## TENNESSEE BAR ASSOCIATION

Home > Groups > eCommunities > Criminal Justice Section eCommunity

## Criminal Justice Section eCommunity

Home	Directory	Files	Calendar	Announcements	Photos	Videos	
Criminal Justice Section News							
All Posts	All Authors		Filter				

## Student Spotlight: Big Data in the Criminal Justice System

Posted by: Madaleine Gray & Kristen Shields on Aug 25, 2020

The TBA Criminal Justice Section Student Spotlight is a quarterly publication featuring a collaborative article authored by a law student with assistance of a section member. This Student Spotlight is authored by Belmont University Law School student <u>Madaliene Gray</u> with assistance from Criminal Justice Section Chair <u>Kristen Shields</u>.

The growing presence of cellular devices, along with ever-changing trends in technology and a globalized economy, has resulted in the innovation of devices that are amply adorned with features for users looking to minimize and multitask – a goal achieved through applications that reduce the need for items such as cameras, maps and address books. Cell phones are "such a pervasive and insistent part of daily life that carrying one is indispensable to participation in modern society."<sup>[1]</sup> On an hourly basis, users create, store and transmit data about nearly every aspect of their personal lives.<sup>[2]</sup> Today, users rely on their devices to store more personal data than ever before.<sup>[3]</sup> What does this mean for the criminals of the world?

Big Data can incriminate or exonerate. Devices are attached to the hip, meaning they are also present during the commission of crime, relied upon for driving directions, internet searches, text message or telephonic communication, and more. Hosted applications capture data that can convict the guilty or free the innocent. Do we, as a society addicted to our devices, even care about any privacy intrusion by data extraction?

Data users entrust to their cell phones more than just photographs, text messages, emails and contact information.<sup>[4]</sup> Technology giants like Facebook, AT&T and Instagram receive an extensive amount of information from a user's daily use of their device and its accompanying applications. "If you start typing something and change your mind and delete it, Facebook keeps those and analyzes them, too," Zeynep Tufekci, a prominent techno-sociologist said in a 2017 TED Talk.<sup>[5]</sup>

As cell phones and social media become an integral part of society, companies know more about consumers than ever before. In 2019, Facebook had 2.4I billion monthly active users, with 2.7 billion people using at least one of the company's core products, such as Facebook, WhatsApp, Instagram or Messenger, each month.<sup>[6]</sup> The daily data collected from 2.7 billion Facebook users helps supplement extensive data sets that, coupled with their subsequent analysis, are referred to as "Big Data."<sup>[2]</sup>

Big Data enables businesses and sociologists to better understand and predict consumers' habits and tendencies, ranging from their dating preferences to shopping habits.<sup>[8]</sup> Big Data specialists express a fundamental belief that scientifically-derived evidence is the most powerful instrument society has "to design enlightened policy and produce a positive social transformation."<sup>[9]</sup>

The number of applications for Big Data analytics is ever-growing: matching partners on online dating sites;<sup>[10]</sup> finding correlations between air quality and health conditions; or using genomic analysis to speed up the breeding of crops like rice for drought resistance.<sup>[11]</sup> In marketing, use of Big Data includes "recommendation engines" like those used by companies such as Netflix and Amazon to make purchase or viewing suggestions based on the prior interests of one customer when compared to millions of other customers' data.<sup>[12]</sup> Interestingly, the superstore, Target, infamously used an algorithm to detect a woman's pregnancy by tracking purchases of items, such as unscented lotions, and used the information gathered to offer special discounts and coupons to the new and easy-to-detect patron.<sup>[13]</sup> Big Data is even capable of sorting through the billion social media posts made daily.<sup>[14]</sup>

Big Data has also become a powerful force in the criminal justice system, driving investigations, solving crimes, and resulting in convictions. Big Data is not only a reactive tool – it has proactive utilities proving to be equally, if not more, valuable. While presently used less in the criminal justice realm as compared to that of the consumer, Big Data has the potential for broad application with crime prevention and in incriminating or exonerating the accused.

Localized application remains largely unknown to the masses, but in due time, its efficacy will be visible. In Los Angeles, California, police use computerized "predictive policing" to anticipate criminal activity and allocate officers and other resources accordingly.<sup>[15]</sup> In Fort Lauderdale,

Florida, algorithms are used to set bail amounts.<sup>[16]</sup> In states across the country, data-driven estimates related to recidivism are being used to set jail sentences.<sup>[17]</sup> Algorithms have even been applied to determine case outcomes.

The predictive function of Big Data transcends beyond the criminal justice realm. Kevin Quinn, former assistant professor of government at Harvard University, hosted a competition comparing his statistical model's performance to the qualitative judgments of 87 law professors to determine which could best predict the outcome of all the Supreme Court cases in a year.<sup>[18]</sup> Despite the law professors' knowledge of each justices' interpretation style and past opinions, the statistical model was more accurate in predicting the Supreme Court case outcomes than the 87 law professors.<sup>[19]</sup>

Nonetheless, within the four corners of the law, Big Data has the potential for greatest impact in the criminal sector. The systems in place for arresting, detaining, trying, imprisoning and releasing criminal defendants are ideal for bulk data analytics.<sup>[20]</sup> Millions of individuals are cycled through the criminal justice system with common data points: charges for each defendant; whether and when counsel was appointed; whether an individual was released on bail or held awaiting trial; the length of pretrial imprisonment; final disposition; and the terms of any sentence imposed.<sup>[21]</sup> The empirical analysis of bulk criminal justice data has produced promising results in the limited contexts where currently employed; however, broader applications are hampered by a lack of access to necessary information.<sup>[22]</sup>

Currently, criminal justice researchers have to rely on limited survey data or engage in the costly and cumbersome task of collecting and coding more complete data sets from a patchwork of federal, state and local systems.<sup>[23]</sup> A large portion of this type of data is already collected and recorded by courts, but is not electronically available.<sup>[24]</sup> Standardizing the use of data will supplement the current lack of information and lead to more transparency on issues within the criminal justice system.<sup>[25]</sup>

Skeptics speculate that the algorithms used to predict future crimes may harbor racial biases. <sup>[26]</sup> Critics argue that the data that algorithms rely on is collected by a criminal justice system in which race makes a difference in the probability of arrest.<sup>[27]</sup> Specifically, concerns center around potential bias in analytics if inputs are derived from a biased system, only to result in skewed data against African American and Latino defendants. The societal danger lies in these groups potentially appearing as a bigger risk than white defendants, should predictions from algorithms rely on data tainted by disparity and exacerbate, rather than eliminate, racial bias in the criminal justice system.<sup>[28]</sup>

Other concerns stem from unfamiliarity with Big Data, as its utility is relatively new with minimal information known about systems' functions.<sup>[29]</sup> However, the existing legal standard for scientific critique is applicable to Big Data analytics.<sup>[30]</sup> In *Daubert v. Merrell Dow Pharmaceuticals* <sup>[31]</sup>, the Supreme Court handed down a landmark ruling concerning the standard for admitting expert scientific testimony in a federal trial.<sup>[32]</sup> In a broader sense, *Daubert* reinforced the idea that scientific evidence should be "not only relevant, but reliable."<sup>[33]</sup> The factors outlined in *Daubert* could be utilized to guide review of algorithms; these factors include: whether the

method producing the evidence, or data in this case, "can be (and has been) tested;"<sup>[34]</sup> whether it has been subjected to peer review and publication;<sup>[35]</sup> whether it has known or potential rate of error; and if the methodology is generally accepted in the relevant scientific community.<sup>[36]</sup>

The extent and nature of challenges to the use of Big Data in criminal justice remain to be seen. Is it more invasive for Big Data to be used to detect a pregnant Target customer or predict or solve crime? Big Data brings along big questions, and your cell phone bill now comes with a hidden cost, your personal information. In an age where we are no longer limited by fixed data allowances, the issue becomes more about weighing data collection against convenience. When the price is privacy, will you make the call?

[1] Carpenter v. United States, I38 S. Ct. 2206, 2220 (2018) (quoting Riley v. California, I34 S. Ct. 2473, 2484 (2014)).

[2] E.g., Kathryn Nobuko Horwath, A Check-In on Privacy After United States v. Jones: Current Fourth Amendment Jurisprudence in the Context of Location-Based Applications and Services, 40 Hastings Const. L.Q. 925, 925-26 (2013) (Describing just a portion of the information a typical person distributes via Twitter, Facebook, etc. throughout a typical day).

[<u>3</u>] Richard M. Thompson II, CRS report (prepared for members and committees of Congress) The Fourth Amendment Third Party Doctrine, CONGRESSIONAL RESEARCH SERVICE (June 5, 2014), <u>https://fas.org/sgp/crs/misc/R43586.pdf</u>; Jonathan Shaw, Why "Big Data" is a Big Deal, Harvard Magazine, <u>http://harvardmag.com/pdf/2014/03-pdfs/0314-30.pdf</u> March-April 2014 (featuring the thoughts of Gary King, director of Harvard's Institute for Quantitative Social Science (IQSS), a hub of expertise for interdisciplinary projects aimed at solving problems in human society.").

[<u>4</u>] Richard M. Thompson II, CRS report (prepared for members and committees of Congress) The Fourth Amendment Third Party Doctrine, CONGRESSIONAL RESEARCH SERVICE (June 5, 2014), <u>https://fas.org/sgp/crs/misc/R43586.pdf</u>.

[5] See <u>https://www.chicagotribune.com/business/ct-facebook-privacy-policy-20l80325-story.html;</u>See

also https://www.ted.com/talks/zeynep\_tufekci\_we\_re\_building\_a\_dystopia\_just\_to\_make\_people\_click\_on\_ad

[6] Active users are those which have logged in to Facebook during the last 30 days. <u>https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/</u>.

[7] Dacia Green, Article, Big Brother is Listening to You: Digital Eavesdropping in the Advertising Industry, I6 Duke L. & Tech. Rev. 352 (2018).

[8] Dacia Green, Article, Big Brother is Listening to You: Digital Eavesdropping in the Advertising Industry, I6 Duke L. & Tech. Rev. 352 (2018). [9]Jonathan Shaw, Why "Big Data" is a Big Deal, Harvard

Magazine, <u>http://harvardmag.com/pdf/2014/03-pdfs/0314-30.pdf</u> March-April 2014 (featuring the thoughts of Gary King, director of Harvard's Institute for Quantitative Social Science, a hub of expertise for interdisciplinary projects aimed at solving problems in human society.").

[10] Lee Rainie & Janna Anderson, Code-Dependent: Pros and Cons of the Algorithm Age, Pew Res. Ctr. (Feb. 8, 2017), <u>http://www.pewinternet.org/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age/</u>.

[<u>II]</u> Jonathan Shaw, Why "Big Data" is a Big Deal, Harvard Magazine, <u>http://harvardmag.com/pdf/2014/03-pdfs/0314-30.pdf</u> (2014)

[12] See generally ARTICLE: BLEEDING DATA IN A POOL OF SHARKS: THE ANATHEMA OF PRIVACY IN A WORLD OF DIGITAL SHARING AND ELECTRONIC DISCOVERY, 64 S.C. L. Rev. 717, 726 (2013).

[13] See generally ARTICLE: BLEEDING DATA IN A POOL OF SHARKS: THE ANATHEMA OF PRIVACY IN A WORLD OF DIGITAL SHARING AND ELECTRONIC DISCOVERY, 64 S.C. L. Rev. 717, 726 (2013).

[14] Jonathan Shaw, Why "Big Data" is a Big Deal, Harvard Magazine, <u>http://harvardmag.com/pdf/2014/03-pdfs/0314-30.pdf</u> (2014)

[15] See Hal Eisner, Predictive Policing, FOX NEWS LOS ANGELES (Nov. 9, 2015, 10:38 PM), <u>http://www.foxla.com/news/local-news/predictive-policing</u>.

[<u>16]</u> See Julie Angwin, Jeff Larson, Surya Mattu & Lauren Kitchner, Machine Bias, PROPUBLICA (May 23, 2016). <u>https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-</u> <u>sentencing</u>.

[17] See Anna Maria Barry-Jester, Ben Casselman & Dana Goldstein, Should Prison Sentences Be Based on Crimes That Haven't Been Committed Yet?, FIVETHIRTYEIGHT (Aug. 4, 2015), <u>https://fivethirtyeight.com/features/prison-reform-risk-assessment/</u>.

[18] Jonathan Shaw, Why "Big Data" is a Big Deal, Harvard

Magazine, <u>http://harvardmag.com/pdf/2014/03-pdfs/0314-30.pdf</u> (Apr. 2014) (featuring the thoughts of Gary King, director of Harvard's Institute for Quantitative Social Science (IQSS), a hub of expertise for interdisciplinary projects aimed at solving problems in human society.").

[<u>19]</u> Id.

[20] Samuel R. Wiseman, ARTICLE: The Criminal Justice Black Box, 78 Ohio St. L.J. 349, 352 (2017).

[21] Samuel R. Wiseman, ARTICLE: The Criminal Justice Black Box, 78 Ohio St. L.J. 349, 352 (2017).

[22] For more, see generally ARTICLE: CUMULATIVE RESEARCH KNOWLEDGE AND SOCIAL POLICY FORMULATION: The Critical Role of Meta-Analysis, 2 Psych. Pub. Pol. and L. 324.

[23] See, e.g., Lauren Sudeall Lucas, Reclaiming Equality to Reframe Indigent Defense Reform, <u>97</u> <u>MINN. L. REV. II97, II98 (2013)</u> (documenting the continuing, widespread failure to provide counsel to indigent defendants and relying primarily on studies that conducted surveys and collected data to provide average estimates of expenditures on indigent defense but did not document the number of defendants who qualify for but do not receive indigent defense); Samuel R. Wiseman, Pretrial Detention and the Right to Be Monitored, <u>123 YALE L.J. 1344, 1352-54 (2014)</u> (describing the frequent jailing of nondangerous defendants pretrial using the available statistics, which are primarily produced from Bureau of Justice Statistics surveys rather than bulk, raw data and systemic analysis of those data); Ann K. Wagner, Comment, The Conflict over Bearden v. Georgia in State Courts: Plea Bargained Probation Terms and the Specter of Debtors' Prison, <u>2010 U. CHI. LEGAL F.</u> <u>383, 384-85</u> (describing the problem of debtors' prison). For the studies, see, for example, NAT'L RIGHT TO COUNSEL COMM., THE CONSTITUTION PROJECT, JUSTICE DENIED: AMERICA'S CONTINUING NEGLECT OF OUR CONSTITUTIONAL RIGHT TO COUNSEL 85-89 (Apr.

2009), <u>http://www.constitutionproject.org/wp-</u>

<u>content/uploads/2012/10/139.pdf</u> [<u>https://perma.cc/M945-2GKC</u>], which provides examples of indigent defendants for whom a court failed to appoint counsel or appointed counsel too late but does not indicate the percentage or number of defendants nationwide for whom counsel was denied or provided too late.

[24] Samuel R. Wiseman, ARTICLE: The Criminal Justice Black Box, 78 Ohio St. L.J. 349, 352 (2017).

[25] Daniel C. Esty, Environmental Protection in the Information Age, 79 N.Y.U. L. REV. 115, 156, 162 (2004).

[26] See Julie Angwin, Jeff Larson, Surya Mattu & Lauren Kitchner, Machine Bias, PROPUBLICA (May 23, 2016). <u>https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing</u>.

[<u>27]</u> ld.

[<u>28]</u> ld.

[29] See Laurel Eckhouse, Big data may be reinforcing racial bias in the criminal justice system, THE WASHINGTON POST (Feb. 10, 2017), <u>https://www.washingtonpost.com/opinions/big-data-may-be-reinforcing-racial-bias-in-the-criminal-justice-system/2017/02/10/d63de518-ee3a-Ile6-9973-c5efb7ccfb0d\_story.html?noredirect=on&utm\_term=.Id5ead53c286.</u>

[30] ARTICLE: Small Data Surveillance v. Big Data Cybersurveillance, 42 Pepp. L. Rev. 773, 806 (2015).

[31] See Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993).

[32] Specifically, the Daubert Court agreed that Federal Rule of Evidence 70I provided the correct standard for scientific testimony admissible in trial. Id. at 587; see also U.S.C.S. Fed. Rules Evid. R. 702.

[33] See Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579, 589 (1993).

[<u>34</u>] Daubert, 509 U.S. at 593.

[<u>35]</u> Id.

[<u>36]</u> ld. at 594.

<u>Madaleine Gray</u> is a recent J.D. graduate of Belmont University College of Law. Originally from Sydney, Australia, she now resides in Nashville. During law school, Gray interned for the Public Defender's Office, small firms, the District Attorney's Office, the State of Tennessee Attorney General's Office and the Hon. Judge Angelita Blackshear Dalton. Gray was also a member and managing editor of the *Belmont Criminal Law Journal*, where two of her publications can be found. Currently, she is employed as a judicial law clerk for the Hon. Judge Barry R. Tidwell, studying for the October Bar Exam, and pursuing a master's degree in Applied Intelligence at Georgetown University.

© Copyright 1996 - 2020 Tennessee Bar Association. All rights reserved.