Berkeley Law | Samuelson Law, Technology & Public Policy Clinic

University of California, Berkeley
School of Law
353 Law Building
Berkeley, CA 94720-7200
(510) 664-4381
mgraham@clinical.law.berkeley.edu
www.law.berkeley.edu/SamuelsonClinic

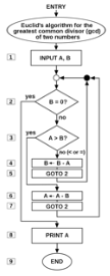
Unlocking the Black Box: Challenging the Use of Secret Algorithms & Technologies in Criminal Cases

Presented on February 17, 2021
National Association of Criminal Defense Lawyers

Megan Graham
Clinical Supervising Attorney
Samuelson Law, Technology & Public Policy Clinic
UC Berkeley School of Law

I. Black Box Algorithms

A. Algorithm:
   1. "A process or set of rules to be followed in calculations or other problem-solving operations, especially by a computer."
   2. Algorithms are mathematical by nature (e.g., Euclid's algorithm for the greater common divisor of two numbers).
   3. Today, we often think about them as computer processes, but they aren't always.
   4. For our purposes, an algorithm is the set of rules that a machine (computer) follows to achieve a particular goal.



B. Black box algorithm:
   1. A black box algorithm is one for which the inputs and outputs are known, but all or part of the set of rules to be followed in the calculation are unknown.
   2. The rules may be unknown because the owner of the algorithm considers the information proprietary and withholds it from the public and outside researchers.
   3. The rules may also be unknown because the algorithm relies on machine learning and the rules are beyond the original programmer's understanding.

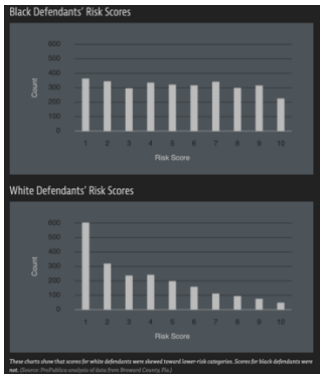
4. Point here is that—without more information—there's no way of knowing what assumptions a black box algorithm is making, or what rules it is following.

II. Where You See Black Box Algorithms

A. This list is not exhaustive and there will be changes as investigations evolve, but three examples of black box algorithms are:
   1. Facial recognition software
   2. Risk assessment tools
   3. Probabilistic genotyping software
   4. Predictive policing algorithms

B. Facial recognition software
   1. We know some information about how facial recognition software works and is trained, but at least parts of how each program runs its analysis is hidden.
   2. Depending on the law enforcement agency, the data used to train a facial recognition may or may not be known.
   3. Depending on the law enforcement agency, details about the dataset that "probe photos" are compared against may or may not be known.
   4. *See* (Mis)identified: The Challenges of Identifying and Litigating Facial Recognition Technology in Criminal Cases (July 23, 2020) (NACDL CLE), https://www.nacdl.org/Content/Webinar-The-Challenges-of-Identifying-and-Litigati

C. Risk assessment tools
   1. Risk assessment tools are algorithms that use statistical modeling to purport to predict the risk of recidivism.
   2. They are often used during bail, sentencing, and supervised release violation hearings, depending on the jurisdiction.
   3. Risk assessment scores have a number of issues, including that very few offer transparency into how various factors (e.g., demographic information, financial and geographic data, criminal history, etc.) are weighted in the analysis.



4. Risk assessment tools also generally rely on criminology studies that—because they involve studies of the prison system—over sample particular populations (i.e., young men of color) and under sample criminological behaviors of groups with less frequent interaction with the police.
5. That is, they rely on aggregated historic population data that may or may not include your client to make a determination about the individual before the court.
6. Julia Angwin et al., *Machine Bias*, ProPublica (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

D. Probabilistic Genotyping Software
   1. Probabilistic genotyping software purports to interpret DNA mixtures (i.e., 2+ contributors to a sample) that are too complex and ambiguous or low quality for human DNA analysts to resolve.
   2. Software generates a likelihood ratio ("LR"), which compares two hypotheses—the likelihood of seeing the evidence if the defendant is one of the contributors to the sample versus the likelihood of seeing the evidence if the defendant is not a contributor to the sample.
   3. The software bring together many disciplines—software engineering, forensic biology, statistical analysis—but is generally only vetted according to some of those disciplines' best practices.
   4. *See* Nathaniel Adams, *What Does Software Engineering Have to Do with DNA?*, The Champion (2018), https://www.nacdl.org/Article/May2018-WhatDoesSoftwareEngineeringHav

E. Predictive Policing Algorithms

   1. Predictive policing algorithms are programs that purport to tell law enforcement where crimes are likely to occur (often described as "hot spots").
   2. These algorithms analyze large sets of data, including historical crime data, reports on interactions between individuals and law enforcement, and other intelligence sources (e.g., social media, news reports).
   3. Law enforcement agencies that use these programs to help decide where to deploy police.
   4. For obvious reasons, a significant amount of the data analyzed can be biased. It is based on who has interactions with the criminal legal system—either as an accused individual, victim, or both—and those interactions are skewed toward certain communities and racial groups.

5. *See* Tim Lau, Brennan Ctr. for Justice, *Predictive Policing Explained* (Apr. 1, 2020), https://www.brennancenter.org/our-work/research-reports/predictive-policing-explained

III. Challenges

A. Types of Challenges
   1. Discovery request and related motions to compel
   2. Pre-trial motions to exclude evidence (e.g., Confrontation Clause, *Brady*, due process concerns)
   3. *Daubert* or *Frye* motion
   4. Requests for jury instructions or adverse inferences
   5. Challenges via objections or experts during trial
   6. Cross-examination

B. How to Tackle These Issues in Your Cases
   1. Experts
   2. Foundational AND as applied challenges
   3. Build a record
   4. Don't forget the non-tech arguments!