DEFENDING AGAINST BIG DATA

Mississippi 2024

Overview

We live in an age of data mining, artificial intelligence, and automated decision systems; a world governed by data science. As defense attorneys, we encounter big data every day in our cases without realizing it: from the case management systems we operate to the facial recognition comparison that is mentioned in discovery, from our probabilistic genotyping DNA cases to our cellphone and social media heavy ones. The prosecution and law enforcement have drunk the big data analytics Kool-Aid.

Unlike some other forms of evidence and police strategies, big data driven information has thus far defied court oversight and defense challenge. And for concerning reasons: prosecutors rarely attempt to present this data as evidence in court and police often hide its use. To defend our cases, zealously represent our clients, and preserve what's left of the First and Fourth Amendments, we have to catch up and get creative.

In this presentation, we will accomplish three objectives:

(1) we will demystify the world of machine learning—learning the vocabulary and getting comfortable with the concepts;
(2) we will identify areas of our practice where artificial intelligence solutions may have been deployed; and
(3) we will find the silver linings in the terrifying body of information coming from these tools by recognizing opportunities in the mess of information.

Outline

a. If you need an introduction or a brush up on these topics:
    i. First, come to the presentation, review the slides, and reach out to me with questions.
    ii. Then, if you still have questions or want to know more, check out these resources:
        1. Paul W. Grimm, Maura R. Grossman, Gordon V. Cormack, *Artificial Intelligence as Evidence*, 19 NW. J. TECH. & INTELL. PROP. 9 (2021).
        2. Maura R. Grossman, Hon. Paul W. Grimm (Ret.), Daniel G. Brown, and Molly (Yiming) Xu, *The GPTJudge: Justice in a Generative AI World*, 23 Duke L. & Tech. Rev. 1 (2023).
        3. David J. Malan, *What's an Algorithm*, YouTube (May 20, 2013).
    iii. If you really want to dive deep:
        1. Brian Jefferson, *Digitize and Punish*
        2. Ruha Benjamin, *Race after Technology: Abolitionist Tools for the New Jim Code*
        3. Virginia Eubanks, *Automating Inequality*

4. Cathy O'Neil, *Weapons of Math Destruction*
5. The Social Dilemma (Netflix)
6. The Great Hack (Netflix)
7. Coded Bias (Netflix)

b. Okay, so I read all that stuff and it was cool, but how does this apply to my cases? Have you seen any of these in your cases?
   i. Cellphone extraction
   ii. Social media monitoring
   iii. License plate readers
   iv. ShotSpotter
   v. Facial recognition
   vi. Probabilistic genotyping DNA analysis
   vii. Investigative genetic genealogy
   viii. Jail call recordings
   ix. Predictive policing
   x. Risk assessments, etc.

c. We'll explore some of these examples in depth together, but you're probably wondering if there's anything you can do about big data when it's screwing your client.
   i. Constitutional Challenges
      1. First Amendment
         a. Daniel J. Solove, *The First Amendment as Criminal Procedure*, 82 NYU L. Rev. 112 (2007).
         b. Alex Abdo, *Why Rely on the Fourth Amendment to do the Work of the First?* 127 Yale L. J. 444 (2017).
      2. Fourth Amendment
         a. Written Testimony of Professor Andrew Guthrie Ferguson before the House of Representatives Committee on Oversight and Reform, Hearing on Facial Recognition Technology (May 2019).
         b. Laura Hecht-Felella, *The Fourth Amendment in the Digital Age*, The Brennan Center (March 2021).
      3. Fourteenth Amendment
         a. NACDL's Task Force on Predictive Policing, *Garbage In, Gospel Out* (Sept. 2021).
   ii. Statutory Challenges
      1. Discovery Litigation
      2. Consider your eavesdropping statutes
   iii. Evidentiary Challenges
      1. Daubert/Frye
      2. Authentication
   iv. Other powers of the court
      1. Protective Orders

a. Even if the Court won't suppress collected data, make sure the Court understands the breadth of the information that has been collected by law enforcement on your client.
b. Request the Court order law enforcement to limit use of that data to the pending matter *only*, and not allow law enforcement to upload any collected data points into any law enforcement or corporate owned databases.

2. Sealing Motions
   a. After your case has come to its conclusion, be sure to make any application you can to limit the retention and re-use of data collected from/about/on your client by law enforcement.
   b. Seal and/or request destruction of as many records as you can, particularly when your case ends favorably, and you have some leverage.

d. And don't forget that we can use these same tools too:
   i. Get aggressive with your discovery and *Brady* demands. If they are going to keep all this data, then they are in possession of *Brady* material.
      1. Andrew Ferguson, *Big Data Prosecution & Brady*, 67 UCLA L. Rev.180 (2020).
      2. Barry Scheck, *The Integrity of Our Conviction: Holding Stakeholders Accountable in an Era of Criminal Justice Reform*, 48 Geo. L.J. Ann. Rev. Crim. Proc. iii (2019).
   ii. Deploy data analysis on behalf of your clients. *See* Brooklyn Defender Services' Science & Surveillance Project.
   iii. Find data analysis out there from other organizations to deploy on behalf of your clients. *See, e.g.,* www.endpolicesurveillance.com (documenting the MacArthur Justice Center and the Invisible Institute's brilliant data work on ShotSpotter's reliability in the Chicago metro area).