUDE IN ANSWERS*

STRmix™

1st Report:

Exclusionary LR of **666**
Called "uninformative" by the lab

2nd Report (more data):

Exclusionary LR of **1980**
"Supports that Michael Robinson is excluded as a contributor to this sample."

Cybergenetics

A match between the fingernail and Michael Robinson is:

1.18 **trillion** times *less* probable than a coincidental match to an unrelated Asian person

to

78.1 **trillion** times *less* probable than a coincidental match to an unrelated African-American person

ACM US Policy Council

1. Awareness
2. Access and redress
3. Accountability
4. Explanation
5. Data Provenance
6. Auditability
7. Validation and Testing

Forensic Statistical Tool (FST)

- 2010 Dec – Approved
- 2011 Apr – Online for casework
- 2011 Apr – Offline for bug fixes
- 2011 Apr-Jun – Bug fixes
- 2011 Jul – Online for casework
- 2016 Sep – First independent review
- 2017 Jan – First acknowledgement of defects
- Today – Still being used

System	Developer	Open-source	Cost	# Cases worldwide
FST	NYC (contractor)	v2.5 since 2017	N/A	1,350+ (2017)
STRmix	Aus/NZ	No	$20,000+	100,000+ (2019)
TrueAllele	Cybergenetics, Inc.	No	$50,000+	Unknown
EuroForMix	EuroForGen	Yes	Free	Unknown

Brown Institute of Columbia and Stanford
Magic Grant
2018-2019

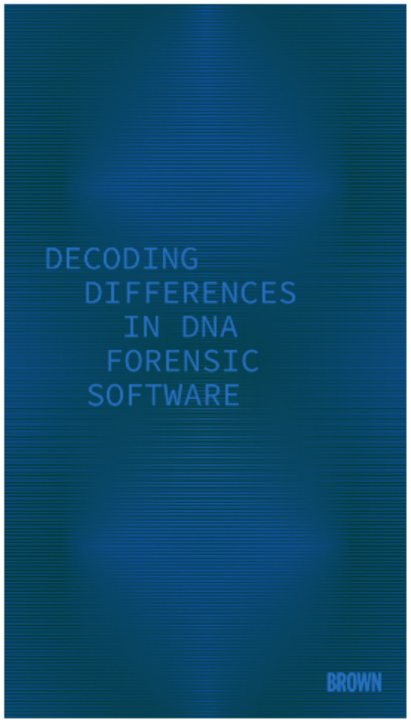
Decoding Differences in Forensic DNA Software

THE BROWN INSTITUTE
FOR MEDIA INNOVATION

Columbia Journalism School
Stanford | ENGINEERING

Decoding Differences in DNA Forensic Software

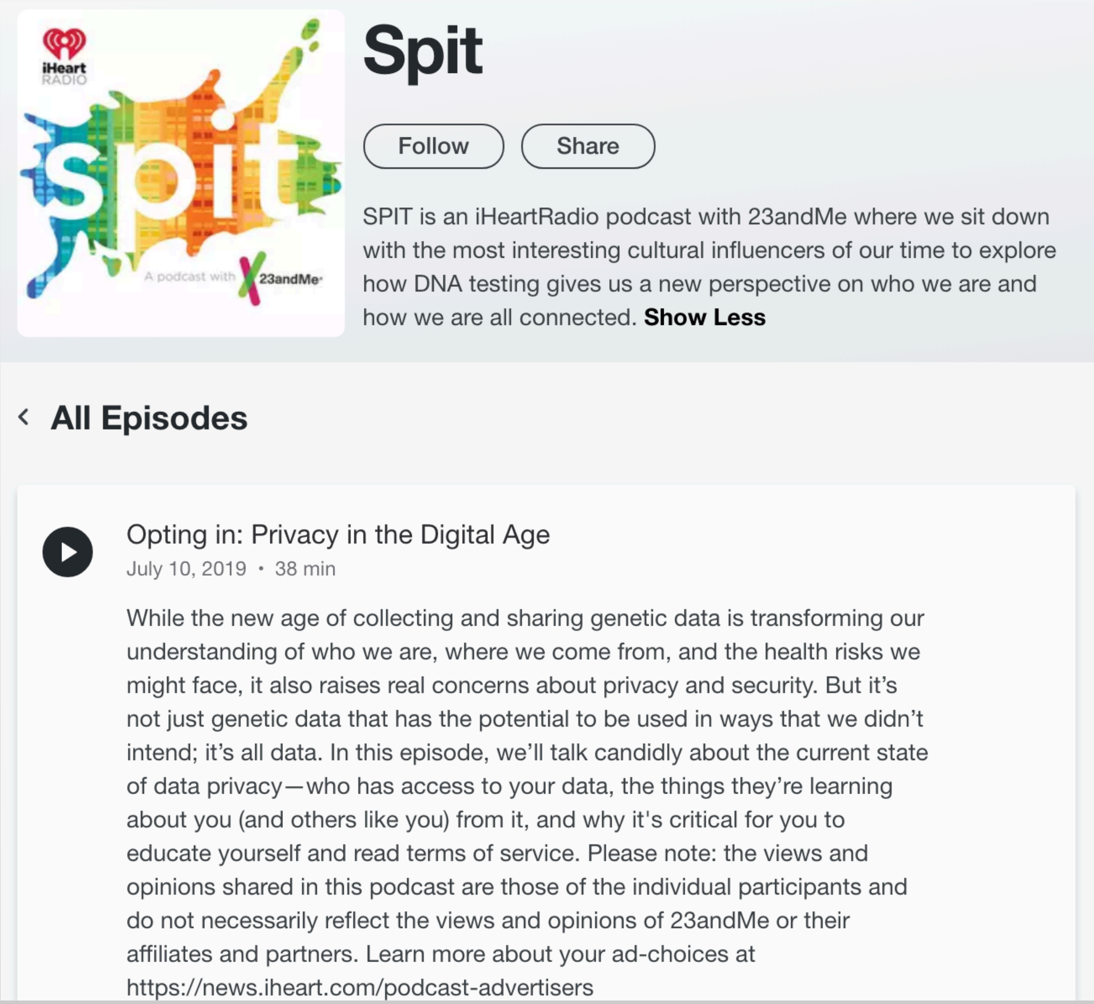
2018-2019



DECODING
DIFFERENCES
IN DNA
FORENSIC
SOFTWARE

BROWN

Imagine testing the fingernail scrapings of a murder victim to determine if a suspect could be the killer, only to have one DNA interpretation software program incriminate the suspect and a different program absolve them. Such a scenario played out two years ago in the widely-publicized murder trial of Oral Nicholas Hillary, raising questions that the criminal justice system still cannot answer: why, when, and by how much do these programs differ from one another? To answer these questions, this Magic Grant assembles a multi-disciplinary team — Jeanna Matthews is a Computer Scientist; Nathan Adams, a DNA investigations specialist; Jessica Goldthwaite with The Legal Aid Society; Dan Krane, a Biologist; Surya Mattu, a Journalist; and David Madigan, a Statistician. This Magic Grant project will systematically compare forensic DNA software, moving the story beyond anecdotal examples to a systematic investigative strategy. In the process, they will explore important issues of algorithmic transparency, and the role of complex software systems in the criminal justice system and beyond.



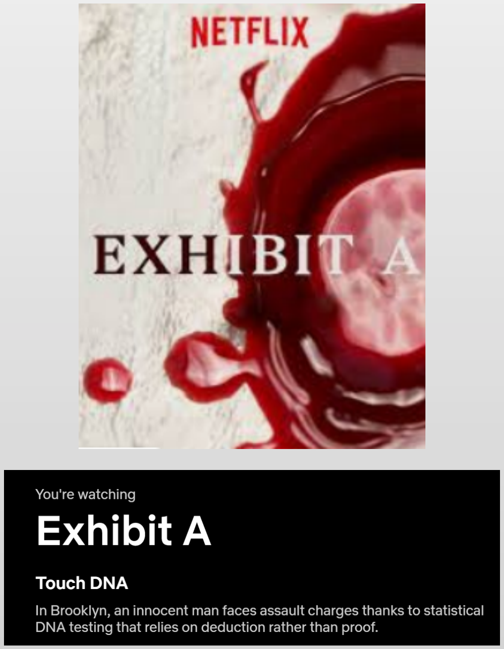
Spit

SPIT is an iHeartRadio podcast with 23andMe where we sit down with the most interesting cultural influencers of our time to explore how DNA testing gives us a new perspective on who we are and how we are all connected.

Opting in: Privacy in the Digital Age

July 10, 2019 • 38 min

While the new age of collecting and sharing genetic data is transforming our understanding of who we are, where we come from, and the health risks we might face, it also raises real concerns about privacy and security. But it's not just genetic data that has the potential to be used in ways that we didn't intend; it's all data. In this episode, we'll talk candidly about the current state of data privacy—who has access to your data, the things they're learning about you (and others like you) from it, and why it's critical for you to educate yourself and read terms of service. Please note: the views and opinions shared in this podcast are those of the individual participants and do not necessarily reflect the views and opinions of 23andMe or their affiliates and partners. Learn more about your ad-choices at https://news.iheart.com/podcast-advertisers

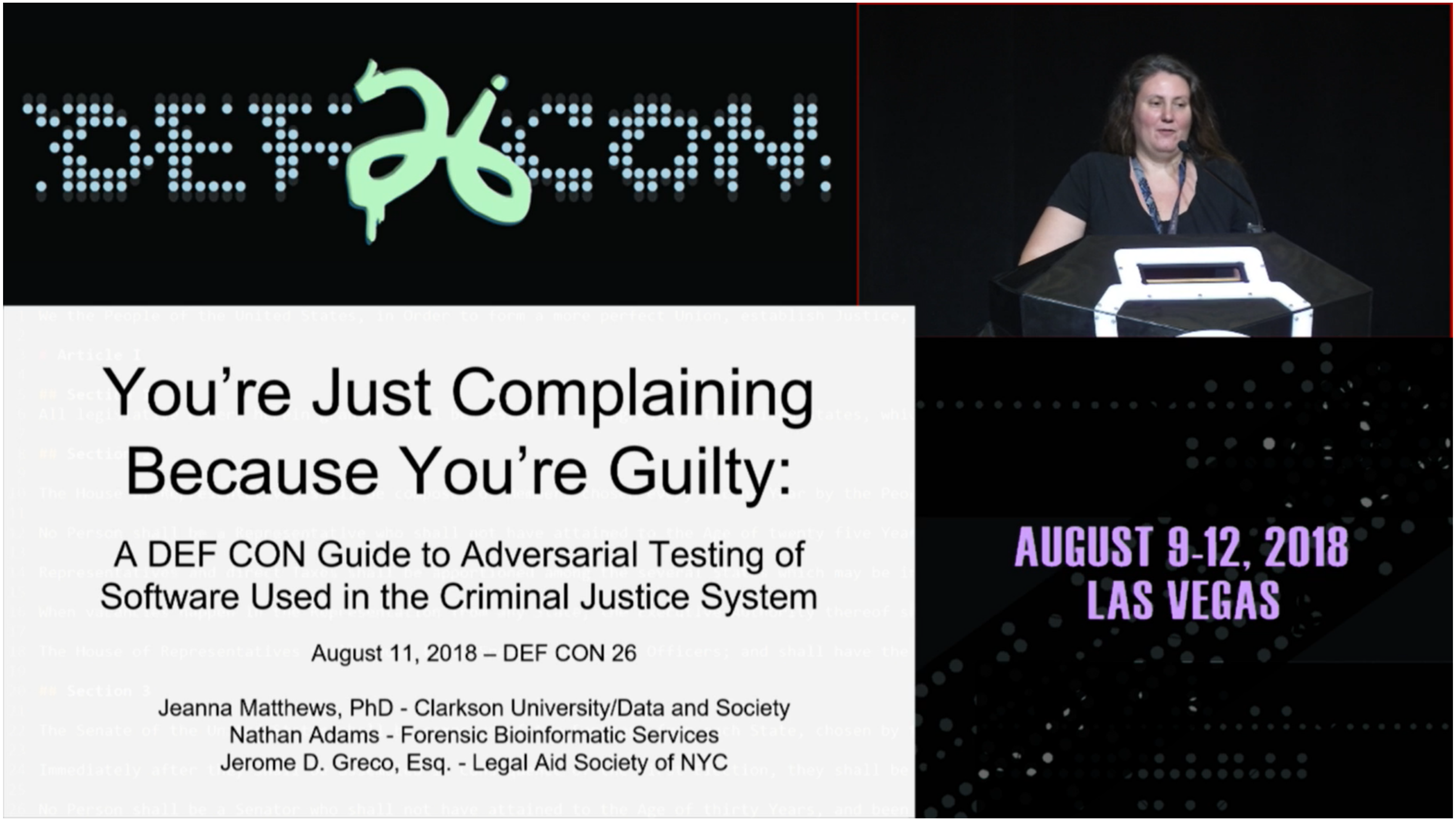


You're watching

Exhibit A

Touch DNA

In Brooklyn, an innocent man faces assault charges thanks to statistical DNA testing that relies on deduction rather than proof.



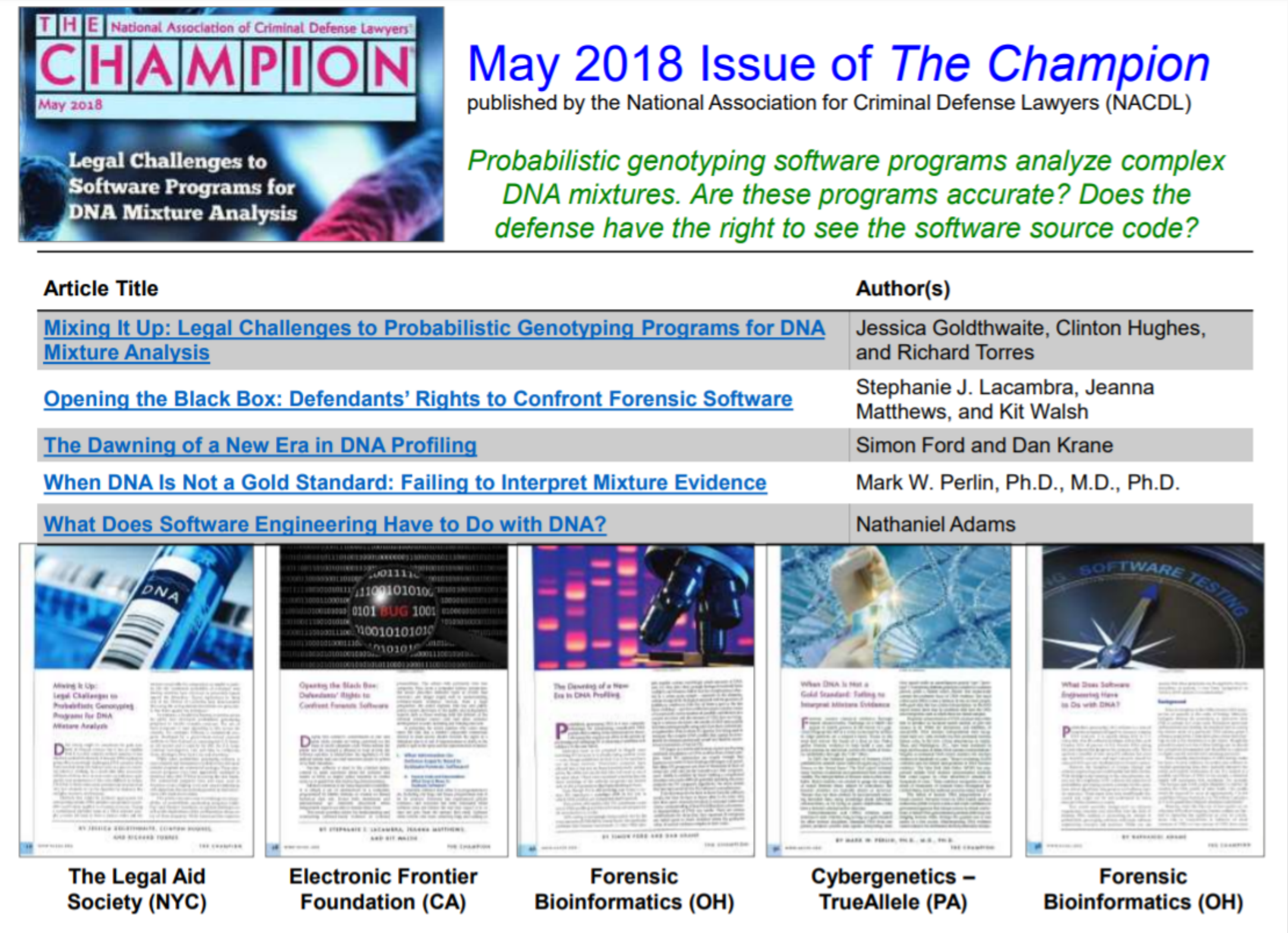
You're Just Complaining Because You're Guilty:

A DEF CON Guide to Adversarial Testing of Software Used in the Criminal Justice System

August 11, 2018 – DEF CON 26

Jeanna Matthews, PhD - Clarkson University/Data and Society
Nathan Adams - Forensic Bioinformatic Services
Jerome D. Greco, Esq. - Legal Aid Society of NYC

AUGUST 9-12, 2018
LAS VEGAS



May 2018 Issue of *The Champion*

published by the National Association for Criminal Defense Lawyers (NACDL)

Probabilistic genotyping software programs analyze complex DNA mixtures. Are these programs accurate? Does the defense have the right to see the software source code?

Article Title	Author(s)
Mixing It Up: Legal Challenges to Probabilistic Genotyping Programs for DNA Mixture Analysis	Jessica Goldthwaite, Clinton Hughes, and Richard Torres
Opening the Black Box: Defendants' Rights to Confront Forensic Software	Stephanie J. Lacambra, Jeanna Matthews, and Kit Walsh
The Dawning of a New Era in DNA Profiling	Simon Ford and Dan Krane
When DNA Is Not a Gold Standard: Failing to Interpret Mixture Evidence	Mark W. Perlin, Ph.D., M.D., Ph.D.
What Does Software Engineering Have to Do with DNA?	Nathaniel Adams

The Legal Aid Society (NYC)

Electronic Frontier Foundation (CA)

Forensic Bioinformatics (OH)

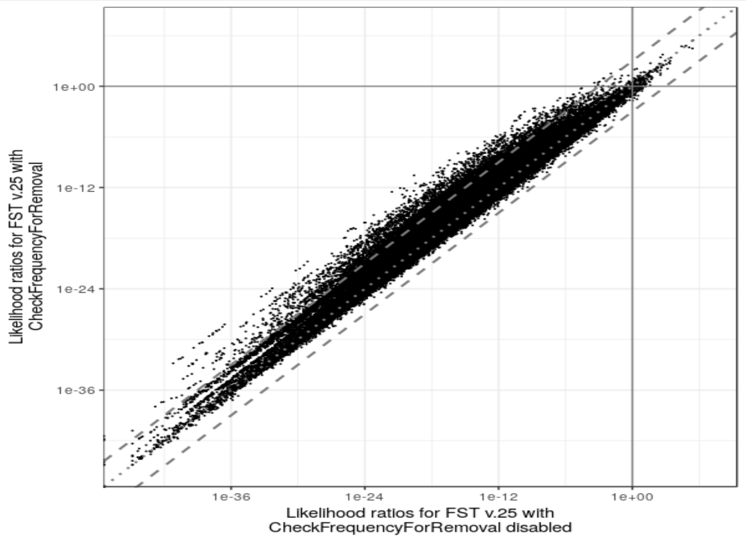
Cybergenetics – TrueAllele (PA)

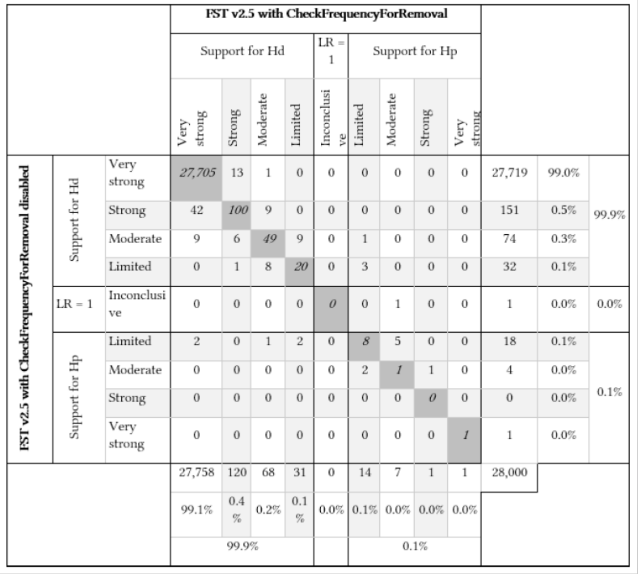
Forensic Bioinformatics (OH)

Slide from John Butler, NIST, DNA Mixture Interpretation Principles: Observations from a NIST Scientific Foundation Review. AAFS 2019 Workshop #10 (February 18, 2019; Baltimore, MD)

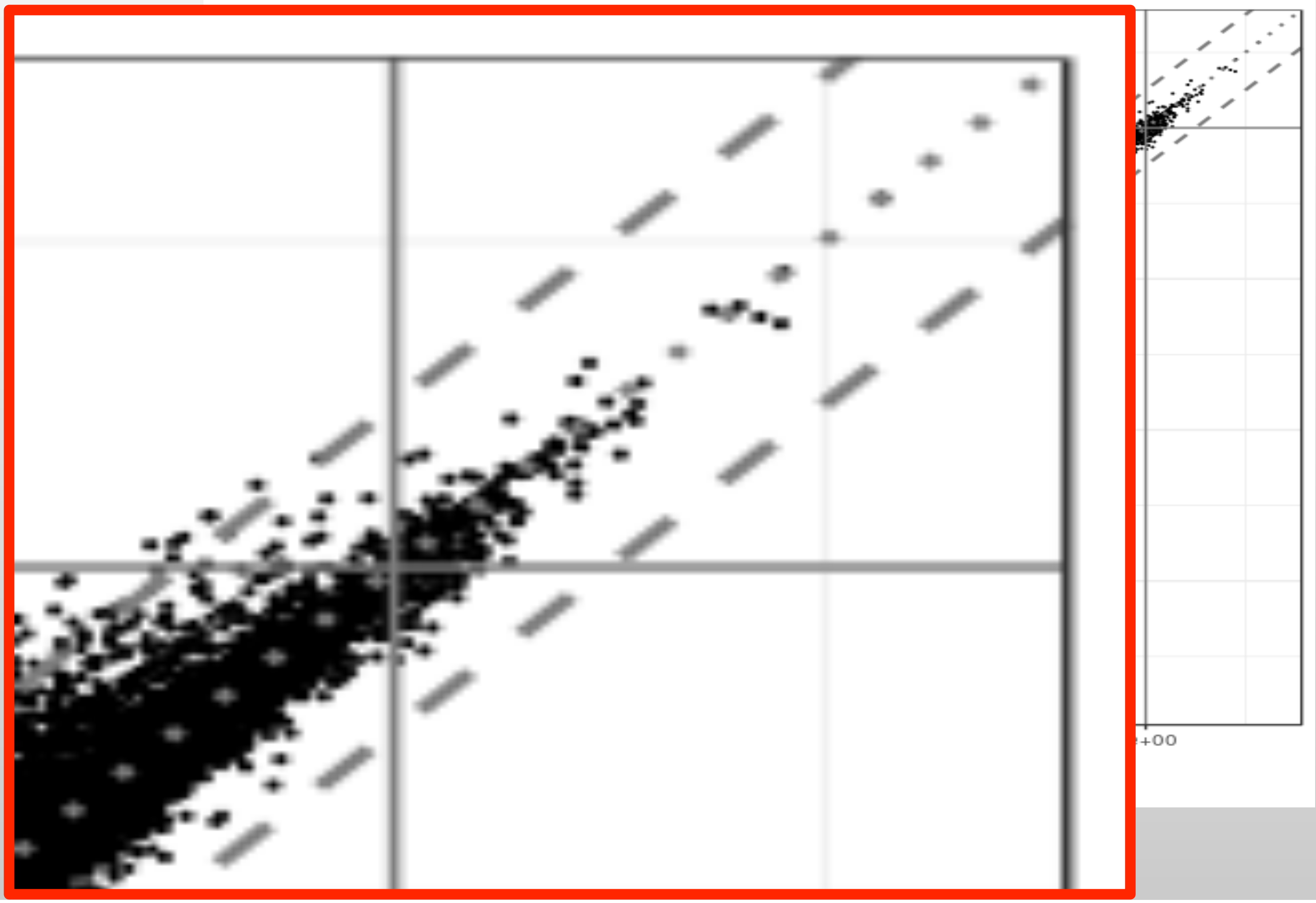
AAAI/ACM Artificial Intelligence, Ethics, and Society 2019

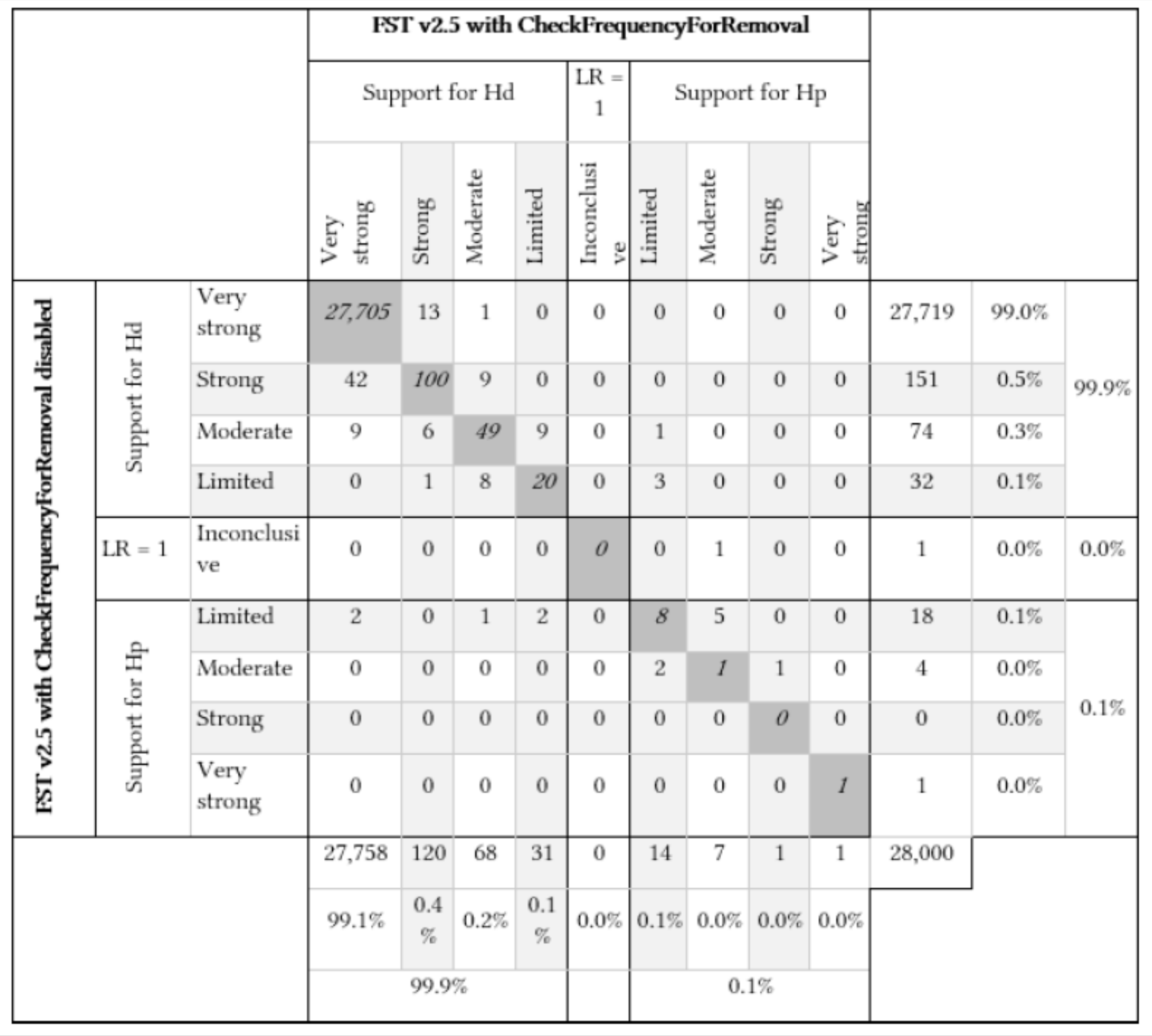
“The Right To Confront Your Accusers: Opening the Black Box of Forensic DNA Software”





			FST v2.5 with CheckFrequencyForRemoval											
			Support for Hd				LR = 1	Support for Hp						
			Very strong	Strong	Moderate	Limited	Inconclusive	Limited	Moderate	Strong	Very strong			
FST v2.5 with CheckFrequencyForRemoval disabled	Support for Hd	Very strong	*27,705*	13	1	0	0	0	0	0	0	27,719	99.0%	99.9%
		Strong	42	*100*	9	0	0	0	0	0	0	151	0.5%	
		Moderate	9	6	*49*	9	0	1	0	0	0	74	0.3%	
		Limited	0	1	8	*20*	0	3	0	0	0	32	0.1%	
	LR = 1	Inconclusive	0	0	0	0	*0*	0	1	0	0	1	0.0%	0.0%
	Support for Hp	Limited	2	0	1	2	0	*8*	5	0	0	18	0.1%	0.1%
		Moderate	0	0	0	0	0	2	*1*	1	0	4	0.0%	
		Strong	0	0	0	0	0	0	0	*0*	0	0	0.0%	
		Very strong	0	0	0	0	0	0	0	0	*1*	1	0.0%	
			27,758	120	68	31	0	14	7	1	1	28,000		
			99.1%	0.4%	0.2%	0.1%	0.0%	0.1%	0.0%	0.0%	0.0%			
			99.9%					0.1%						

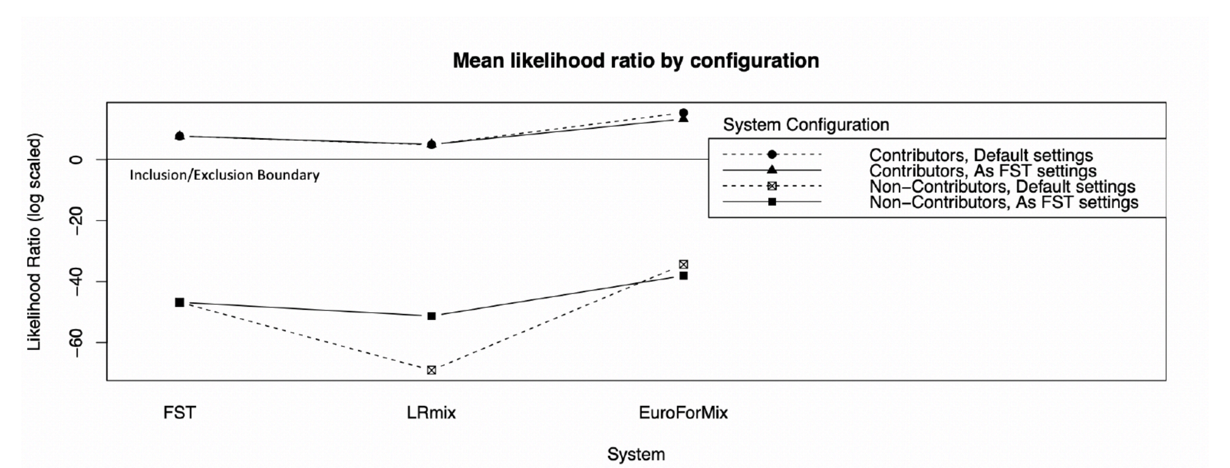

e+00



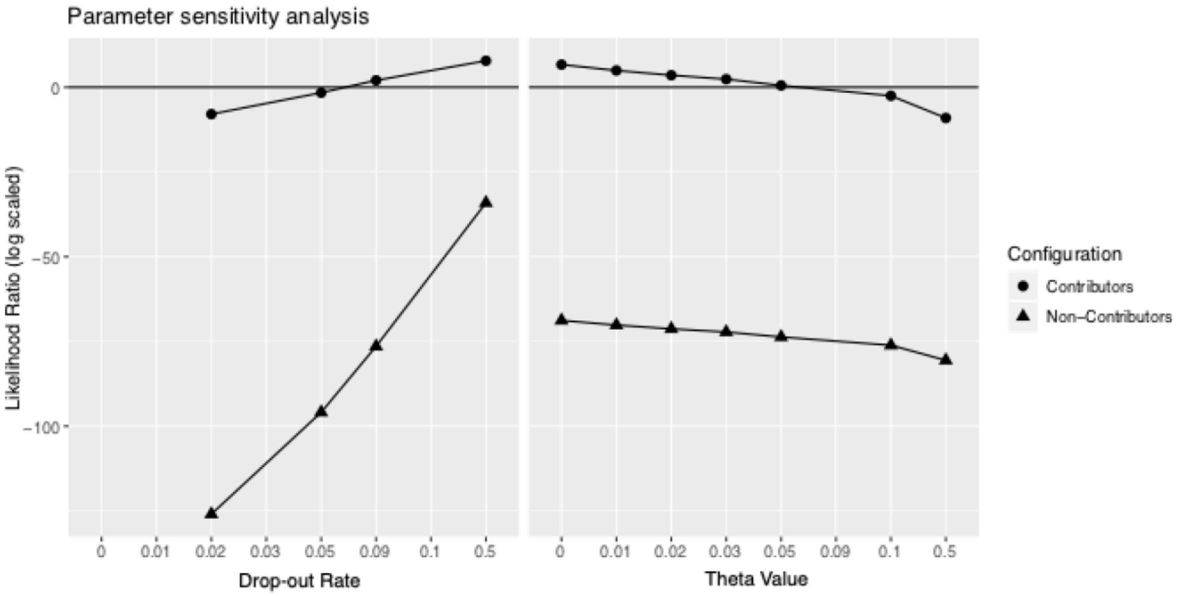
			FST v2.5 with CheckFrequencyForRemoval											
			Support for Hd				LR = 1	Support for Hp						
			Very strong	Strong	Moderate	Limited	Inconclusive	Limited	Moderate	Strong	Very strong			
FST v2.5 with CheckFrequencyForRemoval disabled	Support for Hd	Very strong	*27,705*	13	1	0	0	0	0	0	0	27,719	99.0%	99.9%
		Strong	42	*100*	9	0	0	0	0	0	0	151	0.5%	
		Moderate	9	6	*49*	9	0	1	0	0	0	74	0.3%	
		Limited	0	1	8	*20*	0	3	0	0	0	32	0.1%	
	LR = 1	Inconclusive	0	0	0	0	*0*	0	1	0	0	1	0.0%	0.0%
	Support for Hp	Limited	2	0	1	2	0	*8*	5	0	0	18	0.1%	0.1%
		Moderate	0	0	0	0	0	2	*1*	1	0	4	0.0%	
		Strong	0	0	0	0	0	0	0	*0*	0	0	0.0%	
		Very strong	0	0	0	0	0	0	0	0	*1*	1	0.0%	
			27,758	120	68	31	0	14	7	1	1	28,000		
			99.1%	0.4%	0.2%	0.1%	0.0%	0.1%	0.0%	0.0%	0.0%			
			99.9%					0.1%						

AAAI/ACM
Artificial Intelligence, Ethics, and Society 2020

“When Trusted Black Boxes Don’t Agree: Incentivizing Iterative Improvement and Accountability in Critical Software Systems”

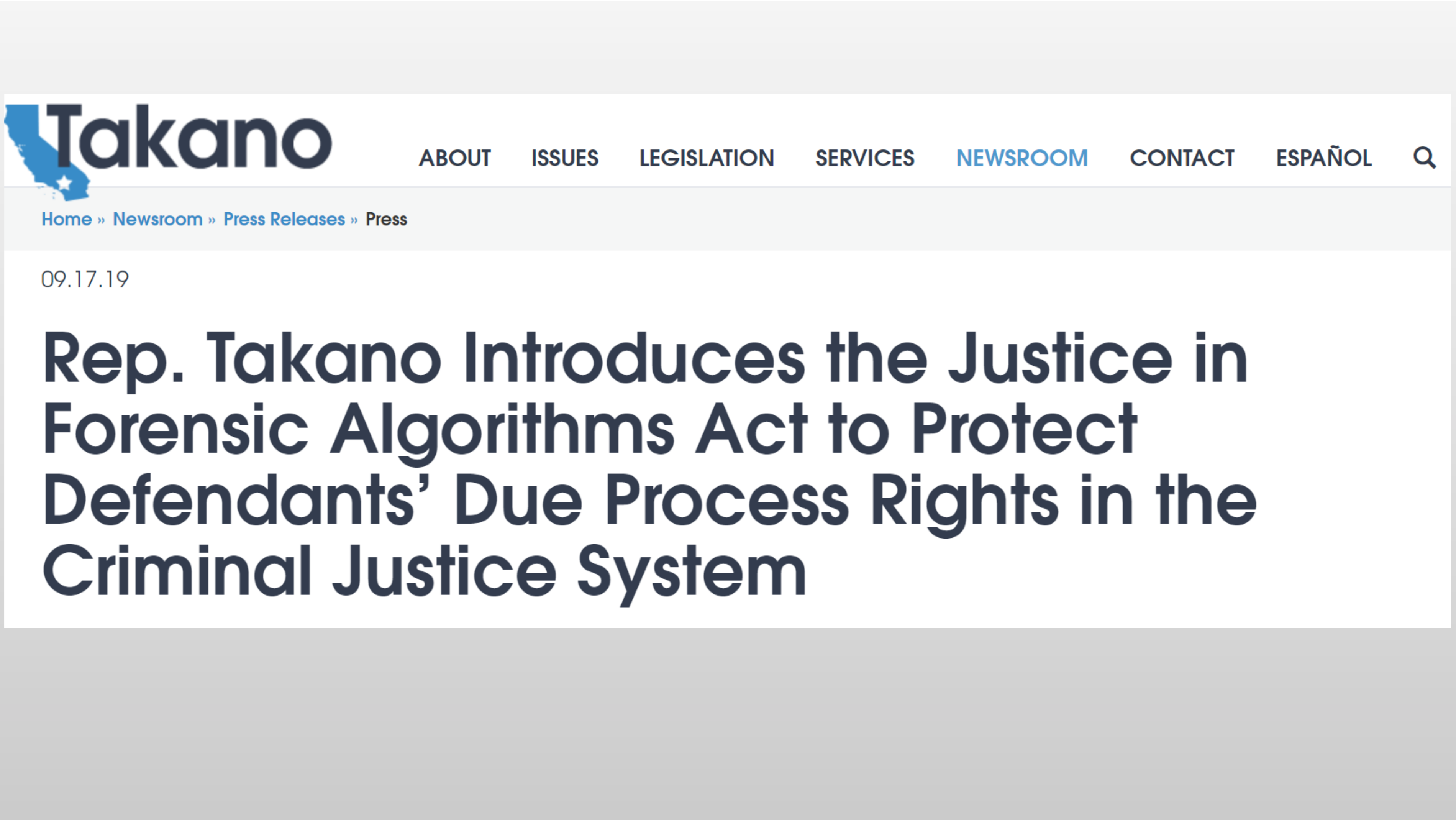


AIES 2020


Parameter sensitivity analysis
Likelihood Ratio (log scaled)
0
−50
−100
0
0.01
0.02
0.03
0.05
0.09
0.1
0.5
Drop-out Rate
0
0.01
0.02
0.03
0.05
0.09
0.1
0.5
Theta Value
Configuration
● Contributors
▲ Non−Contributors

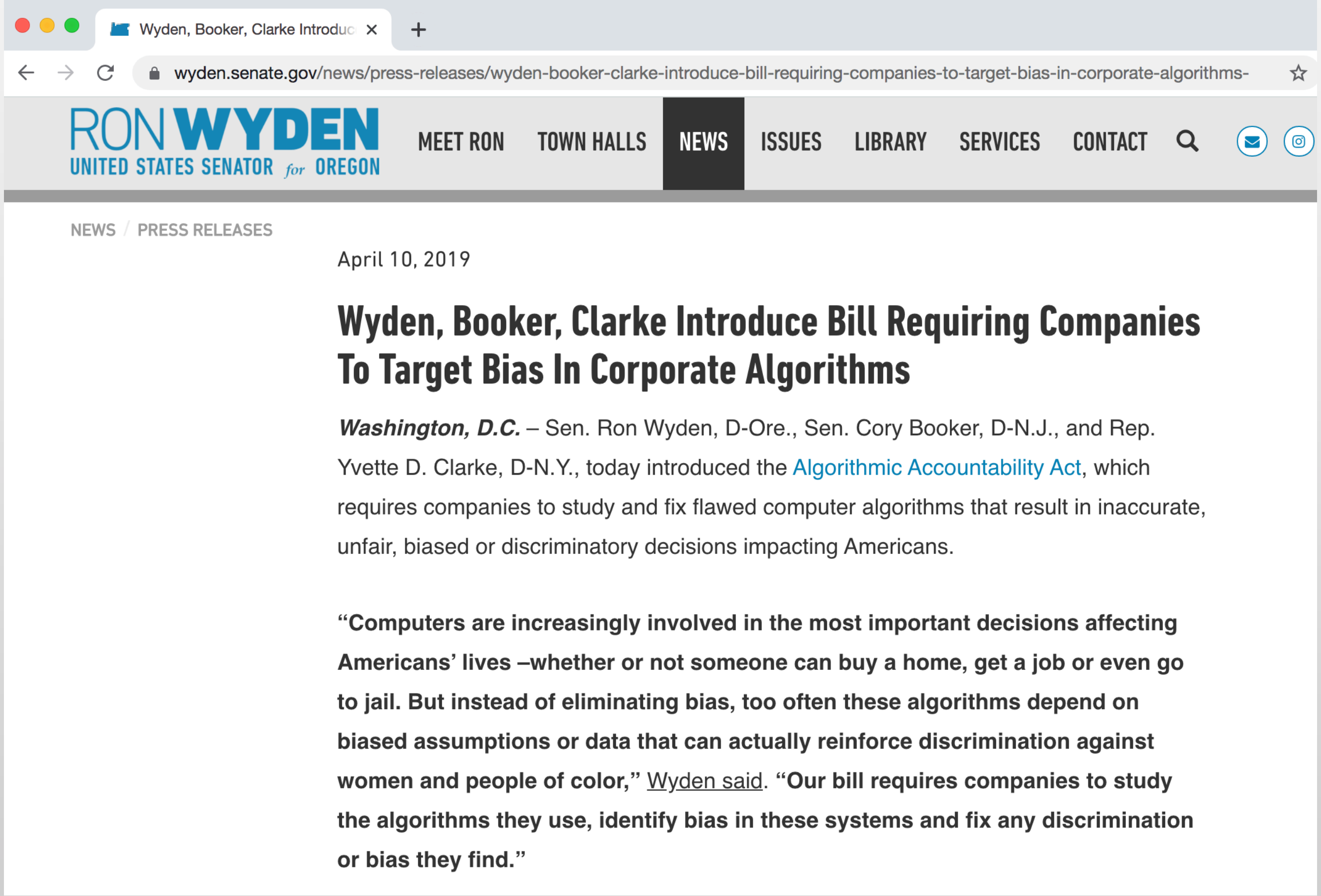
Procurement Contracts in Criminal Justice

- **Transparency** of the technology covered by the contract.
  - Require/reward open-source software, access to software engineering artifacts including bug tracking/change log databases, internal testing plans and results, software requirements and specifications, hazard and risk assessments, design documents, etc
- **Warranties** for fitness for purpose.
- **Preservation** of each successive generation of technology in a format which can be examined at the request of the Jurisdiction, a defense attorney, or a competent Court.
- **Facilitate** adversarial testing.
  - Automated testing interfaces, common input formats and parameters, no prohibition on publishing independent testing results, bug bounties, low cost access to executables for the purpose of testing



09.17.19

Rep. Takano Introduces the Justice in Forensic Algorithms Act to Protect Defendants’ Due Process Rights in the Criminal Justice System



April 10, 2019

Wyden, Booker, Clarke Introduce Bill Requiring Companies To Target Bias In Corporate Algorithms

Washington, D.C. – Sen. Ron Wyden, D-Ore., Sen. Cory Booker, D-N.J., and Rep. Yvette D. Clarke, D-N.Y., today introduced the Algorithmic Accountability Act, which requires companies to study and fix flawed computer algorithms that result in inaccurate, unfair, biased or discriminatory decisions impacting Americans.

"Computers are increasingly involved in the most important decisions affecting Americans' lives –whether or not someone can buy a home, get a job or even go to jail. But instead of eliminating bias, too often these algorithms depend on biased assumptions or data that can actually reinforce discrimination against women and people of color," Wyden said. **"Our bill requires companies to study the algorithms they use, identify bias in these systems and fix any discrimination or bias they find."**

Meta-analysis

Validation review: General operation of software in lab

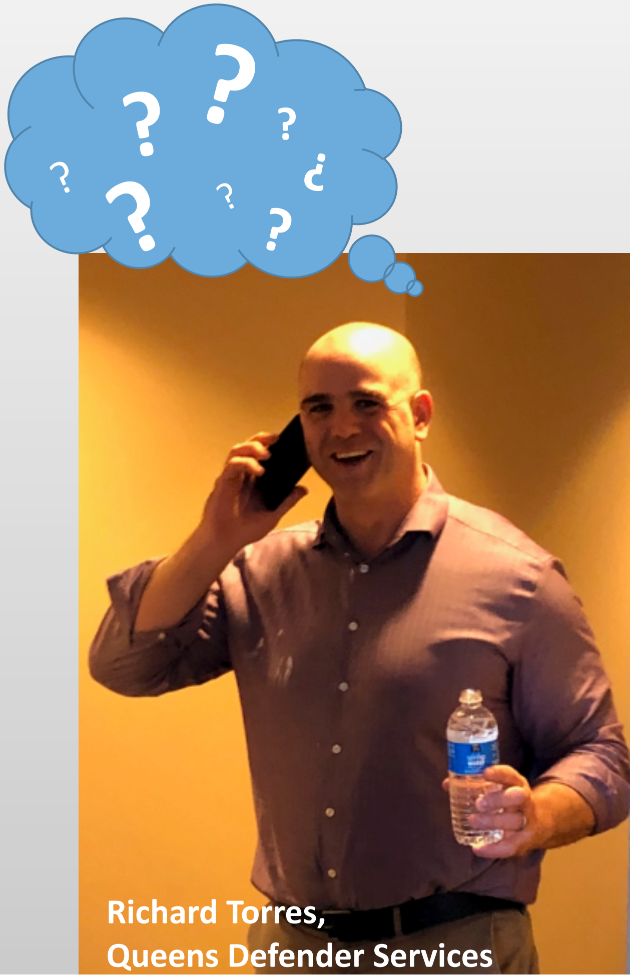
Casework-to-validation comparison: Where do the samples in your case fall in terms of samples studied during validation?

Meta-analysis

How many false inclusions/exclusions are there at particular weights?

Do the STRmix secondary diagnostics give any clue as to whether the results are unreliable?

How closely does STRmix assign mixture proportions?


Richard Torres,
Queens Defender Services

Decisions on PG

People v. Thompson

"…we should not toss unresolved scientific debates into judges' chambers, and especially not into the jury room. That conclusion applies to FST evidence—still."

65 Misc.3d 1206(A) (Sup.Ct. N.Y. Co. September 25, 2019)

US v. Gissantaner

“Here, because the sum of the parts simply does not add up to a reliable whole, the DNA analysis/likelihood ratio resulting from the use of the STRmix probabilistic genotyping software must be excluded.”

1:17-cr-00130-JTN ECF 161 (W. Dist. MI, October 16, 2019)

US v. Elmore

“If SERI could accurately identify five-person mixtures and if it had validated Bullet to analyze them, then it might have a reliable understanding of how underestimating a five-person mixture impacts the likelihood ratio.... But there are simply too many reasons to question the reliability of [analyst]’s conclusion on this foundational issue, which brings the entire analysis outside the parameters of Bullet’s validation at SERI. ”

3:13-cr-00764-WHO Document 2083 (N. Dist. CA, April 29, 2019)

US v. Oldman

“However, Ms. Conway noted that if the apparent number of contributors was five as opposed to four the literature is clear that the result would either be the same likelihood ratio or more conservative.”

“…the FBI lab has only validated STRmix™ up to 10:1 ratio…. In this case, the swab from the basement stairs had a ratio of 18:1… shirt had a ratio of 15:1…. any discrepancy in the ratios does not render [stairs and shirt] inadmissible.”

2:18-cr-00020-SWS Document 227 (Dist. WY, Dec 31, 2018)

US v. Lewis

0:18-cr-00194-ADM-DTS (Dist. MN)

TrueAllele exclusion

PUBLICATION BAN