



1-29-2021

Digitizing and Disclosing Personal Data: The Proliferation of State Criminal Records on the Internet

Sarah Lageson

The State University of New Jersey

Elizabeth H. Webster

Loyola University Chicago, ewebster1@luc.edu

Juan Sandoval

University of California, Irvine

Follow this and additional works at: https://ecommons.luc.edu/criminaljustice_facpubs



Part of the [Criminology and Criminal Justice Commons](#)

Author Manuscript

This is a pre-publication author manuscript of the final, published article.

Recommended Citation

Lageson, Sarah; Webster, Elizabeth H.; and Sandoval, Juan. Digitizing and Disclosing Personal Data: The Proliferation of State Criminal Records on the Internet. *Law and Social Inquiry*, , : 1 - 31, 2021. Retrieved from Loyola eCommons, Criminal Justice & Criminology: Faculty Publications & Other Works, <http://dx.doi.org/10.1017/lsi.2020.37>

This Article is brought to you for free and open access by the Faculty Publications and Other Works by Department at Loyola eCommons. It has been accepted for inclusion in Criminal Justice & Criminology: Faculty Publications & Other Works by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.



This work is licensed under a [Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 License](#).
© The Authors, 2021.

**Digitizing and Disclosing Personal Data: The Proliferation of State Criminal Records on
the Internet**

Forthcoming in *Law & Social Inquiry*

Sarah Esther Lageson

Assistant Professor, School of Criminal Justice, Rutgers University–Newark, United States

sarah.lageson@rutgers.edu

Elizabeth Webster

Assistant Professor, Department of Criminal Justice and Criminology, Loyola University,
Chicago, United States

ewebster1@luc.edu

Juan R. Sandoval

Doctoral Student, Department of Criminology, Law and Society, School of Social Ecology,
University of California, Irvine, United States

jrsando1@uci.edu

ABSTRACT

Digitization and the release of public records on the Internet have expanded the reach and uses of criminal record data in the United States. This study analyzes the types and volume of personally identifiable data released on the Internet via two hundred public governmental websites for law enforcement, criminal courts, corrections, and criminal record repositories in each state. We find that public disclosures often include information valuable to the personal data economy, including the full name, birthdate, home address, and physical characteristics of arrestees, detainees, and defendants. Using administrative data, we also estimate the volume of data disclosed online. Our findings highlight the mass dissemination of pre-conviction data: every year, over ten million arrests, 4.5 million mug shots, and 14.7 million criminal court proceedings are digitally released at no cost. Post-conviction, approximately 6.5 million current and former prisoners and 12.5 million people with a felony conviction have a record on the Internet. While justified through public records laws, such broad disclosures reveal an imbalance between the “transparency” of data releases that facilitate monitoring of state action and those that facilitate monitoring individual people. The results show how the criminal legal system increasingly distributes Internet privacy violations and community surveillance as part of contemporary punishment.

INTRODUCTION

With technological developments, American criminal records are more available than ever before. Digital criminal records are routinely posted online by governmental agencies, sold in bulk through public-private contracts to data brokers, and reposted by a variety of private websites (National Center for State Courts n.d.; Lageson 2020). Once disclosed, criminal records are used in a variety of enterprises, including the growing personal data and commercial background checking industry (Kaspero and Canhasi 2018), and across a suite of websites that offer “people search” services, community notifications, and public data portals (Corda and Lageson 2019). Through these public and private sector channels, the broad dissemination of criminal record data over the Internet has become ubiquitous, not to mention profitable, in a relatively short amount of time.

Survey estimates show that approximately thirty-eight million people (15 percent of Americans) search online for conviction records each year (Lageson, Denver, and Pickett 2019). In response to such consumer demand, personal data companies build marketing platforms on generous data assessments. For instance, a US Department of Justice study reports that state repositories maintain criminal records for approximately 110 million people, constituting roughly 30 percent of the US population (Goggins and DeBacco 2018). Yet Themis Data Solutions claims that its repository holds six hundred million unique criminal records, covering 90 percent of the US population. BackgroundChecks.com maintains 650 million criminal records in their database. The Data Diver Criminal Index claims to contain nearly 500 million individuals, with over two billion criminal records ranging across 1,400+ jurisdictions.¹ These corporate claims point toward the mass collection and online dissemination of public record information (including incorrect, outdated, and duplicate information), raising important law and society questions about the economic value of records that are disclosed as a public good and then commodified into valuable data.

But where and how do private companies and the greater public obtain such voluminous personal information? This study traces the origins of digital criminal records by empirically

¹ Themis Data Solutions, “Products.” <https://www.themisds.com/Home/Products>; Backgroundchecks.com. “National Criminal Databases frequently asked questions.” <https://www.backgroundchecks.com/learningcenter/nationalcriminaldatabases>; Data Diver Technologies. “About us.” <https://www.edatadivers.com/about.cfm>.

analyzing four types of criminal record disclosures in each state. Our data consists of two hundred governmental websites that affirmatively post police, criminal court, correctional agency, and state criminal record repository data for public access. We seek to understand the criminal record disclosure practices of state agencies, identify the types of personal data contained in these increasingly valuable datasets, and estimate the volume of personally identifiable data released on the Internet each year at no cost to users.

Our analysis shows that, while the underlying legal and social reasoning for public disclosure in the name of transparency is a laudable state goal, current record disclosure practices contain voluminous amounts of personally identifiable data, including birth dates, home addresses, physical descriptions, and photographs of people who have been arrested, charged, or detained. These types of personally identifiable information (PII) released by local government are increasingly disconnected from a criminal justice purpose of public notification or agency watchdogging.

This analysis of public records laws “in action” (Pound 1910) shows how government transparency efforts can contribute to community and corporate surveillance through the widespread release of personal data on the Internet. While there are many potential benefits to broad public access to criminal record data, potential consequences include the spread of criminal punishment into civil and corporate sectors, increased social isolation and the stigmatization of criminal record holders, and an inflated sense of the risk of victimization among the general public. These trends may also contribute to a growing “reverse sunshine effect,” where citizens and companies increasingly monitor other members of the public rather than the public agencies that transparency laws were intended to target (Brobst 2015, 191). These digital disclosures also mean that criminal punishment now includes the deprivation of privacy as the justice system distributes personal information across the Internet.

BACKGROUND

Legalities of Criminal Record Disclosure on the Internet

In the United States, criminal records are considered public material. This accessibility reflects two competing interests of the public’s right to know and the accused or convicted person’s right

to privacy. Distinct criminal justice agencies also have individualized purposes for releasing data. For instance, the police blotter has long been publicly available as a log of law enforcement activity. Local jail inmate rosters of daily bookings and current state prison inmate rosters have also been widely available for public inspection and may include the names and photographs of incarcerated people. There is a common law right to “access court records to inspect and to copy,” which is codified through state constitutions or the legislation governing criminal court operations.² These records also contain detailed information about defendants, such as their bail amounts, home addresses, or dates of birth (Conley et al. 2011).

Subsequent disclosure of criminal information in the private sector is legal, provided that the government agency makes the information available first. The First Amendment allows for the reproduction and further publication of these public records.³ Tort law cannot impose damages for truthfully publishing this information⁴ or for failing to remove outdated information, even when an arrest record has been expunged.⁵ Historically, when records of police and court activity existed only on paper, accessibility was limited. Technological advances and the digitization of records has greatly increased public access to criminal record information, diminishing the “practical obscurity” of paper records for the “ubiquity of an electronic record” (Kinstler 2013). The 1996 Electronic Freedom of Information Act encouraged government agencies to “use new technology to enhance public access to agency records and information.”⁶ This legislation was followed by the 2002 E-Government Act, which makes mandatory online access to federal court records and has inspired state courts across the country to follow suit (Jacobs and Crepet 2007).⁷ Yet, while the digitization, duplication, and online indexing of these records has fundamentally changed the scope of public records, public policy has remained rooted in the age of practical obscurity and paper files (Corda 2016). This begs new questions about the benefits and risks of criminal record disclosure on the Internet.

Benefits of Disclosure

² See *Nixon v. Warner Communications*, 435 US 589 (1978).

³ See, for instance, *Cox Broadcasting Corp v. Cohn*, 420 US 469 (1975).

⁴ See, for instance, *Florida Star v. B. J.F.*, 491 US 524 (1989).

⁵ *Martin v. Hearst Corporation*, 777 F.3d 546 (2nd Cir. 2015).

⁶ Electronic Freedom of Information Act Amendments, 110 Stat 3048 (2015).

⁷ E-GOVERNMENT ACT OF 2002, PL 107-347, December 17, 2002, 116 Stat 2899.

Supreme Court Justice Louis Brandeis (1914, 92) argued that “sunlight is said to be the best of disinfectants; electric light the most efficient policeman.” Disclosing records to the general public is often rooted in arguments for government transparency, public notification, agency accountability, and public safety. Like other technologically mediated forms of access, such as police body-worn cameras, granting access to digital criminal legal data allows the public to monitor local police and court operations (Fan 2018). Technological advances in data science and web platforms have allowed state agencies to begin to expand data disclosures to focus increasingly on individuals who have been arrested, detained, or charged with a crime, expanding the original intent of sunshine laws (Brobst 2015). Criminal justice agencies now routinely release arrestee and incarceration rosters on the Internet as a public service, such as providing family members with updates on bail amounts and court appearances. Supporters of these information expansions note how digital record keeping has facilitated public and personal safety goals, such as in the development of services like VINELink, a web- and text-based program that notifies crime victims when a perpetrator is released from incarceration, filling a much-needed gap in victim’s services (APPRISS Safety 2020).

Granting access to criminal records as a public safety measure also has social value and political salience. For instance, legislation proposed in Indiana in 2019 aimed to create an online registry that would list all state residents who have received a felony conviction since 2012. The registry would “fill the gaps” in public access to criminal records since not all counties were submitting records to a centralized state criminal court database (Carden 2019). The Florida legislature passed a set of bills in 2019 to create a “Soliciting for Prostitution Public Database,” which would filter various criminal dockets from around the state into a separate unified database of people charged with sex work-related crimes.⁸

Some criminal justice agencies and websites argue that the public disclosure of criminal records serves as a deterrent and contributes to community values and the maintenance of social norms. For instance, police department representatives have told media outlets that posting arrest information to social media “may in fact be deterring some from committing crimes” (Nashrulla and Grygiel 2019). A website that offers mug shot removal services argues that

⁸ See Florida Senate Bill 540 and House Bill 851, <https://www.flsenate.gov/Session/Bill/2019/00540/>.

<EXT>[p]ublishing official records and mugshots can be a deterrent to committing crimes—if not entirely based on the person’s change of heart about breaking the law, at the very least because of the potential embarrassment. Many publications and websites dedicated to publishing this type of information hope that you will see the people posted and not want to get arrested and earn eternal notoriety by joining their ranks. (*Unpublisharrest.com*, n.d.) <EXT>

Theoretical justifications for criminal record disclosure have been debated in the criminological literature, such as in the rise of sex offender registries that aim to serve both informational and punishment functions at once (Logan 2009; Burchfield 2011). In theory, criminal data disclosure might serve retributive and deterrent purposes, though the specific deterrent value of such disclosure has been questioned in light of the negative stigma of records (Pager 2007; Bushway and Apel 2012; Uggen et al. 2014) and the potentially criminogenic effects of labeling (Chiricos et al. 2007; Restivo and Lanier 2015).

Public opinion is mixed. Surveys show public support for disclosing certain types of criminal justice records when there is a perceived rationale of public benefit or safety; nine out of ten respondents in a 2000 survey supported criminal record disclosure for employment or volunteering purposes (US Department of Justice 2001). However, the same 2001 survey showed that 90 percent of respondents expressed concerns about the misuse of personal information in disclosed criminal records, and 90 percent of respondents preferred that state agencies do not use the Internet to post public criminal history information. In a survey conducted nearly two decades later, the use of criminal records by private companies received similarly low support: 88 percent of respondents opposed the publication of arrest records by private companies (Lageson, Denver, and Pickett 2019). This research suggests a growing public interest in privacy for criminal records, even as transparency efforts have evolved from a focus on state practices to a focus on the individuals who are processed through the criminal legal system.

Risks of Disclosure

While the stigma associated with a criminal record is well established, the digital context raises several new issues. As publicly disclosed criminal records are duplicated across websites and

databases, they are indexed into Internet search results, wherein a Google search for a person's name may yield dozens of websites that repost criminal record data. Mug shot galleries contain millions of booking photographs, searchable by name and city (Stelloh 2017; Solon 2018), leaving digital record subjects to outrun this "digital degradation" (Lageson and Maruna 2018), which might impact their ability to find jobs and housing, stay connected to their loved ones, and engage in pro-social behaviors in their community (Kurlychek, Brame, and Bushway 2006; Thacher 2008; Western 2008; Bushway and Apel 2012). These disclosures also complicate efforts to appeal, exonerate, expunge, or seal eligible criminal records (Yates 2012).

In addition, digital stigma challenges individual due process rights, negatively impacting those who are swept into databases and marked by digital suspicion through "big data blacklisting," surveillance, and monitoring (Hu 2016:1735). The loss to one's digital reputation and its attendant consequences—difficulty finding housing, employment, even a romantic partner—produce a new kind of "digital punishment" that may or may not reflect the state's legitimate punishment goals (Lageson 2020). Racial minorities who have a criminal record may also be more likely to be marked, marginalized, and disenfranchised by criminal record stigma (Pager 2007; Stoll and Bushway 2008). Further, poor people of color may be less likely to have the resources to curate and manage their digital reputation online (Marwick and boyd 2018).

Widespread disclosure of records also contributes to the spread of inaccurate or outdated information. When criminal justice agencies provide agency-specific information online, records often fail to reflect updates to a criminal case, such as a dismissal or an expungement. According to a Bureau of Justice Statistics study, only thirty-one states can say that 60 percent of their records include final disposition information (US Department of Justice 2014). In addition to being outdated, online criminal record data can also be inaccurate—a fact that some state and local criminal justice agencies acknowledge in the form of a website disclaimer. For example, the Harris County, Texas, Sheriff's Office website reads: "While a good-faith effort is made to post accurate information, neither Harris County nor the Sheriff of Harris County makes any representations whatsoever as to the quality, content, accuracy or completeness of any information being posted to this website and contained herein, nor does it guarantee that any such information is current or correct." Such a "blasé acceptance of data error and its negative consequences for individuals" becomes exacerbated online where erroneous information is far easier to reproduce than it is to correct (Logan and Ferguson 2016, 543).

These negative consequences for record holders reflect an expansion of the “culture of control” into virtual spaces (Garland 2001). Fueled by “new levels of fear and insecurity,” the culture of control juxtaposes community safety concerns against offender rights, definitively prioritizing the former (182). As Garland writes, “[t]he assumption today is that there is no such thing as an ‘ex-offender’—only offenders who have been caught before and will strike again. ‘Criminal’ individuals have few privacy rights that could ever trump the public’s uninterrupted right to know” (181). Considering the persistence of digital criminal record data that can be discovered even after sealing, expungement, or exoneration, the difficulty of achieving ex-offender status becomes even more salient. Here, digital stigma represents a new “technology of exile,” in which the cultural project of socially isolating potentially threatening persons is sustained largely by private citizens (for example, the potential employers, landlords, and neighbors of criminal record holders) (Simon 2007, 172).

The accessibility and proliferation of online criminal record data may also foster an exaggerated sense of victimization risk. A private criminal records website posits that,

<EXT>[a]s we continue to bring you publicly available crime incident, sex offender registry and arrest information, we hope to begin changing attitudes about crime information and personal safety. It’s our mission at Busted! to help make crime awareness part of your everyday life in keeping yourself and family safe—just the same as locking your doors, setting your alarms, protecting your identity, hugging your kids, etc. Crimes are occurring around you and your loved ones every second of every day. What are you doing to stay aware, informed, and safe?

(*Bustedmugshots.com* 2020) <EXT>

Such rhetoric transforms the public’s right to know into an urgent need to know, reflecting and producing “new levels of fear and insecurity” (Garland 2001, 182). Moreover, it deflects focus from governmental failings and criminal legal system inefficiencies to the overriding need for governmental protection.

PII, Big Data Commodification, and Surveillance

The disclosure of public criminal records on the Internet has also contributed to the personal data industry, where criminal record data has become a valuable data commodity, leading to a new type of penal entrepreneurialism (Feeley 2002; Corda and Lageson 2019). Stemming from technological innovations and a suite of software platforms that provide online database functionality, governmental agencies have increasingly released digital records on their websites. These records can then be copied, “crawled,” “scraped,” and duplicated by third parties. In particular, the inclusion of personal identifiers (such as names, birthdates, and addresses) on criminal justice agency websites has made such data increasingly attractive to data aggregators, which match public criminal records to other sources of public and consumer data (Kaspero and Canhasi 2018). The commercialization of these public records into background checks and personal data industries has created a type of “Pandora’s Box” scenario for criminal record subjects who now face various versions of their criminal history information kept in a myriad of private databases and duplicated across websites (Lee 2018).

Furthermore, the inclusion of personally identifiable information, such as an arrestee’s home address, widely expands foundational notions of government transparency and foundational arguments for access to criminal histories by employers and concerned members of the public. Put differently, the “right to know” about a person’s criminal legal entanglements typically would include the type and date of offense, but it has now been expanded to include their height, weight, birthdate, and home address. The commodification of personal data further complicates the original intent of transparency laws. For instance, a 2017 analysis of 229,000 Freedom of Information Act requests to eighty-five government agencies found that only 7.6 percent of requests were made by journalists and the news media, while 39 percent were made by commercial businesses (Galka 2017).⁹ This increasing private control over public information is a theme endemic to online criminal records, leading to what legal scholar Jennifer Brobst (2015, 198) calls a “reverse sunshine effect,” where in the digital age, the push for open access to public records has become a “double-edged sword” as “public records primarily contain information on private individuals rather than information on government officials.”

Privacy law scholars have long questioned the surveillance and liberty consequences of personal data aggregation. Jeffrey Rosen (2000, 9) writes of the “unwanted gaze” of data-driven surveillance, leaving record subjects “vulnerable to being misjudged on the basis of our most

⁹ Freedom of Information Act, 5 USC §552 et seq.

embarrassing, and therefore most memorable, tastes and preferences.” Daniel Solove (2001, 1140) argues that if government can share personal data freely with private companies that haphazardly combine various information sources, the result is a “growing dehumanization, powerlessness, and vulnerability for individuals.” The Privacy Rights Clearinghouse (2002) warns of chilling effects stemming from the rampant disclosure of personal information on the Internet, where people begin to avoid contexts where personal information is aggregated and disclosed, effectively avoiding participation in civic and public life to try to maintain control over their privacy.

This is especially true for the highly disadvantaged groups also disproportionately likely to be impacted by the criminal legal system (Western 2006). Studies confirm the disparate impact of big data technologies on already marginalized populations (Barocas and Selbst 2016) and class differentials in privacy intrusions (Gilman 2012). This aspect is further exacerbated for marginalized groups and people of color who are already under intense regimes of surveillance and monitoring through the administration of public benefits, subsidized housing, medical care, and legal aid (Eubanks 2018). Empirical research has shown that people facing online versions of their criminal records “opt out” of social situations that may trigger a Google search in an act of digital avoidance (Lageson 2016; see also Brayne 2014; Goffman 2014).

The affirmative disclosure of criminal records, whether through arrest logs, inmate rosters, or compiled criminal histories, sits at an uneasy crossroads between debates between transparency and privacy and over what constitutes a public record and a criminal record. As digital criminal records continue to evolve in form and function, ranging from applications in public notification services, personal data markets, and commercial background checks, the time is ripe to empirically assess the current state of criminal record disclosure.

RESEARCH QUESTIONS AND METHODOLOGY

Given the evidence on the effects of criminal record disclosure and the growing extralegal uses of criminal record data, we lack a baseline understanding of the volume and qualities of publicly disclosed, digital criminal records across the United States. This analysis aims to provide a foundational understanding of personal data disclosure by the criminal legal system about suspects, defendants, and incarcerated people. Specifically, we ask: what are the digital criminal

record disclosure practices of criminal justice agencies (law enforcement, courts, corrections, and criminal record repositories) across the United States? What are the state-specific data disclosure practices? Given the relative magnitude of criminal justice operations in a given state, how many individual criminal records are created and affirmatively disclosed on the Internet each year?

We employ a content analysis of public criminal record databases to answer these questions, identifying and coding personal data contained on state-run websites. At each stage of criminal justice processing, adult defendants are recorded and cataloged into digital databases.¹⁰ Many criminal justice agencies then also provide this information for free online. Users can search for criminal history information by entering a basic query for a single record or through browsing a bulk database. We classify the types of criminal record data contained in these online databases as evidence of an agency's criminal record disclosure practices. We first identified the main website for each branch of every US state's criminal justice operations. The branches included in the analysis are law enforcement, criminal courts, corrections, and criminal history repository departments. For law enforcement, we analyzed the website for the most populous county in the state. For criminal courts, we used a centralized state database, if available. For states that maintain separate online databases for each jurisdiction, we again used the most populous county's court website. For correctional information, we accessed the website of each state agency responsible for maintaining the prison population. For compiled criminal histories, we identified the website for the agency that operates the criminal records repository. The URLs for each website and the full dataset are available upon request.

Once we accessed each website, we identified whether or not the website disclosed any criminal history information. If so, we ran a basic query using a single alphabetical letter to populate search results. If a full last name was required, we used "Anderson," "Smith," or "Perez" to populate search results (for a similar study approach, see Conley et al. 2011). Once we retrieved publicly available data, we coded the information available on the website, which is detailed below. Our coding scheme was straightforward, detailing the existence or nonexistence of a type of personal identifier or criminal record information. If the results were kept behind a paywall or required an authorized user login, we therefore could not conduct any further coding of personally identifiable information. In other words, we coded based on what was free of

¹⁰ Juvenile records are typically protected from disclosure. For more information, see Juvenile Law Center. "National Scorecard," <https://juvenilerecords.jlc.org/juvenilerecords/#!/map>.

charge and in an open access Internet forum. The first two authors collaboratively coded each website, and then the third author independently verified the entire dataset of results. Any discrepancies were addressed through group consensus by returning to the database and recoding as a group. Each type of web-based database is described below in detail.

Law Enforcement Data

Local law enforcement agencies post information about arrests as a performance indicator demonstrating law enforcement activity to the public through online arrest or jail rosters. Figure 1 displays two examples of common law enforcement websites that display a range of personal information, including photos and personal identifiers. These rosters or galleries are sometimes updated in real time to reflect the current jail inmate population, while others are maintained online as an archive of all prior arrests. Some rosters include booking photographs, physical descriptors, and home addresses, while others reveal very limited information beyond an arrestee's name. These rosters are also important for reporting bail amounts, a service that is helpful for family members to prepare payments as well as a useful resource for bail bond companies. Local media also widely use these arrest logs to report "police blotter"-type articles or to post mug shots on local news websites. Criminal history information available on law enforcement websites was coded for seven dichotomous indicators: (1) where the agency provides online access to booking records; (2) where the results include an arrestee's photo; (3) where the results include an arrestee's date of birth (coded as 0.5 for websites that only disclose birth year); (4) where the results include an arrestee's address (coded as 0.5 for websites that only disclose city rather than full street address); (5) where the results include a description of the arrestee's physical characteristics (such as race); (6) where the results include an arrestee's offense detail; (7) where the results include an arrestee's bail and/or bond amounts; and (8) if the criminal history information stayed online for a period of time after release from jail or booking procedures. Scores for each county ranged from zero to eight, which was converted to a percentage score.

[Insert Figure 1 about here]

Court Data

Courts archive data about defendants and proceedings, and this information has long been accessible to the public for inspecting and copying at the courthouse. Digitization has reduced the need to visit a courthouse in person, though some states offer access to court files only through public access computer terminals. Some states centralize the data through a statewide database. In other states, each individual jurisdiction (such as a municipal or superior court or a county-level court system) maintains its own database. Figure 2 provides several examples of common court record lookup tools. Individual person search results on court websites were coded as follows: (1) where the court provides online access to person-level information; (2) where the court provides free public access (no paywall or paid subscription required); (3) where the information includes the defendant's date of birth (coded as 0.5 for websites that only disclose birth year); (4) where the information includes the defendant's address (coded as 0.5 for websites that only disclose the city rather than the full street address); (5) where the information includes the defendant's physical characteristics; (6) where the information includes the fines owed; (7) where the information includes the defendant's offense detail; and (8) where the information includes the name of the defense attorney. Scores for each state ranged from zero to eight, which was converted to a percentage score.

[Insert Figure 2 about here]

Prison Inmate Data

Online directories of state prison data are typically centralized to a single Department of Corrections website. Several states use a third-party service to maintain their online repository, while other states maintain their own publicly available database. Figure 3 shows several examples of these websites, including one state-run website (Colorado) and one commonly used third-party service, called VINELink (Victim Information and Notification Everyday). It is worth noting that many states have been aided in this effort through victims' services funding to streamline notification services when a prisoner is transferred or released (see, for instance, State of Delaware 2008). State Department of Corrections websites also offer inmate locators for other

audiences. These sites may provide instructions for friends and family of inmates on how to visit or send a package to the location. For our analysis, search results for an inmate were coded as follows: (1) where the state offers online access to an inmate directory; (2) where the results include an inmate's photo; (3) where the results include a description of an inmate's physical characteristics; (4) where the results include a defendant's date of birth (coded as 0.5 for websites that only disclose birth year); (5) where the results include an offense's detail; (6) where the results include the inclusion of probationers and/or parolees; and (7) where the results include the inclusion of former inmates no longer under state supervision. Scores ranged from zero to seven, which was converted to a percentage score.

[Insert Figure 3 about here]

Compiled Criminal Histories

Federal case precedent dictates that conviction histories need not be automatically disclosed in the public record because they represent a compilation of data from various agencies.¹¹ Each state maintains an agency responsible for collecting local criminal record data, such as arrests, charges, and case dispositions, and aggregating them into a full report. By state law, these reports are the most protected, and least accessible, piece of criminal record data. Because this study investigates open access to criminal record data, we did not pay a fee to access protected state-compiled rap sheets. Accordingly, states were coded as follows: (1) where there was online access to a state criminal record summary and (2) where there was no-cost access to a state criminal record summary. Because only one state (Minnesota) provides conviction summaries at no cost, we did not conduct detailed coding. Scores ranged from zero to two, which was converted to a percentage score.

State Disclosure Scale

To assess how a state lands in the continuum of data accessibility, we totaled scores across each branch of criminal justice within each state—law enforcement, court, prison, and state

¹¹ *US Department of Justice v. Reporter's Committee for Freedom of the Press*, 89 US 749 (1989).

repositories of compiled criminal records. This aggregate value represents a state-level disclosure score, where higher scores represent higher levels of disclosure across all criminal justice agencies in the state and greater availability of individual-level, identifiable criminal record data on the Internet. Scores ranged from zero to twenty-five, which was converted to a percentage score.

LIMITATIONS

Our study involved several important limitations. First, we used the criminal record disclosure practices at the county level for all law enforcement agencies and a substantial number of criminal courts as a proxy for state practices. This approach assumes that the practices of the largest and most visible criminal justice agencies within the state represent the practices of smaller counties across the state. A related limitation is that the largest law enforcement agency or criminal court might have substantially more resources devoted to creating and maintaining a website or online database; consequently, smaller counties may have the legal ability to post records but lack the administrative structure to do so. As such, our use of the most populous county as a proxy for state dissemination practices overlooks the possibility that the largest county in a state may differ systematically from less populous counties in the same state. However, the scale of a given state's largest county also differs in important ways from larger and smaller states. For instance, the largest county in Montana may vary considerably in dissemination practices from the largest county in Pennsylvania. In this sense, the county proxy serves our research questions well by showing distinctions between state practices.

Likewise, there is a breadth of literature that documents the spread of criminal justice policy and practice across urban, suburban, and rural contexts (Golden 1981; Feld 1991; Weisheit, Falcone, and Wells 1999). For instance, scholarship by Mona Lynch (2018) demonstrates that federal court districts maintain a high degree of geographic coherence in sanctioning trends over time, with minimal deviation between regions even as local and regional factors exert an influence on these practices. In other words, although research indicates variations among county-level jurisdictions (which may be especially likely for technological and administrative practices), evidence also suggests that less populous counties may engage in a "mimetic process," whereby they replicate the strategies of the most populous counties (or

“leading” counties), establishing coherent standards for interorganizational dissemination practices among intrastate agencies (Grattet and Jenness 2005, 903). Further, the rise in third-party software applications—and the vendors that sell these products to local agencies—might result in even more rapid policy adoption (Brayne 2017).

A second limitation is the cross-sectional nature of our data. Agency websites change constantly; indeed, in the course of our data collection and coding, we were compelled to recode several agency’s disclosure policies to accommodate the release of new online databases and more sophisticated websites as well as noting when a state actively removed information that was once made available (the Arizona Department of Corrections, for instance, removed the full date of birth from inmate records in late 2018 after reports of identity fraud). As such, our data provides a snapshot of websites in late 2019, subject to change in a field of rapid transformation and development. Relatedly, we do not explore predictors of technological adoption in this analysis, which might include punitive and political measures, state budget allocations for criminal justice technologies, and the scale of a state’s criminal justice operations. We wholeheartedly encourage further research in that line of inquiry.

Finally, there are important distinctions between the users of criminal record data, ranging from commercial aggregators, to victims, to employers. Our analysis does not delve into the technical solutions that criminal justice agencies have deployed to block bulk downloading or the purchasing options developed by state commissions for users to buy aggregated data through Freedom of Information Act requests. Instead, we aimed to take a first look at the information that criminal justice agencies affirmatively disclose to Internet users with no fee or registration required.

RESULTS

Public Access to Criminal Record Data

Table 1 shows the baseline descriptive statistics for how many US states release law enforcement, court, corrections, and state criminal record data on the Internet. For conceptual clarity, we organize our findings according to four distinct types of criminal record data. The first group includes law enforcement data that document some sort of interaction with the police.

We found that 82 percent of states' most populous counties (N = 41) provide access to arrest data at no cost to Internet users. Additionally, 42 percent of these counties (N = 21) post booking photos online that can be viewed at no cost to the public. For the second group that includes criminal court data, forty-one states make criminal court data available on the Internet. Twenty-eight states centralize their data to a single website, while thirteen states direct each county to organize their data independently. This results in a patchwork of online databases where several counties within a state might use the centralized website, while other counties in that same state might opt out completely. Nine states do not post any online criminal court data. Of jurisdictions that disclose criminal court data (at either the state or county level), nine states charge a fee for accessing the data.

[Insert Table 1 about here]

The third group includes state prison websites, which often consist of a dynamic public access portal for a shifting set of individuals, such as those who are currently incarcerated, were previously incarcerated and released, or are under probation or parole supervision. Every state provides variations of inmate lookup tools, either centralized through the Department of Corrections website or through the aid of a third-party service, such as an "Inmate Lookup Tool" provided by the software company Digital Solutions or the VINELink service that automates victim notifications. Every state provides a real-time roster of current inmates. Moreover, 76 percent of states (N = 38) also maintain a list of former prisoners who are still under correctional supervision, disclosing a roster of current probationers and parolees. Further, some state correctional websites maintain a roster of former inmates and supervisees who are no longer under some form of supervision. In total, eighteen states disclose these former inmate and supervisee records at no cost.

The fourth group includes state-compiled criminal records, which are managed by the state's central repository. Although these compiled records detail the disposition of each criminal case and, therefore, are technically the most accurate version of a criminal record, these are the most difficult records for the public to obtain. Fifty-six percent of states allow the general public

to access compiled criminal history reports from the state repository.¹² However, of the twenty-eight states that disclose criminal records, only one state (Minnesota) provides these records at no cost, while the others charge users a fee to access criminal records. The remaining twenty-two states either (1) do not make criminal records available to the public at all; (2) offer criminal records only to authorized agencies (such as a government employer); (3) only release records to the subject of the criminal record for personal review; or (4) require the subject of the record to authorize a third party to access their information. Notably, some states that offer public access to criminal records will still charge the user a fee to search the state criminal record database even if no record is found. Casual searchers may turn to the less costly and more open option of law enforcement, prison, or court data, which would also reveal arrests that never led to conviction and charges that have been dismissed or expunged.

Personal Identifiers Disclosed in Criminal Records: Agency-Level

The total amount of information disclosed online about an arrestee, defendant, or incarcerated person varies somewhat, but the majority of agencies post physical characteristics and birth year information. Importantly, when agencies include highly detailed personal identifiers, information about residents are more easily integrated into other sources of public and private records, such as voting records, marriage and property records, eviction databases, traffic court data, consumer data, and credit reports (Kaspero and Canhasi 2018).

Table 2 details the frequency of specific personal identifiers disclosed by the type of criminal justice agencies. Seventy-three percent of law enforcement agencies that post records disclose personal characteristics of the arrested person, and over half (51 percent) disclose photographs online. Eighty-eight percent of law enforcement agency websites provide birthdate information, and 54 percent post an arrestee's entire birthdate. Nearly half of the criminal court websites (44 percent) post a defendant's full year of birth, and 82 percent post at least the year of birth. A defendant's full home address is posted by 31 percent of the criminal court websites,

¹² Note that New York is a special case. While the state repository (NY Division of Criminal Justice Services) does not allow public access to compiled criminal histories (known in the state as CHRI for criminal history record information), the state court system does allow the public to order a "criminal record" for a fee, using only court data. We do not include New York in our estimates of official state criminal record disclosure, but we do include New York in our magnitude estimates of the disclosure of felony convictions.

with another 16 percent posting the city of residence. Nearly half of criminal court websites (44 percent) post physical characteristics of the defendant, such as race, height, and weight. The majority of correctional agencies include physical characteristics (82 percent of states), inmate photographs (60 percent), and year of birth (88 percent).

[Insert Table 2 about here]

We further created standardized scales of disclosure for each criminal justice agency within each state to determine which type of agencies disclose more data at the aggregate, national level (see Table 3). We did not include conviction summaries in this analysis, as the vast majority are behind a paywall and not available to the public free of charge. Of the remaining categories (law enforcement, courts, and corrections), prison inmate records were the most “open,” scoring an average of 7.03 on a ten-point scale, while court records were the most “closed,” at 4.79 on average on a ten-point scale.¹³

[Insert Table 3 about here]

Personal Identifiers in Disclosed Criminal Records: State Level

We now shift our focus to the state level, examining the degree to which a particular state releases personally identifiable criminal record data through any criminal justice platform. Table 4 highlights the nationwide prevalence of the release of full names, home addresses, and birthdates by states. All states affirmatively disclose prisoners’ full names, while the majority of states post the full names of arrestees (82 percent) and criminal defendants (64 percent). The arrestee’s city of residence is disclosed in 20 percent of states (8 percent reporting the full home address), and 29 percent of states disclosing the full address of criminal defendants (with an additional 10 percent disclosing the defendant’s city of residence). The year of birth is reported

¹³ We also analyzed correlations between types of personally identifiable information released (name/date of birth and physical characteristics) types of criminal justice-specific information released (offense type and archived post-release or disposition). We found positive and significant relationships across all data sources for law enforcement and courts but not for prison websites. This may indicate that, if a software platform allows for more data to be posted, agencies are more likely to release all available types. Full results are available from the authors.

for 72 percent of arrestees (44 percent reporting the full date of birth) and for 52 percent of court defendants (28 percent reporting the full date of birth). Over half (52 percent) of US states post the full date of birth of people who are incarcerated, with another 36 percent posting the year of birth. We also examined the release of personally identifiable pre-conviction data released by states through law enforcement and court websites. We found that the majority of states affirmatively post PII on the Internet at no cost. Only four states do not disclose PII in law enforcement and court websites (Delaware, Massachusetts, New Hampshire, and Wyoming). Figure 4 displays the prevalence of releasing pre-conviction records.

[Insert Table 4 about here; Figure 4 about here]

We then analyze within-state variation, analyzing correlations between the four criminal justice agencies. Put differently, do the disclosure practices of one agency (for example, the police) correlate with the data disclosure practices of a different agency (for example, corrections) in the same state? For the correlations, we used the standardized ten-point scale for each branch of criminal justice.¹⁴ Table 5 shows basic correlations and significance levels. Results show a modest, positive, and statistically significant correlation between law enforcement disclosures and prison record disclosures and another modest, positive, and statistically significant correlation between rap sheet disclosure and court record disclosure. These mixed results could represent a lack of agreement between agencies about how to disclose PII and what should be included and about how to reflect the competing legislative and regulatory guidelines that govern each agency separately or they could be driven by differences in county- versus state-level activities (given our methodological choice to use the largest county as a proxy for state law enforcement). Replicating this study with county-level data for each state could better answer questions of within-state variation.

[Insert Table 5 about here]

¹⁴ We conducted additional analysis for our full data set (using twenty-five indicators across all types of criminal justice data, as displayed in Figure 4) and a reduced data set (using eighteen indicators that potentially appear across the three main types of justice data (law enforcement, courts, and prisons). Specifically, indicators include: name, no paywall, date of birth, offense, physical characteristics, and if data remained online post-disposition or release). The results for all three analyses were similar in degree, direction, and statistical significance of correlations.

We then turn to between-state variation. To understand how states compare to one another, we calculated a total criminal record “disclosure score” for each state that ranged from zero to twenty-five and included all types of criminal justice data available across agencies, with higher scores reflecting more disclosure across all branches of criminal justice. Figure 5 shows the aggregate score of all criminal justice agencies within each state, with Florida disclosing the most information across agencies (score = 21) and New Hampshire disclosing the least (score = 2.5). Figure 6 graphically displays disclosure scores on a US map. Several specific criminal justice agencies and states disclosed a significant amount of personally identifying information. High-scoring law enforcement agencies often detail a defendant’s full name, race, sex, date of birth, and home address on its website. Some websites report data beyond our coding scheme, incorporating data including build type, skin tone, eye color, and place of birth. The degree of detail about the case itself also varies, such as including information regarding court personnel, bond amount(s), fees assessed, charge details and types, and links to supporting court documents.

[Insert Figure 5 about here; Figure 6 about here]

Prison records disclose the most data: seven states scored ten out of a possible ten on their disclosure of personal information on prison inmate records.¹⁵ The majority of states provide the full date of birth of individuals, along with hair and eye color. Several states, including Georgia, Michigan, and Montana, include scars, marks, and tattoos as a list of descriptions of an individual. Montana also provides citizenship and resident status, along with the birthplace of an individual (see Figure 7). Offense information and sentence details are also disclosed, all with varying degrees of detail. Importantly, with the exception of Kansas,¹⁶ the seven highest-scoring states supply this information for both individuals who are currently incarcerated or under community supervision (for example, probation or parole) as well as for those who have been discharged from the system entirely.

[Insert Figure 7 about here]

¹⁵ Kansas, Michigan, Nebraska, New Jersey, North Carolina, Ohio, Oklahoma.

¹⁶ This state only lists individuals currently incarcerated and on supervision.

In sum, criminal justice agencies lean toward presumptive disclosure of many types of criminal history information. They also vary somewhat in how agencies and states release detailed personally identifiable data that is arguably non-criminal justice related, such as home address and birthdate. Given the literature on the fragmentation of the criminal legal system even within a single state, the mixed results are perhaps unsurprising (Pfaff 2017). Other research on technological adaptation has shown similar unevenness across agency contexts (see, for instance, Brayne and Christin, forthcoming). For our purposes, divergent practices between branches of the system may be due to third party contractual pressures, the availability of funds, and the specific needs of the agency, which are points we will return to in the discussion. Next, we calibrated the volume of data disclosures, given the data practices we have documented here.

Volume of Data Disclosures

The trend toward increasing digital criminal record disclosures by states and agencies must also be calibrated to the scope of the American criminal legal system. Comparing arrest, criminal caseload, criminal convictions, and incarceration rates to state record disclosure practices begins to answer the question of how many people in America have their arrest and conviction data publicized on the Internet or, more precisely, the raw number of records that are disclosed each year. Combining our public disclosure indicators with state- and county-level arrest, charging, and conviction data, we projected how many records might be created over one-, five-, and ten-year periods. To calculate these projections, we combined administrative criminal justice processing data with our earlier findings of state disclosure practices. For example, we tallied the amount of annual state arrests for each of the forty-one states that disclose law enforcement records to obtain a projection for annual data disclosures. Arrests that occurred in states that do not disclose records online are not included in this count. Specific data sources are described in each of the results presented below.

Table 6 provides estimates for each type of criminal record disclosure, where we projected the sheer number of records that are produced in the United States each year as well as the number of individuals impacted over time, given each state's practices regarding criminal record disclosure. For law enforcement and criminal court records, we did not measure the

number of individuals whose information is disclosed online but, rather, the number of individual records that are created each year (as some individuals may be arrested several times per year, with each arrest creating a new data point, arrest record, or mug shot). This is an important estimate for understanding the potential contribution of criminal justice agencies to the personal data market, particularly through the online disclosure of personally identifiable data that can be utilized by third parties for non-criminal justice ends.

[Insert Table 6 about here]

<H3>Arrest Records and Mug Shots

For arrest records and mug shots, we use uniform crime reporting data obtained through the Inter-university Consortium for Political and Social Research to estimate the adult arrest rate for 2014 for each state. To estimate the magnitude of publicly disclosed arrest records, we calculated the number of arrests for the forty-one states that automatically publish arrest records, measured at the largest county level. Given these parameters, we estimate that approximately 10.16 million digital arrest records are created across the country each year, culminating into over 50.8 million arrest records over five years, and 101 million arrest records over a decade. Booking photos are an especially valuable resource on the criminal record data market, evidenced by the hundreds of mug shot websites that exist (Lee 2018). Across the twenty-one states that automatically disclose booking photos on law enforcement websites, we estimate that 4.57 million mug shots are released annually, culminating in over 22.8 million mug shots in five years and 45.7 million mug shots in over a decade of record disclosure.

<H3>Court Records

For court records, we relied on data obtained from the Court Statistics Project of the National Center for State Courts, utilizing 2010 Statewide Total Criminal Caseload data, the most recent complete dataset available, for all disclosing states (including those we coded at the county level). Court records comprise the most massive criminal record dataset available to companies, consumers, and website users. Thirty-two states post court records to the Internet at no cost.

Using caseload data for each state, we estimate that 14.7 million criminal court records are disclosed on the Internet each year, culminating in 73.5 million records over five years and 147 million records over a decade. Accounting for states that disclose criminal court records to the public and that also charge a fee to access these records (N = 41 total), these estimates rise to 19.5 million criminal court records per year, 97.5 million over five years, and over 195 million criminal court records over a decade.

<H3>Prison Records

For prison records, we used data from the Bureau of Justice Statistics Prisoners in 2016 to measure annual incarceration rates. We estimate the average number of available prison and supervision records using state measures for current incarcerated population and current probation and parole population. These estimates show that, across all fifty states that publish inmate rosters, approximately 1.3 million current prison inmate records are available online. Across the thirty-eight states that disclose current probation and parole populations, 2.4 million individual records are available on the Internet.

<H3>Cumulative Criminal Records

Finally, we provided cumulative measures of prison and conviction records that remain online indefinitely after a sentence has been completed. This estimate more accurately reflects the number of individuals whose criminal histories appear on websites rather than the number of records, which are described in the previous sections. We use Sarah Shannon and colleagues's (2017) aggregate measures of former prisoners and estimates of people with a current or past felony conviction to estimate the number of people whose conviction records remain online, given the data disclosure practices of the individual's state of incarceration or conviction. Shannon and colleagues's approach, using demographic life tables, provides estimates up to 2010; thus, the actual volume is likely even higher than the conservative projections we estimate here. Across the eighteen states that provide rosters of people who were previously convicted of a felony and were incarcerated, we estimate that 6.5 million people have a prior prison record that is available on the Internet at no cost. Across the twenty-eight states that disclose conviction

summaries, 12.4 million people have a felony conviction record that is accessible by the general public through a governmental website.¹⁷ However, the vast majority of states charge a fee for accessing this information.

DISCUSSION

Our analysis highlights several significant trends in state digital criminal record disclosures. First, we show a strong trend toward disclosing criminal record data, with some diversity across and within states in terms of the types of personal information made available online. While some states elect not to disclose any criminal history information, other states err on the side of presumptive and widespread access to records. Thus, in the same manner that all criminal justice practices are “local,” so too are criminal justice data disclosure practices (Lynch 2011). We also calibrate the specific types of personally identifiable information to the scope of a state’s criminal justice operations to provide a baseline sense of the volume and magnitude of criminal record data that is released online every year. Taken together, the breadth and volume of criminal record data are expansive, pushing the boundaries of the types of personally identifiable information now placed under the umbrella of “criminal records” and creating an impressive online archive that has accumulated rapidly given the size of the American criminal justice operations.

Our analysis also shows that a remarkable quantity of PII is disclosed by government agencies through criminal record data publication. While our analysis focuses on the scope and volume of these data, we can only infer the abundance of extralegal uses of such information. Scholarship in privacy law and socio-legal studies provide some insight. Amanda Conley and colleagues (2011, 803) writes: “This information can be aggregated and linked to particular individuals by companies that traffic in the free collection and highly profitable sale of large amounts of personal data,” and they also cite the potential for identity theft, stalking, discrimination, and harassment (Gomez-Velez 2005). Alongside the broader “quantification of criminal court operations” (Lynch 2019, 31), developments in big data policing (Brayne 2017; Ferguson 2019), the use of algorithmic risk assessment tools (Simon 2005), and advances in

¹⁷ This analysis includes New York State, which provides conviction records for a fee via the state court system as well as the state repositories that collect a fee to access conviction information.

facial recognition technologies (Hamann and Smith 2019; Hayes 2019; Gershgorn 2020) there are potentially vast legitimate and illegitimate applications for the volumes of criminal record data, personal information, and booking photos that we describe here. Without careful concerns for accuracy and privacy, many legally innocent people will find their data in a variety of private sector endeavors. Information gleaned online by decision makers has the potential to be outdated or inaccurate—for instance, a police department may arrest a person for a serious charge, which prosecutors decline to pursue or judges dismiss due to a lack of evidence. Having access to only the arrest record will incorrectly imply a serious criminal record, potentially leading to housing or employment discrimination based on faulty information.

This is a troubling finding considering the application of stigma and the attachment of guilt inherent to online criminal record data (Jacobs 2015; Corda 2016). Arrest records lack subsequent charging information, and court records may report dismissed charges or contain incomplete disposition information. From a data integrity standpoint, disclosed records become “stale” as a case proceeds through the justice system, even though the potential for duplication has already taken place (Lageson 2020). Furthermore, a person may have already been labeled as a criminal across a broad spectrum of databases—even if their charges are dropped or their record is later expunged. In this light, the proliferation of online criminal record data reflects larger punitive trends like mass incarceration and, even more so, the mass production of “criminal” subjects that the system “can neither govern adequately nor eliminate permanently” (Simon 2007, 175). As criminal histories occupy larger volumes of virtual space, criminal history data exert greater influence over the lives of private citizens.

These trends also raise important due process questions: given the documented, negative consequences of having a criminal record, how then do divergent disclosure practices produce different patterns of punishment across states? Might different criminal record disclosure practices produce a new form of “justice by geography” when one person’s arrest in one state can remain in practical obscurity, while in another state a person’s record is automatically disclosed on the Internet (Feld and Schaefer 2010, 328)? Geographic variations may accelerate or mitigate the degree of disclosure among jurisdictions within and across states and regions.

While access to technology resources drives some of this variation, policy directives are also important to consider. State constitutional mandates and transparency laws based on the Freedom of Information Act often guide the disclosure of pre-conviction records (such as police

and court data), while state penal codes often regulate access to compiled criminal histories (Lageson 2020). Our analysis confirms that compiled criminal histories are the least publicly accessible type of criminal record data, while pre-conviction police and court records are the most publicly accessible. As a practical consequence, the most “complete” version of a person’s record is less accessible to the broader public. On this score, at least, state practices are remarkably uniform. In addition, this study raises broader questions about public perceptions of crime and safety. The proliferation of criminal record data and the resultant mass production and reproduction of non-conviction data—both justified and unjustified, recent and outdated—might promote false public assessments of victimization risk. Therefore, the public notification function of online criminal record disclosure is compromised by the inconclusive and piecemeal quality of the information. The sheer magnitude of records produced each year would only serve to exacerbate such broader effects.

These results and potential implications show how the criminal legal system has become a key producer of online privacy infringement, placing this governmental institution alongside the technology and data companies that are typically seen as having such power over the distribution of privacy. The rapid pace of the digital environment has created a context collapse between the varied functions of criminal records in American society, where records operate simultaneously as a public record of government agency activity, a valuable personal data commodity, and a stigmatizing label available to the public. In this way, transparency laws and criminal justice data practices may contribute to differential citizenship by turning the lens onto the people processed through the system rather than disclosing criminal justice data for the Brandeisian “sunshine” rationale of governmental watchdogging (Miller and Stuart 2017).

This shifting lens has consequences for actual government transparency. There is often scant information about prison conditions, parole hearings, and prosecutorial charging and bargaining discretion (Sklansky 2018). Law enforcement maintains near total control over other forms of data, such as disciplinary files, police shooting data, or the contents of gang databases. Technologies are even decreasing transparency in some law enforcement domains, such as the recent shift toward encrypted digital communications platforms that eliminate the ability of the public and media to listen to police scanner activity (McCoy 2018). Criminal justice agencies can also invoke Freedom of Information Act exemptions to protect certain types of agency-specific data (US Department of Justice 2019). It is essential to distinguish “transparency”

between data releases that facilitate the monitoring of state action and those that facilitate the monitoring of fellow community members.

In short, the release of criminal record data also circumscribes privacy rights as a new and defining feature of punishment. Inequality and technology research shows how low-income people already face a matrix of vulnerabilities through the collection and aggregation of big data (Madden et al. 2017), especially in their ability to protect personal information online, prevent digital privacy harms, and police their online persona (Madden 2017). Our study is a first step toward understanding how the criminal legal system distributes not only punishment and stigma but also the ability to maintain digital privacy rights at all. Technological innovation has disrupted the delicate balance between maintaining public access to state criminal justice operations and large-scale digital punishment for people ensnared in the system who must also now contend with the online release of personally identifiable data. A better balance between public punishment and surveillance and personal privacy could more accurately reflect both legislative intent and legitimate public interests. Such a balance could be reached by limiting non-criminal justice-related PII in publicly available websites, requiring users to register with the governmental website to access data, implementing restrictions for bulk downloads and web scraping, and reconsidering the public release of pre-conviction records. At the same time, criminal justice agencies should be held to the same standards of openness that have been implemented at the arrestee and defendant levels, where the public can easily retrieve agency-level data about arrest rates, charging decisions, and correctional operations. Digital privacy must also become part of legal reform. As states continue to adopt “clean slate” legislation (Collateral Consequences Resource Center 2020), and bipartisan criminal justice reform increasingly focuses on expungements, record sealing, and second chances, the technological and privacy elements of criminal punishment must become central to the debate.

REFERENCES

- APPRISS Safety*. 2020. "Vine." <https://apprissafety.com/solutions/vine/>.
- Barocas, Solon, and Andrew D. Selbst. 2016. "Big Data's Disparate Impact." *California Law Review* 104, no. 671: 671–732.
- Brandeis, Louis. 1914. *Other People's Money and How the Bankers Use It*. New York: Frederick A. Stokes Company.
- Brayne, Sarah. 2014. "Surveillance and Systems Avoidance: Criminal Justice Contact and Institutional Attachment." *American Sociological Review* 79, no. 3: 367–91.
- . 2017. "Big Data Surveillance: The Case of Policing." *American Sociological Review* 82, no. 5: 977–1008.
- Brayne, Sarah, and Angèle Christin. Forthcoming. "Technologies of Crime Prediction: The Reception of Algorithms in Policing and Criminal Courts." *Social Problems*.
- Brobst, Jennifer. 2015. "Reverse Sunshine in the Digital Wild Frontier: Protecting Individual Privacy against Public Records Requests for Government Databases." *Northern Kentucky Law Review* 42: 191-286.
- Burchfield, Kelly M. 2011. "Residence Restrictions." *Criminology & Public Policy* 10, no. 2: 411–19.
- Bushway, Shawn D., and Robert Apel. 2012. "A Signaling Perspective on Employment-Based Reentry Programming: Training Completion As a Desistance Signal." *Criminology & Public Policy* 11, no. 1: 21–50.
- BustedMugshots.com*. 2020. "About." <https://www.bustedmugshots.com/about>.
- Carden, Dan. 2019. "Every Hoosier Convicted of a Felony Could Find Themselves Listed on New State Website." *Times of Northwest Indiana*, February 18. https://www.nwitimes.com/news/local/govt-and-politics/every-hoosier-convicted-of-a-felony-could-find-themselves-listed/article_a9ee6953-3cd8-576f-ae95-7c7b00819ea8.html
- Chiricos, Ted, Kelle Barrick, William Bales, and Stephanie Bontrager. 2007. "The Labeling of Convicted Felons and Its Consequences for Recidivism." *Criminology* 45, no. 3: 547–81.

- Gershgorn, Dave. 2020. "Clearview AI: We Are 'Working to Acquire All U.S. Mugshots from Past 15 Years.'" *OneZero*, March 5. <https://onezero.medium.com/clearview-ai-we-are-working-to-acquire-all-u-s-mugshots-from-past-15-years-645d92319f33>.
- Gilman, Michele Estrin. 2012. "The Class Differential in Privacy Law." *Brooklyn Law Review* 77, no. 4: 1389–1445.
- Goffman, Alice. 2014. *On the Run: Fugitive Life in an American City*. Chicago: University of Chicago Press.
- Golden, Kathryn. 1981. "Rural Courts: An Illinois Study." *Criminal Justice Review* 6, no. 1: 38–42.
- Gomez-Velez, Natalie. 2005. "Internet Access to Court Records: Balancing Public Access and Privacy." *Loyola Law Review* 51: 365–438.
- Goggins, Becki R., and Dennis A. DeBacco. 2018. *Survey of State Criminal History Information Systems, 2016: A Criminal Justice Information Policy Report*. Washington, DC: Bureau of Justice Statistics. <https://www.ncjrs.gov/pdffiles1/bjs/grants/251516.pdf>.
- Grattet, Ryken, and Valerie Jenness. 2005. "The Reconstitution of Law in Local Settings: Agency Discretion, Ambiguity, and a Surplus of Law in the Policing of Hate Crime." *Law & Society Review* 39: 893–942.
- Hamann, Kristine, and Rachel Smith. 2019. "Facial Recognition: Where Will It Take Us?" *American Bar Foundation Criminal Justice Magazine* 34, no. 1: https://www.americanbar.org/groups/criminal_justice/publications/criminal-justice-magazine/2019/spring/facial-recognition-technology/.
- Hayes, Michael. 2019. "NYPD Is Using Sealed Mugshots In Its Facial Recognition Program." *OneZero*, August 27. <https://onezero.medium.com/exclusive-the-nypd-is-using-sealed-mug-shots-in-its-facial-recognition-program-bd5678ad5632>.
- Hu, Margaret. 2016. "Big Data Blacklisting." *Florida Law Review* 67: 1735–1809.
- Jacobs, James B., and Tamara Crepet. 2007. "The Expanding Scope, Use and Availability of Criminal Records." *New York University Journal of Legislation and Public Policy* 11: 177–213.
- Jacobs, James B. 2015. *The Eternal Criminal Record*. Cambridge, MA: Harvard University Press.

- Kaspero, Esq, and Darsei Canhasi. 2018. "Criminal Record Privacy and the Structural Risks Inherent within Commercial Storehouses in the Consumer Data Industry." *Social Science Research Network*. <http://dx.doi.org/10.2139/ssrn.3130418>.
- Kinstler, Linda. 2013. "When Openness Backfires: Is There Room for More Gray Area in How Court Records Are Made Public?" *NeimanLab*, July 19.
<https://www.niemanlab.org/2013/07/when-openness-backfires-is-there-room-for-more-gray-area-in-how-court-records-are-made-public/>.
- Kurlychek, Megan C., Robert Brame, and Shawn D. Bushway. 2006. "Scarlet Letters and Recidivism: Does an Old Criminal Record Predict Future Offending?" *Criminology & Public Policy* 5, no. 3: 483–504.
- Lageson, Sarah E. 2016. "Found Out and Opting Out: The Consequences of Online Criminal Records for Families." *Annals of the American Academy of Political and Social Science* 665, no. 1: 127–41.
- . 2020. *Digital Punishment: Privacy, Stigma, and the Harms of Data-Driven Criminal Justice*. Oxford: Oxford University Press.
- Lageson, Sarah E., Megan Denver, and Justin T. Pickett. 2019. "Privatizing Criminal Stigma: Experience, Intergroup Contact, and Public Views about Publicizing Arrest Records." *Punishment & Society* 21, no. 3: 315–41.
- Lageson, Sarah E., and Shadd Maruna. 2018. "Digital Degradation: Stigma Management in the Internet Age." *Punishment & Society* 20, no. 1: 113–33.
- Lee, Eumi K. 2018. "Monetizing Shame: Mugshots, Privacy, and the Right to Access." *Rutgers Law Review* 70, no. 3: 557–645.
- Logan, Wayne A. 2009. *Knowledge as Power: Criminal Registration and Community Notification Laws in America*. Stanford, CA: Stanford University Press.
- Logan, Wayne A., and Andrew Guthrie Ferguson. 2016. "Policing Criminal Justice Data." *Minnesota Law Review* 101: 541–616.
- Lynch, Mona. 2011. "Mass Incarceration, Legal Change, and Locale." *Criminology and Public Policy* 10, no. 3: 673–98.
- . 2018. "94 Different Countries? Time, Place, and Variations in Federal Criminal Justice." *Berkeley Journal of Criminal Law* 23, no. 3: 134–64.
<https://doi.org/10.15779/Z38KW57J58>.

- . 2019. “The Narrative of the Number: Quantification in Criminal Court.” *Law & Social Inquiry* 44, no. 1: 31–57.
- Madden, Mary. 2017. “Privacy, Security, and Digital Inequality.” *Data & Society*, September 27. <https://datasociety.net/output/privacy-security-and-digital-inequality/>
- Madden, Mary, Michele E. Gilman, Karen Levy, and Alice E. Marwick. 2017. “Privacy, Poverty and Big Data: A Matrix of Vulnerabilities for Poor Americans.” *Washington University Law Review* 95: 53-125.
- Marwick, Alice E., and danah boyd. 2018. “Understanding Privacy at the Margins.” *International Journal of Communication* 12: 1157–65.
- McCoy, Terrence. 2018. “Last of the Scanners: Are Police Security Measures and New Technologies Killing an American Obsession?” *Washington Post*, December 29. https://www.washingtonpost.com/local/social-issues/last-of-the-scanners-are-police-security-measures-killing-an-american-obsession/2018/12/28/4e344a62-09f5-11e9-a3f0-71c95106d96a_story.html.
- Miller, Reuben Jonathan, and Forrest Stuart. 2017. “Carceral Citizenship: Race, Rights and Responsibility in the Age of Mass Supervision.” *Theoretical Criminology* 21 no. 4: 532–48.
- Nashrulla, Tasneem, and Jennifer Grygiel. 2019. “She Stole Something While Struggling with Heroin Addiction. Cops Turned Her Mugshot into a Meme.” *Buzzfeed*, November 6. <https://www.buzzfeednews.com/article/tasneemnashrulla/cops-humiliate-meme-petty-crime-suspects-facebook>.
- National Center for State Courts. n.d. “Privacy/Public Access to Court Records: Rules on Bulk Data.” <https://www.ncsc.org/Topics/Access-and-Fairness/Privacy-Public-Access-to-Court-Records/State-Links.aspx?cat=Rules%20on%20Bulk%20Data>.
- Pager, Devah. 2007. *Marked: Race, Crime, and Finding Work in an Era of Mass Incarceration*. Chicago: University of Chicago Press.
- Pfaff, John. 2017. *Locked In: The True Causes of Mass Incarceration—And How to Achieve Real Reform*. New York: Basic Books.
- Pound, Roscoe. 1910. “Law in Books and Law in Action.” *American Law Review* 12, no. 1: 12–36.

- Privacy Rights Clearinghouse. 2002. "Public Records on the Internet: The Privacy Dilemma." July 23. <https://www.privacyrights.org/blog/public-records-internet-privacy-dilemma>.
- Restivo, Emily, and Mark M. Lanier. 2015. "Measuring the Contextual Effects and Mitigating Factors of Labeling Theory." *Justice Quarterly* 32, no. 1: 116–41. <https://doi.org/10.1080/07418825.2012.756115>.
- Rosen, Jeffrey. 2000. *The Unwanted Gaze: The Destruction of Privacy in America*. New York: Vintage Books.
- Shannon, Sarah K. S., Christopher Uggen, Jason Schnittker, Melissa Thompson, Sara Wakefield, and Michael Massoglia. 2017. "The Growth, Scope, and Spatial Distribution of People with Felony Records in the United States, 1948–2010." *Demography* 54, no. 5: 1795–1818. <https://doi.org/10.1007/s13524-017-0611-1>.
- Simon, Jonathan. 2005. "Reversal of Fortune: The Resurgence of Individual Risk Assessment in Criminal Justice." *Annual Review of Law and Social Science* 1, no. 1: 397–421. <https://doi.org/10.1146/annurev.lawsocsci.1.041604.120006>.
- . 2007. *Governing through Crime: How the War on Crime Transformed American Democracy and Created a Culture of Fear*. Oxford: Oxford University Press.
- Sklansky, David Alan. 2018. "The Problems with Prosecutors." *Annual Review of Criminology* 1, no 1: 451–69. <https://doi.org/10.1146/annurev-criminol-032317-092440>.
- Solon, Olivia. 2018. "Haunted by a Mugshot: How Predatory Websites Exploit the Shame of Arrest." *The Guardian*, June 12. <https://www.theguardian.com/technology/2018/jun/12/mugshot-exploitation-websites-arrests-shame>.
- Solove, Daniel J. 2001. "Access and Aggregation: Public Records, Privacy and the Constitution." *Minnesota Law Review* 86, no. 6: 1137–1218.
- State of Delaware. 2008. "Delaware Launches Statewide Victim Notification Program," press release. <https://deljjs.delaware.gov/pdfs/Delaware%20Launches%20Statewide%20Victim%20Notification%20Program.pdf>.
- Stelloh, Tim. 2017. "Innocent until Your Mugshot Is on the Internet." *New York Times*, June 3. <https://www.nytimes.com/2017/06/03/opinion/sunday/innocent-until-your-mug-shot-is-on-the-internet.html>.

- Stoll, Michael A., and Shawn D. Bushway. 2008. "The Effect of Criminal Background Checks on Hiring Ex-Offenders." *Criminology & Public Policy* 7, no. 3: 371–404. <https://doi.org/10.1111/j.1745-9133.2008.00516.x>.
- Thacher, David. 2008. "The Rise of Criminal Background Screening in Rental Housing." *Law & Social Inquiry* 33, no. 1: 5–30. <https://doi.org/10.1111/j.1747-4469.2008.00092.x>.
- Uggen, Christopher, Mike Vuolo, Sarah Lageson, Ebony Ruhland, and Hilary K Whitham. 2014. "The Edge Of Stigma: An Experimental Audit of the Effects of Low-Level Criminal Records on Employment." *Criminology* 52, no. 4: 627–54.
- US Department of Justice. 2001. Bureau of Justice Statistics. *Public Attitudes toward Uses of Criminal History: A Privacy, Technology, and Criminal Justice Information*. Washington, DC: Office of Justice Programs, 2001. <https://www.bjs.gov/content/pub/pdf/pauchi.pdf>.
- . 2014. *Survey of State Criminal History Information*. <https://www.ncjrs.gov/pdffiles1/bjs/grants/249799.pdf>.
- . 2019. *Guide to the Freedom of Information Act, Exemption 7(A)*. Washington, DC: Office of Information Policy. <https://www.justice.gov/oip/page/file/1197166/download>.
- Unpublisharrest.com*. n.d. "Mugshots: Legitimate Public Interest." <http://unpublisharrest.com/mugshots-legitimate-public-interest/>.
- Weisheit, Ralph A., David N. Falcone, and L. Edwards Wells. 1999. *Crime and Policing in Rural and Small-Town America*. Prospect Heights, IL: Waveland Press.
- Western, Bruce. 2006. *Punishment and Inequality in America*. New York: Russell Sage Foundation.
- . 2008. "Criminal Background Checks and Employment among Workers with Criminal Records." *Criminology & Public Policy* 7, no. 3: 413–18. <https://doi.org/10.1111/j.1745-9133.2008.00518.x>.
- Yates, Jon. 2012. "Problem Solver: Exonerated Man Feels Mugged by Website." *Chicago Tribune*, March 3. <https://www.chicagotribune.com/business/ct-xpm-2012-03-03-ct-biz-0304-problem-swift-20120303-story.html>.

Table 1. Criminal Record Disclosure Practices		
Type of Access	N of states	% of states
County-Level Law Enforcement Data		
Arrest records (county)	41	82%
Booking photos (county)	21	42%
County and State-Level Criminal Court Data		
Court records (all types)	41	82%
Court records (state centralized)	28	56%
Court records (county centralized)	13	26%
Court records (fee and no fee; county and state)	32	64%
Court records (fee; county and state)	9	18%
No records disclosed	9	18%
State Prisoner and Probationer Data		
Current inmate population	50	100%
Current parole population	38	76%
State Criminal Record Data		
Prison records (ever incarcerated)	18	36%
Compiled criminal record repository (fee and no fee)	28	56%
Compiled criminal record repository (no fee)	1	2%

Table 1. Patterns of agency-level criminal record disclosure practices across the United States

Table 2. Frequency of Personal Identifiers Disclosed Online			
Type of Access	N	% of reporting states	% of all states
<i>Law Enforcement Agencies (N=41 disclosing states)</i>			
Photograph	21	51%	42%
Full date of birth	22	54%	44%
Year of birth or age	14	34%	28%
Full home address	4	10%	8%
City of residence only	6	15%	12%
Physical characteristics	30	73%	60%
Information on released arrestees	15	37%	32%
Offense detail	38	93%	76%
Bail/bond amount	34	83%	68%
<i>Criminal Courts (N=32 disclosing states)</i>			
Full date of birth	14	44%	28%
Year of birth or age	12	38%	24%
Full home address	10	31%	20%
City of residence only	5	16%	10%
Physical characteristics	14	44%	28%
Offense detail	29	91%	58%
Fines owed	28	90%	56%
Attorney named	24	75%	48%
<i>Corrections Agencies (N=50)</i>			
Photograph	30	60%	60%
Full date of birth	26	52%	52%
Year of birth or age	18	36%	36%
Physical characteristics	41	82%	82%
Offense detail	34	68%	68%
Includes parolees/probationers	38	76%	76%
Includes former prisoners no longer in custody	18	36%	36%

Table 2. Frequency of personally identifiable information (PII) disclosed by types of criminal justice agencies across the United States

Table 3. Average State Disclosure Scores (10-point scale)		
Type of Data	Score	Std. Dev
Law Enforcement	5.38	2.98
Courts	4.79	3.93
Prison	7.03	2.16

Table 3. Average degree of PII disclosures by criminal justice agency, standardized to a 10-point scale, where 10 is highest degree of disclosure

Table 4. Online Release of Personally Identifiable Information in Public State Criminal Record Data		
	N of States	% of States
Full name		
Arrestees	41	82%
Criminal court defendants	32	64%
People with convictions	1	2%
Prisoners	50	100%
Home address		
Arrestees (full address)	4	8%
Arrestees (city/state)	6	12%
Criminal court defendants (full address)	10	20%
Criminal court defendants (city/state)	5	10%
People with convictions	0	0%
Prisoners	NA	NA
Birthdate		
Arrestees (full DOB)	22	44%
Arrestees (year only)	14	28%
Criminal court defendants (full DOB)	14	28%
Criminal court defendants (year only)	12	24%
People with convictions (full DOB)	1	2%
Prisoners (full DOB)	26	52%
Prisoners (year only)	18	36%

Table 4. Number and percentage of U.S. states that disclose personally-identifiable data in criminal records posted to criminal justice agency websites

Table 5. Correlations between CJ Agencies within States				
	Law Enforcement	Criminal Courts	Criminal Record Repository	Prisons
Law Enforcement				
Criminal Courts	.07			
Criminal Record Repository	.08	.38**		
Prisons	.43**	.05	.25*	

** p < .05

* p < .1

Table 6. Magnitude of Criminal Record Disclosure					
Type of Record	N of states	% of states	Annual Est.	5-year Est.	10-year Est.
Law Enforcement Information*					
Arrest records	41	82%	10,160,728	50,803,640	101,607,280
Booking photos	21	42%	4,574,740	22,873,700	45,747,400
Criminal Court Information					
Court records (no fee)	32	64%	14,705,837	73,529,185	147,058,370
Prison, Probation and Supervision Information					
Currently incarcerated	50	100%		1,316,205	
Current parolee	38	76%		2,410,000	
State Criminal Records					
Prison record (ever incarcerated)	18	36%		6,551,923	
Felony conviction record	28	56%		12,477,133	

Table 6. Projections of criminal records disclosed on the internet each year. Estimates for law enforcement records and county-centralized court records use the disclosure practices of the largest county as proxy for state-level disclosure practices. State administrative data derived from a variety of public sources; see text for description

Fig. 1 Law Enforcement Websites

Inmate Summary - [REDACTED]

Demographic Information

Name: [REDACTED]
 Global Subject Number: [REDACTED]
 Address: BESSEMER, Alabama 35020



12/20/2016 12:23 AM

Booking History

2016- [REDACTED]

Booking Date: 12/19/2016 12:28 PM
 Release Date:
 Prisoner Type: Pretrial Inmate
 Housing Facility: Jefferson County Jail - Bessemer
 Total Bond Amount: \$265,000.00
 Total Bail Amount: \$0.00

Bond Number	Bond Type	Bond Amount
2016-[REDACTED]	Good Bond	\$60,000.00
2016-[REDACTED]	Good Bond	\$15,000.00
2017-[REDACTED]	Good Bond	\$250,000.00

Charge Description	Offense Date	Attempt/Commit
CAPITAL MURDER DURING ROBBERY 1ST ATTEMPT	12/19/2016 12:30 PM	Commit
ATTEMPT	12/20/2016 12:15 AM	Attempt
Criminal Possession of a Forged Instrument 3	2/13/2017 12:00 AM	Commit

Charges/Warrants/Bond Information

Date: 9/13/2018 5:29:20 PM

Name: [REDACTED]
 ID: [REDACTED]
 Booking Number: [REDACTED]
 YOB: 1993
 Age: 25
 Sex: MALE
 Race: HISPANIC
 Arrival Date: 7/14/2018 4:11:59 PM
 Description: D UNIT

07/14/2018



Fig. 3 Prisoner database maintained by the State of Colorado and Louisiana database maintained by VINE

SMITH, [REDACTED]

Name: SMITH, [REDACTED] DOC Number: [REDACTED]
 Age: 36 Est. Parole Eligibility Date: 06/04/2017
 Ethnicity: WHITE Next Parole Hearing Date:
 Gender: MALE
 Hair Color: BROWN This offender is scheduled on the Parole Board agenda for the month and year above. Please contact the facility case manager for the exact date.
 Eye Color: HAZEL
 Height: 5' 08" Est. Mandatory Release Date: 10/22/2018
 Weight: 185 Est. Sentence Discharge Date:
 Current Facility Assignment: STERLING CORRECTIONAL FACILITY

CURRENT CONVICTIONS

Sentence Date	Sentence	County	Case No.
07/07/2016	3Y-3Y	GARFIELD	16CR100
07/07/2016	18M-18M	GARFIELD	16CR100
07/07/2016	3Y-3Y	GARFIELD	16CR166
07/07/2016	18M-18M	GARFIELD	16CR166

VINE® Victim Information and Notification Everyday Language Louisiana

HOME SEARCH ALERTS

[REDACTED] SMITH
 Custody Record

Age 38
 Gender Male
 Race White

LA: Jackson Parish Sheriff's Office
 ID Number [REDACTED]
 Custody Status **In Custody**
 Custody Detail **Jackson Parish Sheriff's Office**

[RECORD DETAILS](#) [GET NOTIFIED](#)

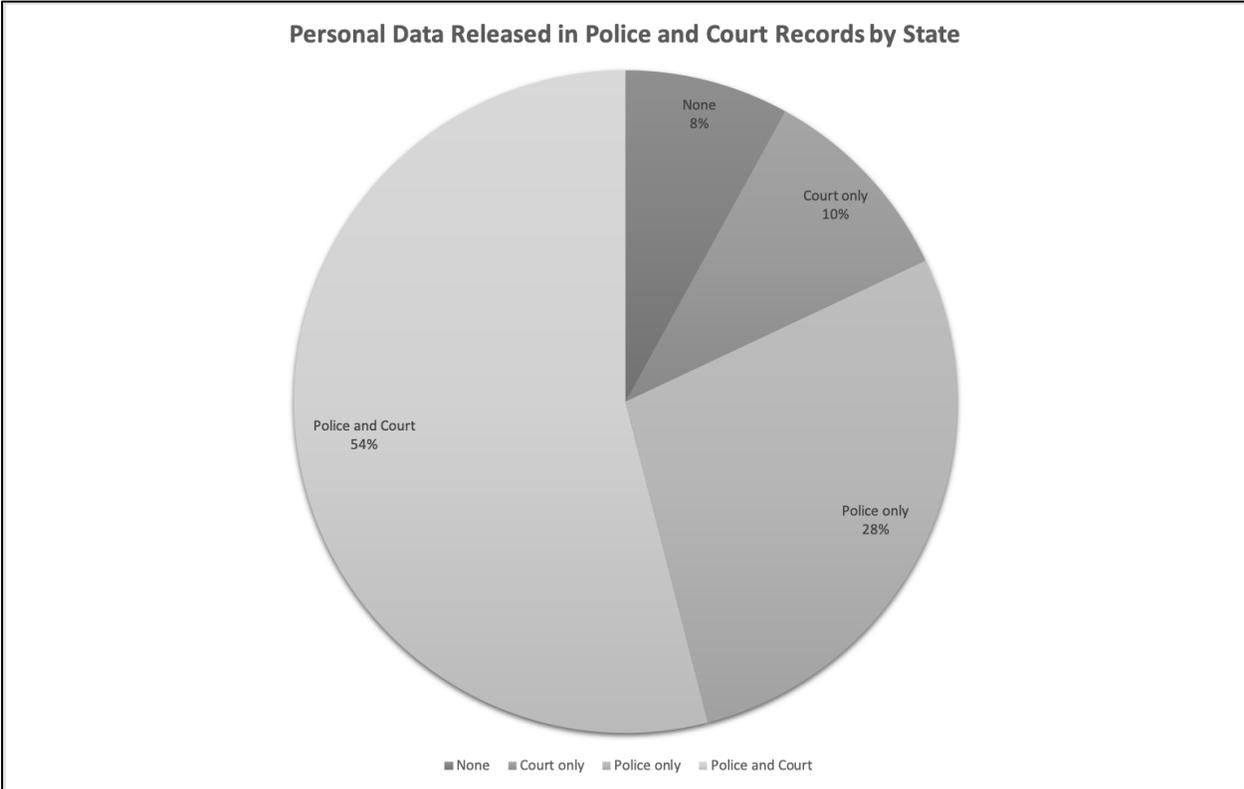


Figure 4. Percentage of states that release personal identifiers in law enforcement and criminal court records that include pre-conviction data.

Figure 5. Disclosure of Personal Identifiers in Online Criminal Record Data by State

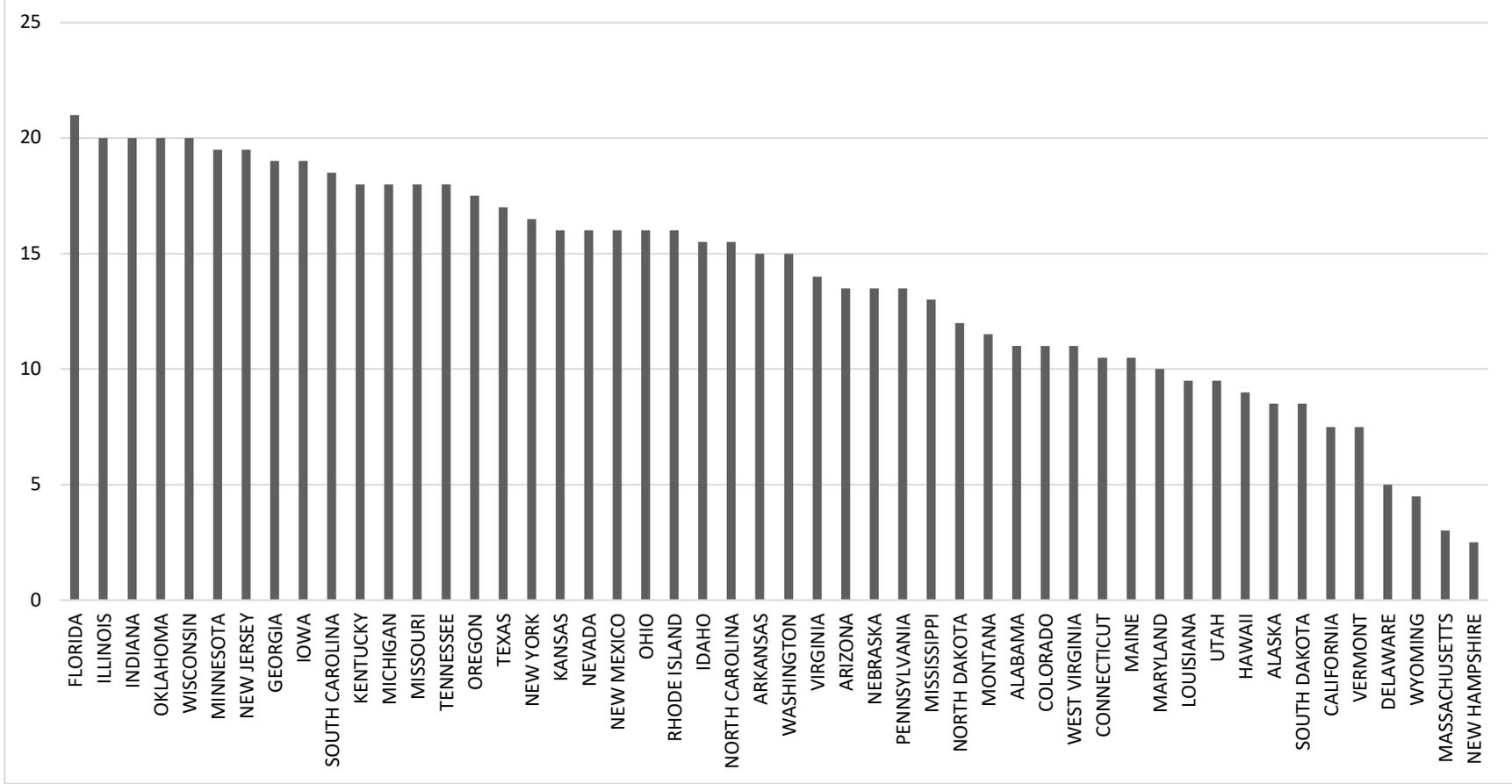
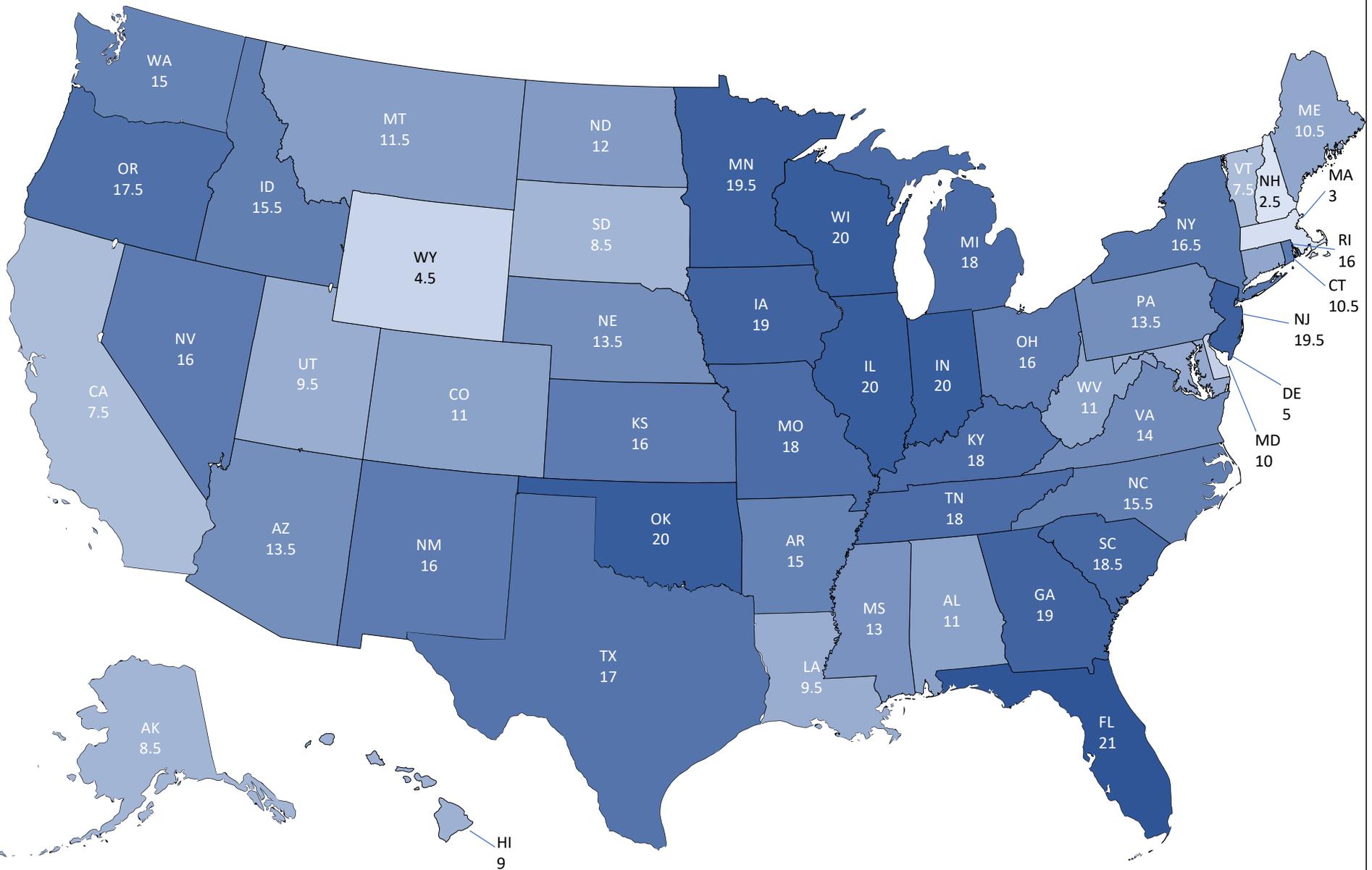


Figure 5. Degree of PII disclosed by states across criminal justice agencies on 0-25 scale, with 25 indicating more disclosure. Scores range from 2.5-21.

Figure 6. Disclosure of Personal Identifiers in Online Criminal Data by State



Powered by Bing
© GeoNames, HERE, MSFT

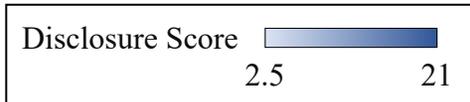


Figure 7. Personal identifiers in Montana prison inmate online database

PHYSICAL AND DEMOGRAPHIC CHARACTERISTICS		
HAIR COLOR: Brown EYE COLOR: Brown HEIGHT: 5 FT 8 IN WEIGHT: 200 lbs BUILD: Stocky	RACE: White SKIN TONE: Light Brown L/R HANDED: RIGHT	YEAR OF BIRTH: 1976 BIRTH PLACE: Unknown CITIZENSHIP: United States MT RESIDENT: Y

AKAs
██████████ Aventura, ██████ Guiteres, ██████ Gutierrez, ██████ Perez, ██████████ Reyes

SCARS, MARKS, TATTOOS AND OTHER PHYSICAL CONDITIONS	
TYPE:	DESCRIPTION:
TATTOO NECK	
TATTOO SHOULDER, RIGHT	
TATTOO ANKLE, RIGHT	
TATTOO ARM, RIGHT, NONSPECIFIC	
TATTOO SHOULDER, LEFT	
TATTOO CALF, RIGHT	