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

Motivation/overview of the problem

Black box decision making

- Software is increasingly used to make important decisions about people's lives
  - Hiring, housing, how we make friends, find partners, navigate city streets, get our news, …
  - The weightier the decision the more crucial it is that we understand and can question it
  - What input is used to make the decision? Is it correct? Do we have other information that should be considered?
  - Are protected attributes like race and gender used? What about proxies for those characteristics?
- Criminal justice system
  - Software/algorithmic decision making used increasingly throughout the criminal justice system
  - Often black boxes for which trade secret protection is claimed to be more important than rights of individual defendants or citizens to understand the decisions
  - Evidence of problems
  - How can we find bugs and fix problems if the answer is always "you can't question" and "you are just complaining because you are guilty"?

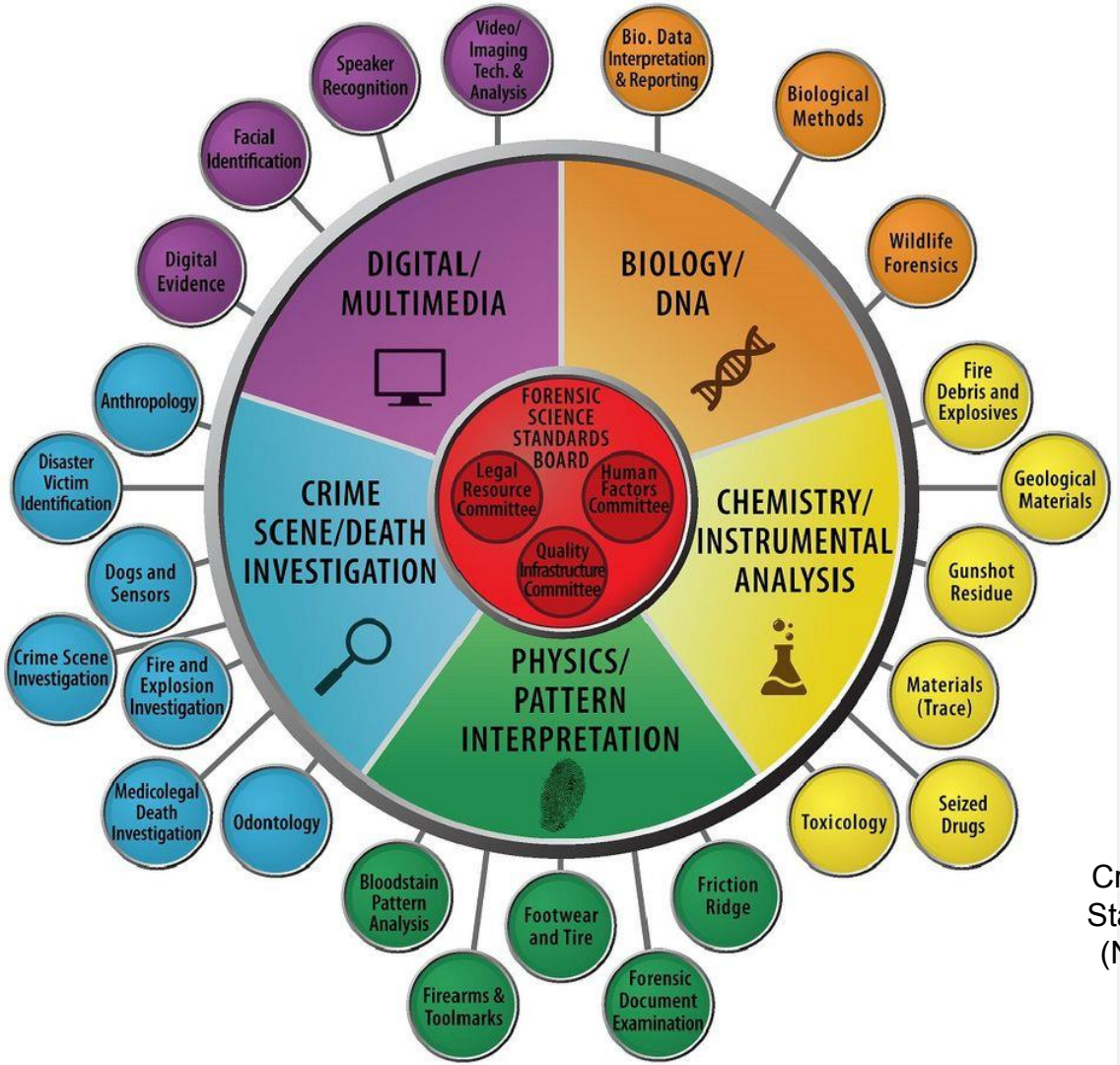
Can you imagine...

- Being sent to prison rather than given probation because proprietary software says you are likely to commit another crime?
  - But you can’t ask how the software makes its decisions. (Eric Loomis)
- Having the primary evidence against you being the results of DNA software?
  - But one program says you did it and another says you didn’t. (Nick Hillary)
- Being accused of murder solely because of DNA transferred by paramedics?
  - But they don’t figure that out for months. (Lukis Anderson)

- Software and complex systems need an iterative process of debugging and improvement!
- Anyone who has used technology knows that there are glitches and bugs and unintended consequences!
- Anyone who builds technology knows how easy it is for there to be substantial bugs you did not find!
- Huge advantages to independent, third-party testing aimed at finding bugs!
- If only those with interests in the success of software see the details, we have a huge problem and a recipe for injustice!



An Overview of Problematic Technology Used in the Criminal Justice System



Credit: National Institute of Standards and Technology (NIST) - The Organization of Scientific Area Committees (OSAC)


Individualized assessment
Predictive policing
Evidence assessment
Evidence gathering

Law Enforcement Tech by Secrecy Level*

Secret

We don't want you to know it exists and/or that we have it.

- Cell-Site Simulators
- Hemisphere Project
- PRISM
- Backscatter X-Ray Vans
- Drone Surveillance

Secret as Applied

We have it but we won't tell you when and/or how we used it.

- Automated License Plate Readers
- Facial Recognition/Capture
- Domain Awareness System
- Police Internal Databases
  - Real Time Crime Center
  - Gang databases
  - Social media analytics
  - Etc.
- Predictive Policing

Trust Us

We have it. We used it here. Stop asking questions.

- DNA Probabilistic Genotyping Software
- Bail/Parole/Sentencing Determination Algorithms
- ShotSpotter
- Cellebrite Advanced Services and Graykey
- P2P/Child Pornography Investigative Software
- Network Investigative Techniques (NITs)
- Alcohol breath testing

*Not comprehensive of all available technology. Some technologies fit under different levels based on the jurisdiction and agency.

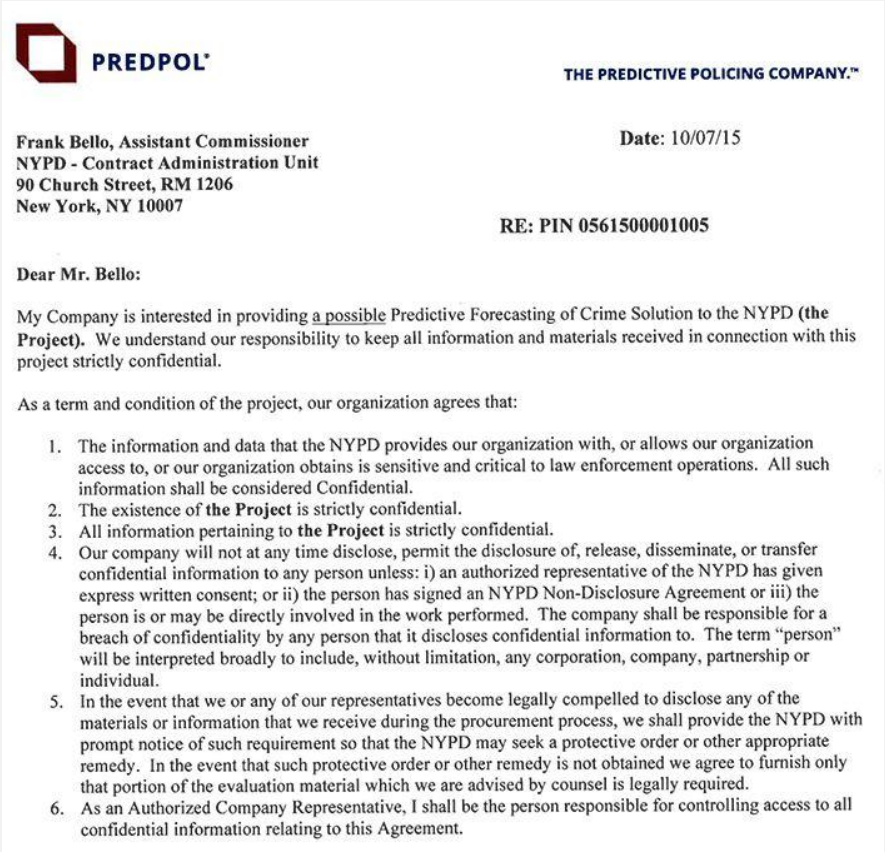
Predictive policing



- Caution needed during a stop
- Types of policing needed
- Areas to police
- People to stop
- Gang and affiliation databases
- People likely to become victims

Predictive Policing, Flawed Data, and Flawed Results

- Bad data in = bad data out
- Racial disparities
- Sources of data
- Presumption of guilt by association
- Constitutional rights of individuals
- Lack of Transparency and Public Debate
  - Non-Disclosure Agreements (NDAs)
  - Proprietary trade secrets
  - Sensitive data



PREDPOL®

THE PREDICTIVE POLICING COMPANY.™

Frank Bello, Assistant Commissioner
NYPD - Contract Administration Unit
90 Church Street, RM 1206
New York, NY 10007

Date: 10/07/15

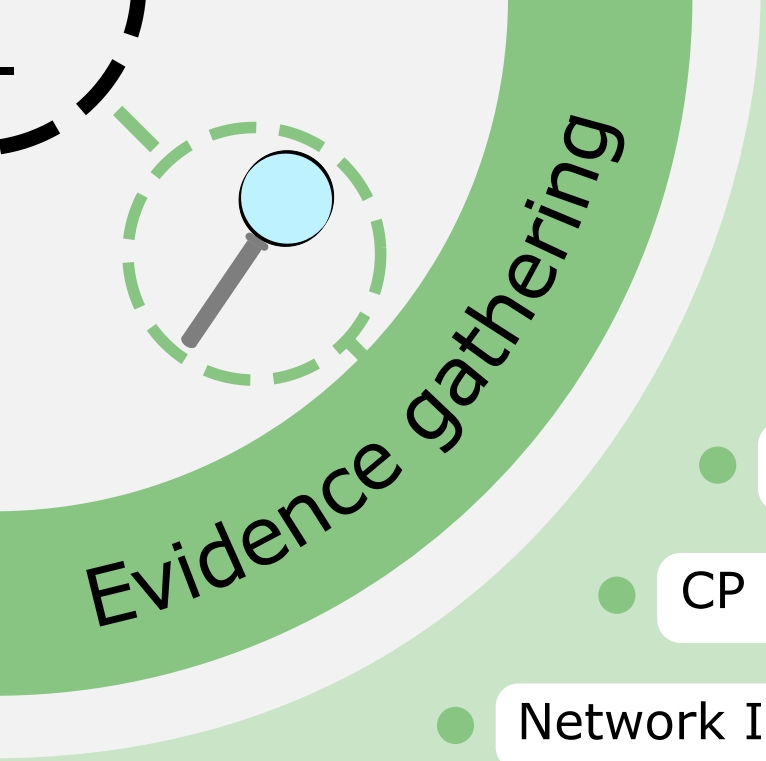
RE: PIN 0561500001005

Dear Mr. Bello:

My Company is interested in providing a possible Predictive Forecasting of Crime Solution to the NYPD **(the Project).** We understand our responsibility to keep all information and materials received in connection with this project strictly confidential.

As a term and condition of the project, our organization agrees that:

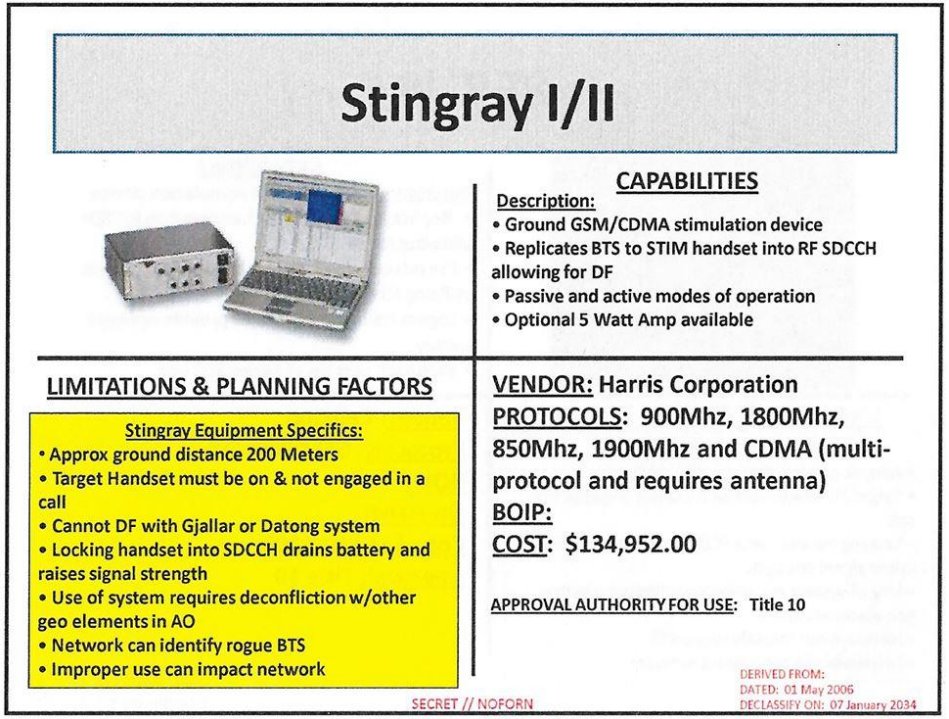
1. The information and data that the NYPD provides our organization with, or allows our organization access to, or our organization obtains is sensitive and critical to law enforcement operations. All such information shall be considered Confidential.
2. The existence of **the Project** is strictly confidential.
3. All information pertaining to **the Project** is strictly confidential.
4. Our company will not at any time disclose, permit the disclosure of, release, disseminate, or transfer confidential information to any person unless: i) an authorized representative of the NYPD has given express written consent; or ii) the person has signed an NYPD Non-Disclosure Agreement or iii) the person is or may be directly involved in the work performed. The company shall be responsible for a breach of confidentiality by any person that it discloses confidential information to. The term "person" will be interpreted broadly to include, without limitation, any corporation, company, partnership or individual.
5. In the event that we or any of our representatives become legally compelled to disclose any of the materials or information that we receive during the procurement process, we shall provide the NYPD with prompt notice of such requirement so that the NYPD may seek a protective order or other appropriate remedy. In the event that such protective order or other remedy is not obtained we agree to furnish only that portion of the evaluation material which we are advised by counsel is legally required.
6. As an Authorized Company Representative, I shall be the person responsible for controlling access to all confidential information relating to this Agreement.



- ShotSpotter
- Cell-site simulators (Stingray)
- Facial recognition
- Automated License Plate Readers (ALPRs)
- CP investigation (P2P, IP tracing)
- Network Investigation Techniques (NITs)
- Mobile device cracking (Cellebrite, Gray Key)

Cell-Site Simulators (aka Stingray Devices)

- Mimics a cell phone tower and emits a signal that compels cell phones in the area to connect to it rather than a legitimate tower
- Not all cell-site simulators are “Stingrays”
- Non-Disclosure Agreements (NDAs)
- NYPD used 1,000+ times from 2008 to 2015 without once getting a warrant
- *U.S. v. Lambis*, 197 F. Supp. 3d 606 (S.D.N.Y. 2016)
- *People v. Gordon,* 58 Misc. 3d 544 (N.Y. Sup. Ct. 2017)
- *Carpenter v. United States*, 16-402, 2018 WL 3073916 (2018)



People v. Gordon and the Use of Cell-Site Simulators

The Concession

Sections I-V of The Defendant's motion contend that a Cell Site Simulator was used, without a court order, to locate The Defendant. The Defendant further asserts that the use of the Cell Site Simulator, for various reasons, violated The Defendant's Constitutional rights. In this case, a Cell Site Simulator was used pursuant to lawfully Court Orders issued by the Honorable Justice Alan Marrus. Attached as *In Camera* Exhibit 1 and *In Camera* Exhibit 2 are copies of the Sealed Order to Sprint Corporation and the Order to Authorize.

People v. Gordon and the Use of Cell-Site Simulators

The Decision

Therefore, the failure to obtain a proper eavesdropping warrant here prejudiced the *defendant*since the most useful-and needed information-*ie*. his location-was procured from the unlimited use of the *cell site simulator*.

The NYPD's Post-Decision Denial*

But the New York City Police Department on Wednesday took issue with Murphy's decision, arguing that the judge was simply wrong on key factual points, including about whether a cell site simulator was used to locate the defendant in the case and potentially about the type of warrant issued in the investigation.

*Probable-Cause Warrant Needed for Cell-Tracking, Brooklyn Judge Rules by Jason Grant (New York Law Journal) (November 15, 2017)

“He thought of the telescreen with its never-sleeping ear. They could spy upon you night and day, but if you kept your head you could still outwit them. With all their cleverness they had never mastered the secret of finding out what another human being was thinking.”

Quote from 1984 by George Orwell

Mobile Digital Forensics and the Encryption War

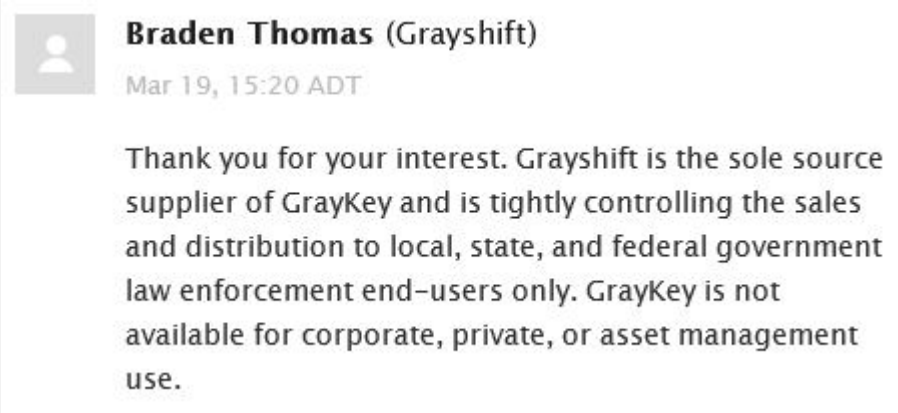
- *Riley v. California*, 134 S. Ct. 2473 (2014)
- Cellebrite UFED Touch2
  - Cellebrite is a digital forensics company specializing in mobile devices
  - UFED = Universal Forensic Extraction Device
- Magnet Axiom
- Paraben E3
- Extraction of data (extraction of your life)
- Available Outside of Law Enforcement



IT IS FURTHER ORDERED that a search of all files and data stored in the target devices is authorized, irrespective of how the data is filed, labeled, designated, encrypted, hidden, disguised or otherwise stored.

Cellebrite Advanced Services (CAS) and GrayKey

- 2015 Attack in San Bernardino
- Cellebrite Advanced Services (CAS)
  - Secret process performed by Cellebrite at a Cellebrite lab
  - Reportedly $1,500 per phone or a $250,000 a year subscription
- GrayKey by Grayshift
  - Secret tool only sold to law enforcement
  - Reportedly two models available for $15,000 or $30,000 per GrayKey device
- Defense has no access, can't verify, can't test, and is limited in challenging their use



Braden Thomas (Grayshift)

Mar 19, 15:20 ADT

Thank you for your interest. Grayshift is the sole source supplier of GrayKey and is tightly controlling the sales and distribution to local, state, and federal government law enforcement end-users only. GrayKey is not available for corporate, private, or asset management use.

Probabilistic genotyping

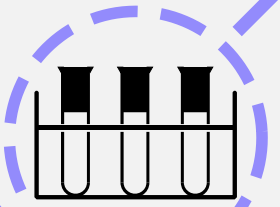
Facial recognition

Latent prints (AFIS)

Social media analytics

Ballistics and toolmarks

Breath alcohol (Alcotest)



Evidence assessment

Facial Recognition

DETECTIVE BUREAU
REAL TIME CRIME CENTER – FACIAL IDENTIFICATION SECTION

FACIAL IDENTIFICATION SECTION
SEARCH RESULT REPORT

POSSIBLE MATCH

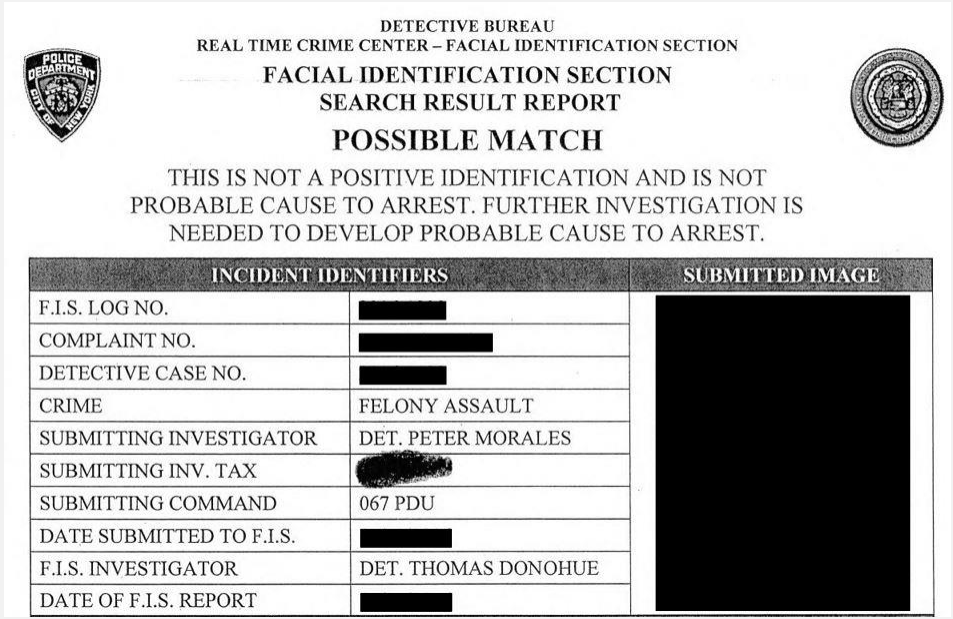
THIS IS NOT A POSITIVE IDENTIFICATION AND IS NOT PROBABLE CAUSE TO ARREST. FURTHER INVESTIGATION IS NEEDED TO DEVELOP PROBABLE CAUSE TO ARREST.

INCIDENT IDENTIFIERS		SUBMITTED IMAGE
F.I.S. LOG NO.	[redacted]	[redacted]
COMPLAINT NO.	[redacted]	
DETECTIVE CASE NO.	[redacted]	
CRIME	FELONY ASSAULT	
SUBMITTING INVESTIGATOR	DET. PETER MORALES	
SUBMITTING INV. TAX	[redacted]	
SUBMITTING COMMAND	067 PDU	
DATE SUBMITTED TO F.I.S.	[redacted]	
F.I.S. INVESTIGATOR	DET. THOMAS DONOHUE	
DATE OF F.I.S. REPORT	[redacted]	

SEARCH RESULTS/SUBJECT INFORMATION		
NAME	[redacted]	[redacted]
NYSID NO.	N/A	
D.O.B.	[redacted]	
PRECINCT OF ARREST	N/A	
DATE OF ARREST	N/A	
TOP CHARGE	N/A	
ARREST NO.	N/A	
ACTIVE WARRANT	N/A	
ACTIVE I-CARD	N/A	
RACE	BLACK / NON-HISPANIC	
SEX	MALE	
HAIR COLOR/LENGTH	N/A	
HEIGHT	N/A	
WEIGHT	N/A	
EYE COLOR	N/A	
SOCIAL MEDIA	[redacted]	
SOURCE OF IMAGE	SOCIAL MEDIA	

New York City Police Department - Facial Identification Section -One Police Plaza New York, NY Rm. 905 [redacted]
LAW ENFORCEMENT SENSITIVE

Facial Recognition



DETECTIVE BUREAU
REAL TIME CRIME CENTER – FACIAL IDENTIFICATION SECTION

FACIAL IDENTIFICATION SECTION
SEARCH RESULT REPORT

POSSIBLE MATCH

THIS IS NOT A POSITIVE IDENTIFICATION AND IS NOT PROBABLE CAUSE TO ARREST. FURTHER INVESTIGATION IS NEEDED TO DEVELOP PROBABLE CAUSE TO ARREST.

INCIDENT IDENTIFIERS		SUBMITTED IMAGE
F.I.S. LOG NO.	[redacted]	[redacted]
COMPLAINT NO.	[redacted]	
DETECTIVE CASE NO.	[redacted]	
CRIME	FELONY ASSAULT	
SUBMITTING INVESTIGATOR	DET. PETER MORALES	
SUBMITTING INV. TAX	[redacted]	
SUBMITTING COMMAND	067 PDU	
DATE SUBMITTED TO F.I.S.	[redacted]	
F.I.S. INVESTIGATOR	DET. THOMAS DONOHUE	
DATE OF F.I.S. REPORT	[redacted]	

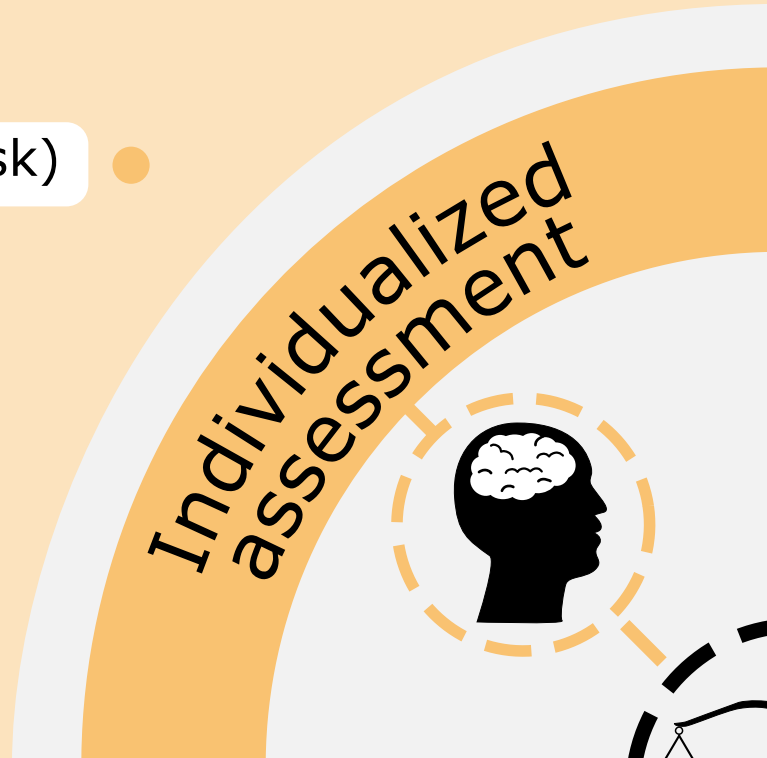
- What company?
- What algorithm?
- What qualifies as a match?
- Procedures, rules, guidelines, etc.
- Source of images?
- The Perpetual Line-Up: Unregulated Police Face Recognition in America (2016) by Georgetown Law Center on Privacy & Technology (Clare Garvie, Alvaro Bedoya, & Jonathan Frankle)
  - perpetuallineup.org

Bail determinations (flight risk)

Parole determinations (reoffense risk)

Sentencing

Parole/probation monitoring



State v. Loomis and Sentencing Algorithms

- *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016)
- Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) by Northpointe, Inc.
  - Risk Assessment Tool
- Are gender or race acceptable factors to consider?
- How are the factors weighed?
- How is that weighing determined?
- Proprietary trade secrets

¶6 The court of appeals certified the specific question of whether the use of a COMPAS risk assessment at sentencing "violates a defendant's right to due process, either because the proprietary nature of COMPAS prevents defendants from challenging the COMPAS assessment's scientific validity, or because COMPAS assessments take gender into account."[12]

Case study:
Forensic Statistical Tool (FST)

Office of the Chief Medical Examiner (OCME), NYC

Forensic Statistical Tool (FST)

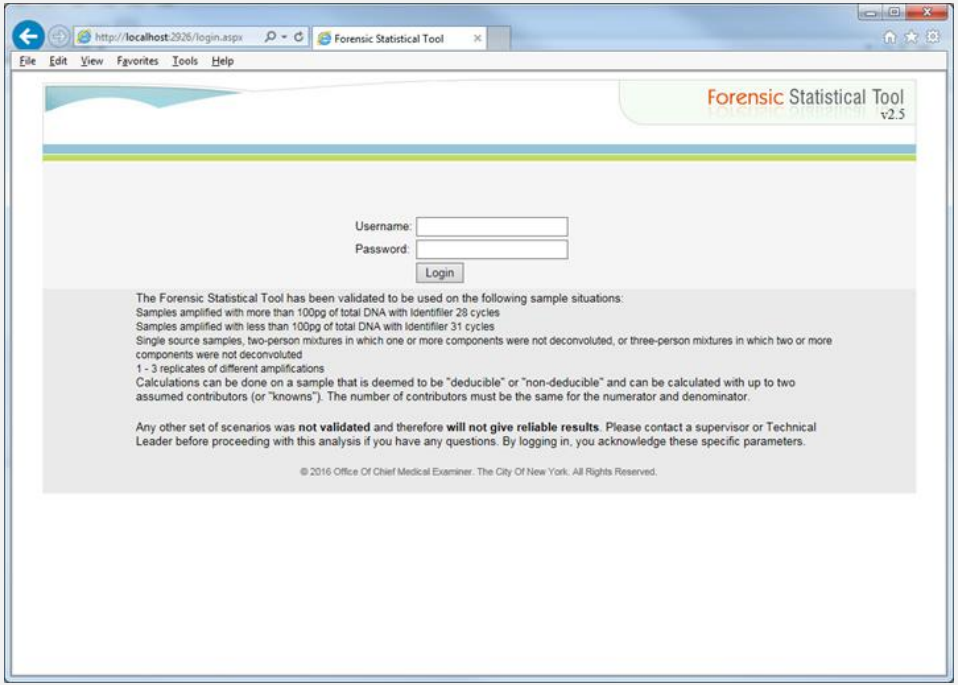
Probabilistic genotyping software

- Mixtures of DNA from 2-3 people
- Allows for dropout (missing data) and drop-in (artifactual data)
- Reports “likelihood ratio” statistic as a weight of evidence

Developed in-house

- C#, MS SQL back-end
- Browser interface for casework

Commercial sales to other labs never succeeded



FST

- 2010 Dec - Approval

NY State Commission on Forensic Science approves FST for use in casework

FST is cleared to be used to evaluate 15 genetic locations (sing. locus; pl. loci) for mixtures of up to 3 people.

FST

- 2010 Dec - Approval
- 2011 Apr - Online

OCME brings FST online for casework

FST

- 2010 Dec - Approval
- 2011 Apr - Online
- 2011 Apr - Offline

“FST went online for casework in April 2011, following its approval for use by the Commission. Shortly thereafter, also in April 2011, some functions were updated by the programmers and a small, unrelated change was inadvertently made, causing OCME to take FST off-line.”

-Florence Hutner, OCME General Counsel, October 18, 2017 letter to Brian Gestring, Director, Office of Forensic Services, NYS Division of Criminal Justice Services, “Re: Allegations by Legal Aid Society/Federal Defenders of New York to the Honorable Catherine Leahy-Scott, NYS Inspector General (September 1, 2017)”

FST

- 2010 Dec - Approval
- 2011 Apr - Online
- 2011 Apr - Offline
- 2011 Apr-Jun - Modifications

For some samples reanalyzed post-modification, likelihood ratio “values were slightly modified as expected.”

-Quality Control Test of Forensic Statistical Tool (FST) Version 2.0, June 30, 2011

“Because this modification did not affect the methodology of the program, it did not require submission to the Commission on Forensic Science or the DNA Subcommittee.”

-Affidavit of Eugene Lien, OCME Assistant Director, July 17, 2017

FST

- 2010 Dec - Approval
- 2011 Apr - Online
- 2011 Apr - Offline
- 2011 Apr-Jun - Modifications
- 2011 Jul - Online

Following performance checks, FST is reauthorized for casework.

FST

- 2010 Dec - Approval
- 2011 Apr - Online
- 2011 Apr - Offline
- 2011 Apr-Jun - Modifications
- 2011 Jul - Online
- 2016 Oct - Independent report

Source code provided under protective order in *United States v. Kevin Johnson*

Reference

Evidence

Statistical Weight

Genetic locations (loci)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Profile	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	D5S818	FGA
John Butler test (Comparison)															
	12,14	28,30	9,9	10,10	16,17	6,6	11,14	11,13	22,23	12,14	17,18	8,8	14,16	12,13	21,22
Evidence															
1	8, 12, 13, 14, 15	28, 29, 30, 30.2	8, 9, 10, 11	9, 10, 12	14, 15, 16, 17, 18	6, 8, 9, 9.3	8, 9, 14	9, 10, 11, 12, 13	19, 20, 22	12, 13, 13.2, 14, 15.2	16, 17, 18	6, 8, 9, 11	12, 13, 14, 17, 18	11, 13	19, 21, 22, 23, 24, 25, 26
2	8, 10, 12, 13, 14, 15	28, 29, 29.2, 30, 30.2, 31	8, 9, 11	10, 12	14, 15, 16, 17, 18, 19	6, 8, 9, 9.3	8, 9, 11, 12	10, 11, 12, 13	19, 20, 22, 25	12, 13, 13.2, 14, 15.2	16, 17, 18	8, 9	12, 14, 17, 18	11, 13	19, 21, 22, 23, 25
3	8, 13, 14	29, 30, 30.2, 31	8, 9	9, 10, 12	14, 15, 16, 17, 18, 19	6, 8, 9.3	8, 9, 11, 13	10, 12, 13	19, 27	12, 13, 13.2, 14, 15.2	16, 17	8, 9	12, 13, 14, 16, 17, 18	10, 11, 12, 13, 14	19, 21, 22, 23, 25, 26

Comparison Result

	Asian	Black	Caucasian	Hispanic
Likelihood Ratio	3.03e+04	339.47	70.6	419.32

Lowest is reported

Weight: The Evidence is approximately 70.6 times more probable

Hp: if the sample originated from Reference profile and two unknown,
unrelated persons

Hd: than if it originated from three unknown, unrelated persons.

2010 Validation
(non-contributor)

≠

2016 Review
Same data
(15/15 locations)

=

2016 Review
Same data
(12/15 locations)

3E	60	D	3E_3PCPenB_60pg_3p_D_1000RR	'JB	1.57E+02

157

≠

70.6

=

70.6

Profile	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	D5S818	FGA
John Butler test (Comparison)															
	12,14	28,30	9,9	10,10	16,17	6,6	11,14	11,13	22,23	12,14	17,18	8,8	14,16	12,13	21,22
Evidence															
1	8, 12, 13, 14, 15	28, 29, 30, 30.2	8, 9, 10, 11	9, 10, 12	14, 15, 16, 17, 18	6, 8, 9, 9.3	8, 9, 14	9, 10, 11, 12, 13	19, 20, 22	12, 13, 13.2, 14, 15.2	16, 17, 18	6, 8, 9, 11	12, 13, 14, 17, 18	11, 13	19, 21, 22, 23, 24, 25, 26
2	8, 10, 12, 13, 14, 15	28, 29, 29.2, 30, 30.2, 31	8, 9, 11	10, 12	14, 15, 16, 17, 18, 19	6, 8, 9, 9.3	8, 9, 11, 12	10, 11, 12	19, 20, 22, 25	12, 13, 13.2, 14, 15.2	16, 17, 18	8, 9	12, 14, 17, 18	11, 13	19, 21, 22, 23, 25
3	8, 13, 14	29, 30, 30.2, 31	8, 9	9, 10, 12	14, 15, 16, 17, 18, 19	6, 8, 9.3	8, 9, 11	10, 12, 13	19, 27	12, 13, 13.2, 14, 15.2	16, 17	8, 9	12, 13, 14, 16, 17, 18	10, 11, 12, 13, 14	19, 21, 22, 23, 25, 26

Comparison Result

	Asian	Black	Caucasian	Hispanic
Likelihood Ratio	3.03e+04	339.47	70.6	419.32

Profile	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	D5S818	FGA
John Butler test (Comparison)															
	12,14	28,30	9,9	10,10	16,17	6,6	11,14	11,13	22,23	12,14	17,18	8,8	14,16	12,13	21,22
Evidence															
1	8, 12, 13, 14, 15	28, 29, 30, 30.2	8, 9, 10, 11	9, 10, 12		6, 8, 9, 9.3			19, 20, 22	12, 13, 13.2, 14, 15.2	16, 17, 18	6, 8, 9, 11	12, 13, 14, 17, 18	11, 13	19, 21, 22, 23, 24, 25, 26
2	8, 10, 12, 13, 14, 15	28, 29, 29.2, 30, 30.2, 31	8, 9, 11	10, 12		6, 8, 9, 9.3			19, 20, 22, 25	12, 13, 13.2, 14, 15.2	16, 17, 18	8, 9	12, 14, 17, 18	11, 13	19, 21, 22, 23, 25
3	8, 13, 14	29, 30, 30.2, 31	8, 9	9, 10, 12		6, 8, 9.3			19, 27	12, 13, 13.2, 14, 15.2	16, 17	8, 9	12, 13, 14, 16, 17, 18	10, 11, 12, 13, 14	19, 21, 22, 23, 25, 26

Comparison Result

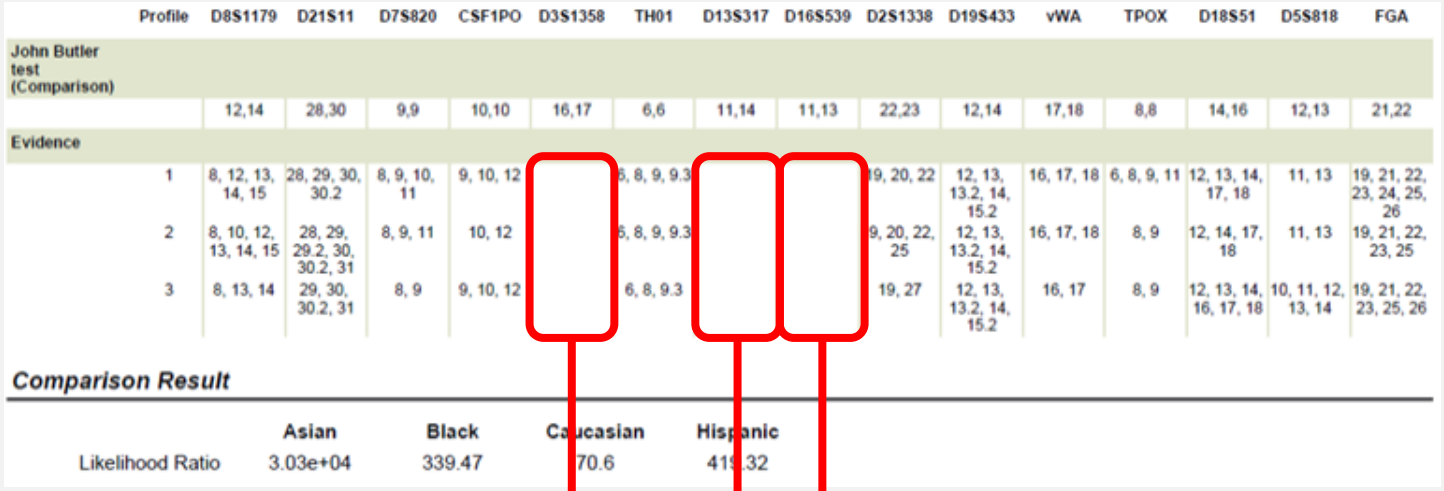
	Asian	Black	Caucasian	Hispanic
Likelihood Ratio	3.03e+04	339.47	70.6	419.32

A false positive value became less incriminating?

Why we can't tell if this is a good thing -

LR > 1 supports inclusion as a contributor

LR < 1 supports exclusion as a contributor



Profile	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	D5S818	FGA
John Butler test (Comparison)															
	12,14	28,30	9,9	10,10	16,17	6,6	11,14	11,13	22,23	12,14	17,18	8,8	14,16	12,13	21,22
Evidence															
1	8, 12, 13, 14, 15	28, 29, 30, 30.2	8, 9, 10, 11	9, 10, 12		6, 8, 9, 9.3			19, 20, 22	12, 13, 13.2, 14, 15.2	16, 17, 18	6, 8, 9, 11	12, 13, 14, 17, 18	11, 13	19, 21, 22, 23, 24, 25, 26
2	8, 10, 12, 13, 14, 15	28, 29, 29.2, 30, 30.2, 31	8, 9, 11	10, 12		6, 8, 9, 9.3			9, 20, 22, 25	12, 13, 13.2, 14, 15.2	16, 17, 18	8, 9	12, 14, 17, 18	11, 13	19, 21, 22, 23, 25
3	8, 13, 14	29, 30, 30.2, 31	8, 9	9, 10, 12		6, 8, 9.3			19, 27	12, 13, 13.2, 14, 15.2	16, 17	8, 9	12, 13, 14, 16, 17, 18	10, 11, 12, 13, 14	19, 21, 22, 23, 25, 26

Comparison Result

	Asian	Black	Caucasian	Hispanic
Likelihood Ratio	3.03e+04	339.47	70.6	41[illegible].32

Exclusionary 0.53

3.1 1.3 Inclusionary

Removing data at 3 loci that is...

FST

- 2010 Dec - Approval
- 2011 Apr - Online
- 2011 Apr - Offline
- 2011 Apr-Jun - Modifications
- 2011 Jul - Online
- 2016 Oct - Independent report
- 2017 Jan - Acknowledgement

“FST disregards the information from any locus in a sample if the alleles present at that locus reflect 97% or more of the alleles in the overall population for that locus.”

-Assistant US Attorneys, Jan. 2017

FST

- 2010 Dec - Approval
- 2011 Apr - Online
- 2011 Apr - Offline
- 2011 Apr-Jun - Modifications
- 2011 Jul - Online
- 2016 Oct - Independent report
- 2017 Jan - Acknowledgement
- 2017 Oct - Protective order vacated

ProPoublica and Yale Media Freedom and Information Access Clinic request that the protective order be vacated.

OCME does not oppose.

Order vacated, reports unsealed, and code posted by ProPublica: https://github.com/propublica/nyc-dna-software

Quality Control Test of Forensic Statistical Tool (FST) Version 2.0 - June 2011

First made public in October 2017:

“Twelve samples that were previously evaluated with FST in August 2010 were re-evaluated….

Two samples had one locus each that displayed such values [i.e. were removed].”

Only 12/439 mixtures studied in validation were re-evaluated. Only two of those exhibited data-dropping behavior (at one locus each).

In June 2018, records from 16 additional “Quality Control Test” were produced under NY’s Freedom of Information Law (FOIL).

checkFrequencyForRemoval()

~70 lines, including comments and whitespace

```
246         public void CheckFrequencyForRemoval(DataTable dtFrequencies)
247         {
248             // if our db connection isn't initialized, do it. then, get all the ethnicities (races)
249             myDb = myDb ?? new Database();
250             DataTable raceTable = myDb.getAllEthnics();
251             int intsr = 0;
252             string[] srem = new string[comparisonLoci.Count];
253 
254             // we go through all the comparison loci and check whether the sum of the frequencies for that locus is greater than 0.97.
255             // if it is, we remove the locus. frequencies are only used for the alleles in the evidence replicates.
256             for (int i = 0; i < comparisonLoci.Count; i++)
257             {
258                 bool blRemove = false;
259                 // get a CSV list of alleles for all the replicates at a locus
260                 IEnumerable<string> unknownPair = EvidenceAllelesAtLocus(evidenceAlleles[comparisonLoci[i]]);
261                 // check if the frequency is greater than 0.97 for any of the races. frequencies are values for an allele at a locus for a
```

https://github.com/propublica/nyc-dna-software/blob/master/FST.Common/Comparison.cs#L246

Unfortunately, this is not entirely surprising

Washington v. Emmanuel Fair

In a case involving evidence analyzed by the TrueAllele® system, Mr. Fair's team requested the TrueAllele® source code and development materials in 2016.

Responses included…

Washington v. Emmanuel Fair

Declaration of Dr. Mark Perlin, TrueAllele® developer

“There is no way to actually use source code in a validation study, which tests the reliability of an executable computer program.”

Washington v. Emmanuel Fair

Declaration of Dr. Michael Gorin, Professor of Medicine, UCLA

“Since it is essential that one conducts testing with a compiled and operational version of the software, there is no benefit (nor justification) in providing individuals with the source code unless they intend to modify it.”

Washington v. Emmanuel Fair

Declaration of Thomas Hebert, DNA Technical Leader for Baltimore Police

“In my opinion, I do not believe the source code is necessary for determining the reliability of TrueAllele because source code is not normally used in the validation of software programs for forensic use.”

Washington v. Emmanuel Fair

Declaration of Dr. Kevin Miller, former Lab Director of Kern Regional Crime Lab (CA)

“In fact, DNA analysts are required by national mandate to have taken only one statistics class and they have no computer science educational requirements. Therefore, this level of mathematics and engineering is above most individuals who work in the field.”

Washington v. Emmanuel Fair

Declaration of Dr. Kevin Miller, former Lab Director of Kern Regional Crime Lab (CA)

“Moreover, it strikes me has highly irregular that any one particular step in any one particular workflow would suddenly become singled out as an issue for source code revelation. If one is to discuss error in DNA testing, then would one not want to capture an error rate for the entire workflow?”

“If one is to discuss error in DNA testing, then would one not want to capture an error rate for the entire workflow?”

Why would one not?

Magic Grant

- Brown Institute Magic Grant
  - Journalism - tell new stories in new ways with technology (General Audience)
  - Technology Audience
  - Legal Audience
- Independent, third-party testing
- FST testing and FST source code review
- Comparison to other probabilistic genotyping systems



What makes independent testing hard?

- Access to executables of the software
  - Cost
  - Sometimes not even sold to individuals or groups outside law enforcement
  - Difficulty in getting old copies of software
  - Let alone source code, bug databases, testing plans, design documentation...
- Terms of service that limit publishing of results
- Trade secret protection claimed over rights of defendants
  - To shield from legitimate questions of quality and fairness more than to protect from competitors?
  - Thwarting essential iterative improvement! and accountability to stakeholders beyond buyers
- Need for natural repositories to share results/connect audiences
  - How would a defense team connect with experts? someone who found a relevant bug?

We want you to help!



Procurement Phase Wishlist

- When public money used for criminal justice software, require! or at least give credit for:
    - Source code
    - Software artifacts: bug reports, internal testing plans and results, software requirements and specifications, risk assessments, design documents, etc.
        - Lack of software standards in traditionally non-computing fields (e.g. DNA)
    - No clauses preventing third party review or publishing of defects found
    - Access to executables for third party testing
    - Scriptable interfaces to facilitate automated testing
    - Bug bounties
- Fund non-profit third party entities to do independent testing!

Be a third-party reviewer

- Criminal justice software that is open source now
  - DNA: FST and LabRetriever (US); LRmix, LikeLTD and EuroForMix (Europe)
  - Predictive policing: CivicScape
- Take a look!
  - Find bugs or bad code? Please let us know!
- Construct software yourself for alternatives and comparisons
  - Many programs have algorithms published - replicate.

Bigger picture

- Black box decision making all around us
  - Hiring, housing, how we make friends, find partners, navigate city streets, get our news, …
  - The weightier the decision the more crucial it is that we understand and can question it
- US-ACM/EU-ACM Principles for Algorithmic Transparency and Accountability
  - Awareness
  - Access and redress
  - Accountability
  - Explanation
  - Data provenance
  - Audit-ability
  - Validation and testing
- Provide the evidence needed to improve systems for all stakeholders so we don’t run our society on buggy or even malicious algorithms hidden from view

Our work wouldn’t be possible without:

- Legal Aid Society
  - DNA Unit, especially:
    - Jessica Goldthwaite
    - Clint Hughes
    - Richard Torres
  - Digital Forensics Unit, especially:
    - Lisa Brown
    - Aaron Flores
    - Shannon Lacey
    - Brandon Reim
  - Cynthia Conti-Cook
- Eli Shapiro
- Rebecca Wexler, Visiting Fellow at Yale Law School
- Federal Defenders of New York: Chris Flood, Sylvie Levine
- Clarkson University
  - Marzieh Babaeianjelodar
  - Stephen Lorenz
  - Abigail Matthews
  - Anthony Mangiacapra
  - Graham Northup
  - Mariama Njie (Iona College, McNair Scholar at Clarkson summer 2018)
  - COSI/ITL labs
- Data and Society
- Dan Krane, Wright State University
- The Brown Institute at Columbia University
  - Funding provided by a 2018-19 Magic Grant!