2010

DATABASES, DETERRENCE & THE EXCLUSIONARY RULE

Erin E. Murphy
University of California, Berkeley, erin.murphy@nyu.edu

Follow this and additional works at: https://fordham.bepress.com/ulj

Recommended Citation
Available at: https://fordham.bepress.com/ulj/vol37/iss3/4

This Article is brought to you for free and open access by DigitalResearch@Fordham. It has been accepted for inclusion in Fordham Urban Law Journal by an authorized editor of DigitalResearch@Fordham. For more information, please contact considine@fordham.edu.
DATABASES, DETERRENCE & THE EXCLUSIONARY RULE

Erin Murphy*

Over the past twenty years there has been an explosion in the creation, availability, and use of criminal justice databases. Large scale database systems now routinely inform law enforcement decisions as formal as whether to arrest or convict an individual and as informal as whether to subject a person to secondary pre-flight screening or investigate possible gang membership. Much has been written about the failure of constitutional law to adequately respond to the threat to constitutional rights posed by databases, but less attention has focused upon the awkward fit between database error rights violations and the traditional remedy for constitutional violations: the exclusionary rule. This Essay examines the rise of databases as a tool of law enforcement and sets forth tentative steps toward a theory of constitutional violations in this area.

* Visiting Professor of Law, NYU Law School and Assistant Professor of Law, UC Berkeley School of Law. [THANKS FORDHAM]. I am grateful to the NYU Center on the Administration of Criminal Law for its support of this project, and to Chris Slobogin and the members of the Young Criminal Law Professor Workshop at Vanderbilt Law School for their helpful comments on an earlier draft.
INTRODUCTION

Around 1964, military information systems employees coined the term “data base” to describe repositories of data accessible by users across time-shared computer systems. Shortly thereafter, the commercial data processing world appropriated the term to signify the aggregated information accessed by individual applications. Today, “database” is a word that requires no definition. In thirty short years it progressed from two disconnected descriptive nouns, to a hazy concept stitched together by a hyphen, to a single word instantly understandable to every member of a modern society. Indeed, databases are such an integral component of contemporary life that it is easy to forget just how new they are. Like other currently indispensable technologies – personal computers, cell phones, e-mail -- it can be hard to remember just how things worked before they existed, or how much has changed as a result of their availability.

The relative novelty of databases is just one of the many qualities that makes them difficult to assess. Another is that they are wholly conceptual. Databases do of course manifest physically – the data does exist in ones and zeros out in the world – but their true essence is intangible. Typically, they constitute an expression of collated and diffused pockets of information for dissemination: the code is a database’s only binary aspect. Their value is in their potential to convey – the ability of a search or query to cull and produce information from otherwise difficult to access or survey sources.

A wide variety of government databases have flourished in the criminal justice field over the past ten to twenty years. Law enforcement has benefited from systems of its own design,

---

2 Id.
such as compilations of criminal records, along with the development of general government databases like repositories of highway toll payment systems. Private sector databases, such as those related to mortgages or credit cards, have likewise become common sources for criminal justice investigations. And much legal scholarly attention, including some of my own, has been devoted to the impact of these databases on individual privacy, whether as a general normative matter or in relation to specific constitutional doctrines such as the Fourth Amendment.

This Article assumes a focus different than general concerns about privacy. Indeed, this Article takes as a given that databases affect personal privacy, without quantifying how severe or grave those implications might be. Instead, this Article identifies the challenges that large scale databasing poses to conventional constitutional analysis – though less as a matter of literal doctrine, than as both a theoretical and pragmatic concern. That is, this Article considers the investigative and adjudicative restraints imposed by the Constitution – the Fourth, Fifth, and Sixth Amendments, generally speaking – and aims to sketch ways in which information and evidence drawn from databases conflicts with the assumptions underlying those constitutional regimes. Part One begins by sketching, very briefly, the evolution of databases in criminal justice by providing a sampling of databases in existence. Part Two then examines the few major instances in which the Supreme Court has commented on databases or databasing in the context of criminal justice, in order to glean different approaches or themes that have emerged. Part Three then identifies five presumptions that seem to attach to database-related issues, while Part Four sets out some thoughts about how constitutional doctrine might evolve to the particular needs of databases, in order to better regulate and safeguard their use in the criminal justice system.

---

See, e.g., Erin Murphy, Paradigms of Restraint, 57 DUKE L.J. 1321 (2008).
I. THE RISE OF DATABASES IN CRIMINAL JUSTICE.

A. Proto-Databases.

The emergence of the notion of the formal police force in 1829, when Robert Peel organized the London bobbies, is well known. The first detective unit in the United States was formed in 1846 in Boston, at a time when tracking down criminals largely remained a private sector gig, dominated by companies such as the (in)famous Pinkertons. Their work consisted largely of pounding the pavement (and suspects). Indeed, many of the modern tools of detecting -- “[s]ophisticated criminal investigation techniques- - well organized crime records systems, fingerprints, crime labs – did not appear until the 20th century.” Although Alphonse Bertillion had pioneered his anthropometrical system of identification in the late 1800s, the system depended largely upon manual recording and comparison of measurements.

The first primitive databases emerged around the turn of the century. For instance, as early as 1919, the California State Bureau of Identification introduced a punch-card system for storing and retrieving modus operandi information. But perhaps the most landmark development in government databasing occurred in the early 1930s, around the time that J. Edgar Hoover opened the Federal Bureau of Investigation’s first criminal evidence laboratory, which included fingerprint processing capacities, hair, blood, and firearm analysis. In 1934, the FBI implemented its first fingerprint database – a card sorter that capitalized on the technology

---

5 LAWRENCE MEIR FRIEDMAN, CRIME AND PUNISHMENT IN AMERICAN HISTORY 204. Peel had organized a small detective unit in 1842 in London. Sklansky, Private Police, supra note X, at 1204.
7 SAMUEL WALKER, POPULAR JUSTICE: A HISTORY OF AMERICAN CRIMINAL JUSTICE 160.
8 SIMSON GARFINKEL, DATABASE NATION 40.
9 [restore source]
created to tabulate the census and that led to the formation of IBM.  As a punch card system that simply organized characteristics, however, the FBI system still provided very little proactive or interactive investigative capacity.

Without sophisticated computer systems, or the technological advances that would make it possible and affordable for even the most remote sheriff’s office to have a globally-connected desktop computer, databases as a major innovation for law enforcement remained largely an elusive dream. The development of the mainframe computer in 1946 and the replacement of punch cards with magnetic tape significantly advanced databasing possibilities, but it was still a primitive technology. With specific regard to law enforcement, the fingerprint database continued to mature and develop, but as late as 1984 it was still depended primarily on manual recording and retrieval. As described in the next section, the automated fingerprint identification system, or AFIS, first went online in 1985. Prior to that, linking two fingerprints required manual comparison of an unknown scene sample with, for instance, the 23 million criminal cards on file with the FBI.

The 1980s, however, initiated a period of rapid change. Personal computers became commonly available. Law enforcement began to recognize and harness the potential of electronic storage and retrieval. Around the same time, the internet was born. Connectivity became possible in ways previously unimagined, and storage capacity reached new heights. The foundations for the modern criminal justice databases had been set.

---

11 GARFINKEL, supra note X, at 18.
13 http://www.fbi.gov/hq/cjis/iafis.htm (describing how “[j]ust a few years ago, … fingerprint cards had to be physically transported and processed,” and thus a check “could often take three months to complete”).
14 GARFINKEL, supra note X, at 45.
15 GARFINKEL, supra note X, at 44.
B. Databases Today.

The past ten to twenty years have witnessed an explosion in criminal justice databasing. Although public records for law enforcement purposes have been kept for over a century, their manner of storage rendered them of little actual utility. Most were paper files or cards that were stored in diffused, and at times difficult to access, locations. An act as simple as switching locations might effectively obscure identity, since accessing a record created and stored even in the next town over could be uncommon and unlikely. Proactive searching was likewise difficult, because technology could not automatically sift through huge volumes of standardized material. Thus, most records served a largely confirmatory function – if law enforcement had a suspect and a specific record, then a direct comparison could be undertaken. Thus, although it is true that the government has long collected a variety of information from individuals, it is imperative to remember that it has really only been in the past ten or so years that this information has become widely useful and easily accessible.

Today, the largest single biometric database in the world is the fingerprint system operated by the federal government.17 The first major advance occurred in the mid-1980s, when technology simplified the creation of digital images from physical fingerprint cards. In 1985, AFIS software allowed a technician to conduct both confirmatory matches as well as automated comparisons between known and unknown prints, a landmark advance over the old manual methods of comparison. A 1987 report on the creation of AFIS noted that a San Francisco investigation had expended thousands of man-hours searching fingerprint records for the one clear print an alleged killer had left behind – when that print became the subject of San

---

16 SOLOVE, SUPRA note X, at 15.
Francisco’s first automated search, a match was found in six minutes. But AFIS still relied upon manual scanning of a card ink stain image, and thus that slowed response time and transmission capacity. The imperfection of the images also diminished automated matching capacity.

In July of 1999, the Integrated Automated Fingerprint Identification System, or IAFIS, became operational. It replaced the old ink-and-card system with a new, immediate digital image that could be easily stored, transmitted, and searched. Rather than locate a physical card and conduct a visual comparison, or even – as with the precursor AFIS system – electronically search digital images of physical cards, IAFIS kept an actual direct digital image of the print itself in high enough resolution to record and classify characteristics. By allowing any remote location to immediately upload or retrieve direct records, IAFIS transformed the significance of the FBI’s cache of information. Sharing and searching functions were dramatically enhanced. As it currently stands, IAFIS contains the records of over 55 million subjects in its Master File, which can be automatically searched 24 hours a day every day of the year, and that number grows daily.

As a companion to AFIS, and later IAFIS, the FBI in 1983 created the Interstate Identification Index (III), a pointer index which linked states to criminal records of both arrest and conviction that are associated with uploaded fingerprints. The Index is operated through the National Criminal Information Center, or NCIC, which has served since the late 1960s as the central federal criminal records database. Around its creation in 1967, it served roughly 2

18 GARFINKEL, SUPRA note X, at 45.
22 Congress had authorized the collection of criminal records and fingerprints as early as 1924, 43 Stat. 217, but it did not computerize them centrally until after the midcentury.
million requests annually; that number has now climbed to 2.5 million per day. In July of 1999, however, the FBI implemented a new generation of NCIC technology, which represented a series of advances over the prior system. New features included “the addition of image processing (i.e., mugshots, signatures, and identifying marks); automated single-finger fingerprint matching; and information linking, which provides the ability to associate logically related records across NCIC files for the same criminal or the same crime.” Most importantly, the new system “automates functions that employees previously had to perform manually,” including collection and evaluation of the system.

The NCIC consists of a composite of many sub-databases. For instance, in support of the 1993 Brady Handgun Violence Protection Act (popularly known as the Brady Act), which necessitated the creation of a reliable national background check system, Congress allocated funds to help update the over seventy million criminal records held by the fifty states and the District of Columbia. As of 1998, that goal had largely been achieved, and a rapid and quick national records system, called the National Instant Criminal Background Check System (NCIS) had been created. The NCIC 2000 system also added the Convicted Sexual Offender Registry and the Convicted Persons on Supervised Release database, in addition to standards such as the outstanding warrant files.

In addition, around the same time that IAFIS went online, the FBI launched CODIS, the combined DNA index system. This database is essential a pointer system to DNA profiles typed

---

24 Id. at 13.
25 Id.
26 Id. at 180.
27 Id.
28 Hitt, supra note X, at 14.
by state and federal agents, and it currently contains over 7.5 million offender known profiles and almost 300,000 forensic unknown profiles.  

But IAFIS and CODIS are simply the tip of the iceberg. It was estimated in 2001 that federal agencies and departments today maintain roughly 2000 databases. Those databases cover a wide variety of topics, ranging from those directly related to criminal justice purposes to those applicable only in the most specialized circumstances. There are gang databases, terrorist watch lists, violent criminal databases, forensic reference databases, corrections databases, and a wide variety of public and private databases ranging from security industry to identity theft to gaming industry databases. As a tool for criminal justice, among other things, it seems clear that the database is here to stay.

C. Databases Tomorrow.

In the wake of the September 11, 2001 terrorist attacks, and in response to congressional and executive mandates to facilitate information sharing among and within the criminal justice and national security communities, a great deal of attention has surrounded the creation of communication structures across various state and federal entities. Most notably, forty-three “fusion centers” across the nation have received roughly $380 million dollars in federal grants. These centers integrate federal, state, and local officials under one roof, and although their initial

---

34 Robin Bowen & Jessica Schneider, Forensic Databases: Paint, Shoe Prints, and Beyond, NIJ Journal No. 258 (Oct. 2007).
mandate was to focus specifically on terrorism-related investigations, “they have increasingly gravitated toward an all-crimes and even broader all-hazards approach.”

Fusion centers rely intensively on information gathered by other entities. That is, they execute their mission in part by accessing and coordinating an extraordinary number of public and private database systems, including motor vehicle, location, criminal records, corrections, sex offender, public health, and industry databases.

According to some reports, the FBI has also announced plans to replace IAFIS with a new multimodal biometric database that will incorporate DNA, facial imaging, iris scans, palm and voice prints, and identifying information in one integrated system. This announcement follows on the wake of the REAL-ID Act of 2005, which aimed to create a de facto national identity card, and has been met with strong opposition from both states and individual privacy advocates. Again, however, these examples are simply two among many, intended only to provide some insight into the breadth of databasing likely to occur in the future. After all, at present there are even plans for a doggie DNA database.

II. THE DATABASE IN CONSTITUTIONAL CRIMINAL PROCEDURE.

A wide variety of criminal justice-related databases currently operate to house, filter, and regurgitate information -- as the foregoing (even cursory) overview attests. Any person who has witnessed the past fifteen years of technological advancement knows, without reading a law review article, that online databases have transformed modern life. Yet surprising few changes

37 Id. at i.
38 Id. at 33-34.
have occurred in actual constitutional doctrine as a result of widespread databasing. Last Term, the Court confronted the applicability of the Fourth Amendment exclusionary rule to a police database error in *Herring v. United States*. Considered a watershed moment in criminal procedure, inasmuch as the Court approved the admission of evidence acquired from an unconstitutional search motivated by a database error committed by police personnel, *Herring* offers a glimpse of the kinds of cases that the Court may need to confront across a variety of constitutional criminal procedure doctrines.

Before turning to the Court's reasoning in that case, however, it may be useful to look at previous instances in which the Court addressed – even outside of the criminal context – law enforcement-related databases, in order to see the different ways in which the Court has assessed the constitutional significance of databasing generally.

A. **Major database cases.**

The Supreme Court has directly addressed and confronted the significance of databases on only a handful of occasions, few of which were very recent. In fact, perhaps the two most significant cases issued over twenty years ago. The few remaining cases only peripherally address the significance of databasing, despite being of more recent vintage. Thus, the current

---

43 In addition to the cases discussed at length below, the Court also addressed or mentioned databases in two additional cases: Ornelas v. United States, 517 U.S. 690, 693 (1996) (referring to Narcotics and Dangerous Drugs Information System database); Hiibel v. Sixth Jud. Dist. Ct. of Nev., 542 U.S. 177, 196 (2004) (Stevens, J., dissenting) (“A name can provide the key to a broad array of information about the person, particularly in the hands of a police officer with access to a range of law enforcement databases.”).
Court’s view of databases, and how they might or might not be subject to constitutional regulation, remains largely obscured.

In 1977, in *Whalen v. Roe*, the Court confronted a state statute that ordered the creation of a central database containing the names and addresses of persons who had obtained certain prescription drugs in order to combat prescription fraud.\(^{45}\) The challengers, a group of doctors and patients, alleged that that statute violated substantive due process because it impinged upon “their interest in the nondisclosure of private information and also their interest in making important decisions independently.”\(^{46}\) The government, in turn, argued that the scheme was necessary to control the illegal market in such drugs.

In upholding the statutory scheme, the Court described the databasing process in detail. It noted that the prescriptions are written in triplicate, and one copy transported to Albany monthly.\(^{47}\) They were then sorted, coded, and logged into a computer and stored on magnetic tape. The physical forms are then kept for five-years, whereupon they are destroyed. The tapes themselves are kept in a locked cabinet, and are only run “off-line” in a single room. Only seventeen employees could access the files, and only 24 investigators could investigate. In other words, the system in place closely regulated both the physical copy of the information as well as the data storage mechanism, search method, and retrieval. The Court also observed that penalties applied to any person who willfully violated the system.\(^{48}\) In light of the myriad safeguards, the Court proclaimed that it “therefore need not, and do not, decide any question which might be presented by the unwarranted disclosure of accumulated private data whether


\(^{46}\) *Id.* at 877. Because the doctrine was still evolving, the litigants framed their interest under several constitutional provisions, *see id.* at 877 nn. 23-24, including the First, Fourth and Fourteenth amendments, *id.* at 32, but the claim sounds in due process today and that was the basis upon which the Court ruled, *id.* at 604.

\(^{47}\) *Id.* at 593-95.

\(^{48}\) *Id.* at 594 n.12.
intentional or unintentional or by a system that did not contain comparable security provisions.”

In a closing passage oft cited today, Justice Stevens, writing for the Court, declared:

We are not unaware of the threat to privacy implicit in the accumulation of vast amounts of personal information in computerized data banks or other massive government files. The collection of taxes, the distribution of welfare and social security benefits, the supervision of public health, the direction of our Armed Forces, and the enforcement of the criminal laws will all require the orderly preservation of great quantities of information, much of which is personal in character and potentially embarrassing or harmful if disclosed. The right to collect and use such data for public purposes is typically accompanied by a concomitant statutory or regulatory duty to avoid unwarranted disclosures. Recognizing that in some circumstances that duty arguably has its roots in the Constitution…[the New York statutory scheme adequately protects privacy.]

Thus, the Court in 1977 not only contemplated the databasing currently underway in the nation, but also anticipated the widespread databasing that would come in the future. Both it deemed apparently necessary, although potentially threatening to personal privacy. Most intriguingly, the Court seemed unconvinced by, although open to, the idea that the Constitution would regulate the compilation of state databases. Instead, it seemed to view the enabling statutes as the proper source of primary regulation – a proposition that assumes that databases come about only as a result of statutory mandate.

In contrast, Justice Brennan, concurring, explicitly stated that “[b]road dissemination by state officials of such information… would clearly implicate constitutionally protected privacy rights.” Justice Brennan also specifically addressed the significance of databasing itself. He identified the troubling aspect of the scheme as “the central computer storage of the data thus

49 Id. at 606.
50 Id. at 605-06.
51 Id. at 606 (Brennan, J., concurring). The other concurer, Justice Stewart, sharply disputed this claim, criticizing its citation to Griswold v. Connecticut, 381 U.S. 479, and Stanley v. Georgia, 394 U.S. 557, as inadequate support for the proposition. Id. at 608 (Stewart, J., concurring).
collected.”

Although he deemed it “[o]bvious[]” that “collection and storage of data…is not rendered unconstitutional simply because new technology makes the State’s operations more efficient,” he recognized that the “central storage and easy accessibility of computerized data vastly increase[s] the potential for abuse of that information…” Citing, interestingly, to the Fourth Amendment as an example of a means restriction on information-gathering (as opposed to simple kind-restrictions), Justice Stevens remarked that “future developments” may “demonstrate the necessity of some curb on such technology.”

Depending on one’s perspective, the Court either made good on its promise or tacked the other direction ten years later in its next major judgment concerning large scale databases, United States Department of Justice v. Reporters Committee for Freedom of the Press. In Reporters Committee, the Court addressed a claim by a reporter and journalists’ association for access under the Freedom of Information Act (FOIA) to FBI “rap sheets.” The FBI had compiled “rap sheets” that contained the name, birth date, physical characteristics and criminal records (including arrest, charges, convictions, and incarcerations) of over 24 million individuals. The challenge related to the denial of a request by a reporter who sought access to the records of members of a reputed mafia family under the federal FOIA.

As in Whalen, the focus of the legal analysis was largely statutory. The enabling statutes that authorized the creation of the rap sheets had been amended to limit disclosure, and the Department of Justice had a policy of generally treating them as confidential, except for

---

52 Id. at 606.
53 Id. at 607.
54 Id. at 607.
56 Id. at 751.
57 Id. at 751.
58 Id. at 757.
disclosure to the subject of the sheet or for press releases for wanted persons.\textsuperscript{59} The FOIA, in contrast, sets forth a presumption of disclosure of agency records.\textsuperscript{60} However, three exemptions arguably required the DOJ to withhold the rap sheets: Exemption 3, requiring compliance with other statutory nondisclosure provisions; Exemption 6, protecting certain files that could create “a clearly unwarranted invasion of privacy”; and Exemption 7(C), applying to law enforcement records that would likewise invade privacy.\textsuperscript{61}

Again writing for the Court, Justice Stevens held that Exemption 7(C) prohibited the disclosure of the rap sheet. Interestingly, the Court first had to acknowledge the claimed invasion of privacy was to material that, technically speaking, consisted of almost entirely public information. Indeed, the reporter specifically sought only that information that constituted “matters of public record.”\textsuperscript{62} Of course, the Court noted that a few states prohibited disclosure of out-of-state information, or restricted the conditions under which non-conviction information would be disclosed.\textsuperscript{63} A larger number had restrictions on issuing criminal history summaries.\textsuperscript{64} But individually, each piece of information was almost always technically open and available as a matter of public record.\textsuperscript{65} Thus, in order to find that the rap sheet implicated any privacy interest, the Court had to acknowledge that “much rap-sheet information is a matter of public-record,” and specifically draw a distinction between the technical availability of the information

\textsuperscript{59} Id. at 752. The statute had further been amended to permit disclosure of rap sheets to banking institutions, securities industry regulators, and the Nuclear Regulatory Commission. Id. at 753.

\textsuperscript{60} Id. at 755.

\textsuperscript{61} Id. at 755-56 (citing 5 U.S.C. § 552(b)(3), (6), (7)(C)).

\textsuperscript{62} Id. at 757.

\textsuperscript{63} Id. at 753-54 & n.2.

\textsuperscript{64} Id. at 753.

\textsuperscript{65} Id. at 753.
and its actual availability ("practical obscurity," in the Court’s parlance), which it described as "limited." 

Rejecting a "cramped notion of personal privacy," Justice Stevens observed that "[i]n an organized society, there are few facts that are not at one time or another divulged to another." Yet such disclosures did not render that information no longer private, given the passage of time and the practical impediments to broad dissemination. Noting that the "very fact that federal funds have been spent to prepare, index, and maintain these criminal-history files demonstrates that the individual items of information . . . would not otherwise be ‘freely available,’" Justice Stevens distinguished between "scattered disclosure" of bits of information and complete and total revelation. 

While the gravamen of the Court’s analysis rested on a statutory analysis of the provisions involved and the propriety of issuing a categorical rule, the manner in which it assessed the interests implicated by disclosure granted some insight into its view of databasing generally. The Court noted that the Privacy Act had been enacted specifically to counter concerns about "computer data banks on individual privacy." In the Court’s estimation, electronic storage altered the nature of the interest affected, in that "the computer can accumulate and store information that would otherwise have surely been forgotten long before...." By

---

66 Id. at 762.
67 Id. at 753.
68 Id. at 762.
69 Id. at 763. Justice Stevens identified the interest as in “selective” versus total nondisclosure. Id. at 763 n.14 (citing Karst, “The Files”: Legal Controls Over the Accuracy and Accessibility of Stored Personal Data, 31 L. & Contemp. Probs. 342, 343-44 (1966)).
70 Id. at 763-64.
71 Id. at 779-80.
72 Id. at 766-67 (citing H.R. Rep. No. 93-1416, p. 7 (1974)). The Court also cited to Department of Air Force v. Rose, 425 U.S. 352 (1976), a sort of proto-database case in which law students requested records of disciplined cadets. Despite the publicity of some of the cadets’ names, the Court noted that the passage of time and informality of the public knowledge continued its character as “private” sufficiently to deny disclosure under FOIA. Id. at 768-69.
73 Id. at 771.
holding the records unobtainable under the law enforcement exemption, while in no way casting aspersions on the statutory mandate to centralize them in the first place, the Court seemed to say that such long memories were properly the province of the state alone.

The holding in Reporters Committee contrasts remarkably with that in Smith v. Doe, in which the Court upheld the Alaska Sexual Offender Registration Act against an Ex Post Facto constitutional challenge.74 As in Reporters Committee, the case involved a statute that authorized the compilation of publicly available information – here, convictions for sexual offenses.75 Unlike Reporters Committee, however, the purpose of the statute was public dissemination.76 Finding that the statute was civil and regulatory, as opposed to punitive, in character, the Court found no violation of the constitutional bar against retroactive punishments.77

Most relevant for this Article, however, were the Court’s treatment of the arguments raised by the respondents. The respondents analogized the creation of a centralized public database available online to the public stigma and shaming of the pillories and whippings of old, in order to liken the Act’s requirements to “punishment.”78 Rejecting that argument, however, the Court refused to consider mere “dissemination of information” or the publication of truthful, formally available information a punishment.79 Writing for the Court, Justice Kennedy distinguished the physical, corporeal “direct confrontation” from the “dissemination of accurate information about a criminal record, most of which is already public.”80 Moreover, to the extent

---

74 538 U.S. 84 (2002). Smith was also decided with a companion case, Connecticut Dep’t of Public Safety v. Doe, 538 U.S. 1 (2003).
75 Id. at 89.
76 Id. at 93.
77 Id. at 92-93. For a discussion of the civil-punitive divide as regards this case, see generally Murphy, Paradigms, supra note X, at 1351-58.
78 Id. at 98-99.
79 Id. at 98-99.
80 Id. at 98-99.
that “adverse consequences” flowed from that dissemination, such results were not the direct purpose of the regulatory scheme.\footnote{Id. at 99.}

Thus, unlike the Court in both Reporters Committee and Whalen, in Smith the Court seemed indifferent to the particular implications of widespread and ready distribution of information, choosing instead to focus on its technical availability. That point is particularly salient with regard to the Court’s dismissal of the impact of the internet publication, which it said served only to “increase[] in proportion” the degree of humiliation.\footnote{Id. at 99.} The corresponding ease and breadth of disclosure was wholly irrelevant to the analysis; the Court likened the online registry to “a visit to an official archive of criminal records,” which was simply “more efficient, cost effective, and convenient for Alaska’s citizenry” by virtue of being online.\footnote{Id. at 99-100; see also id. at 105 (further dismissing the significance of the breadth of disclosure).} The Court later identified the negative consequences of the statute as flowing from “the fact of conviction, already a matter of public record” rather than from the broad disclosure and broadcast provisions of the statute.\footnote{Id. at 101.} In sum, unlike the analysis in Reporters Committee and Whalen, the Court in Smith seemed wholly uninterested in the significance of databasing as an act that changes the nature of information in terms of its meaning and character. Rather, the database was simply another form of the same information, with no particular significance or power merely on account of its technical capacity.

However, two justices maintained that the compilation and ready availability of the information constituted a significant change from its prior paper-and-archive existence.\footnote{Justice Ginburg, writing in dissent for herself and Justice Breyer, found the Act unconstitutionally punitive in an opinion that did not specially address the significance of the creation of a centralized database of information.} Justice Souter, concurring in the judgment, disputed that the Act “simply makes public information

\footnote{Id. at 99.}
available in a new way,” noting that “[i]ts point, after all, is to send a message that probably
would not otherwise be heard.” Justice Stevens likewise concluded that “there can be no doubt
that ‘[w]idespread public access]…to this personal and constantly updated information has a
severed stigmatizing effect.” Overall, however, Smith, marked a fairly sharp departure from
Reporters Committee and Whalen in its seeming indifference to arguments based on the special
nature or character of information once collected in centralized databases.

B. Evans and Herring.

The two most significant criminal procedure cases related to databasing occurred more
recently, in the context of Fourth Amendment challenges to evidence seized as a result of
erroneous information from a database. Both Arizona v. Evans, decided in 1994, and Herring
v. United States, decided in 2009, held the exclusionary rule inapplicable to the fruits of
unconstitutional searches conducted in reasonable reliance on a database entry later proven to
have been inaccurate. In Evans, the error stemmed from the failure of a judicial court clerk to
clear a quashed warrant from a computer database; in Herring, the same mistake occurred, but
this time under the watch of a police department clerk. The Court in both cases ruled the
exclusionary rule inapplicable, citing deterrence as the primary purpose for the rule and the
absence of a deterrence rationale for reasonable record-keeping errors.

In both cases, the Court briefly discussed the procedures used to enter and clear warrants
in the computer, although in both cases those procedures relied upon factual development, as

86 Id. at 109 (Souter, J., concurring).
87 Id. at 111-12 (Stevens, J., concurring) (internal citation omitted).
89 129 S. Ct. 695 (2009).
90 Evans, 514 U.S. at 4.
91 Herring, 129 S. Ct. at 698-99.
92 514 U.S. at 15-16; 129 S. Ct. at 703-04.
opposed to statutory or regulatory analysis.\textsuperscript{93} In both cases, the Court mentioned the error rate for the database in question, although in both cases that error rate was the subject of dispute.\textsuperscript{94} And in both cases, the justices making up the majority paused to consider the particular implications of the fact that a database was the source of error, even while ultimately dismissing those implications for purposes of resolving the instant case.

Specifically, in \textit{Evans} three concurring justices (Justices O’Connor, Souter, and Breyer) expressed considerable concern about the nature of the error as a “recordkeeping error” that occurred as a result of departure from established protocol.\textsuperscript{95} They specifically worried about databases -- “the advent of powerful, computer-based recordkeeping systems that facilitate arrests in ways that have never before been possible.”\textsuperscript{96} Cautioning that such technology, while beneficial, ought not to be relied upon “blindly,” they warned that reliance on a system with “no mechanism to ensure its accuracy over time” would invoke the exclusionary rule.\textsuperscript{97}

Similarly, in \textit{Herring}, Chief Justice Roberts stated that a showing that the police were “reckless in maintaining a warrant system,” or had “knowingly made false entries to lay the groundwork for future false arrests” would invoke exclusionary rule protection.\textsuperscript{98} Demonstrated “systemic errors” in a system might also render reliance on that system objectively unreasonable.\textsuperscript{99} Absent evidence – by implication, offered by the defendant challenging the action – that a system had “routine or widespread” deficiencies, however, the evidence seized by

\textsuperscript{93} 514 U.S. at 5 (describing testimony); 129 S. Ct. at 698.
\textsuperscript{94} Compare 514 U.S. at 15 (citing testimony that “this type of error occurred once every three or four years”); 129 S. Ct. at 704 (citing testimony that no other errors of this kind had ever occurred), with 514 U.S. at 27-28 (Ginsburg, J., dissenting) (observing both that the witness later admitted to three of the same kind of errors that day, and also recounting prominent cases of similar error); 129 S. Ct. at 700 n.1 (debating characterization of error as “negligent” or not); id. at 704 & n.5 (noting dispute in testimony about whether the error occurred with greater frequency than suggested).
\textsuperscript{95} 514 U.S. at 16-17 (O’Connor, J., concurring).
\textsuperscript{96} Id. at 17.
\textsuperscript{97} Id. at 17.
\textsuperscript{98} 129 S. Ct. at 703.
\textsuperscript{99} 129 S. Ct. at 704.
an officer in reliance on information gleaned from an erroneous database entry would be admissible.

The dissenters in *Evans* each focused particularly on the specific implications of database error. Justice Stevens observed that the exclusionary rule operated to deter “systemically,” not just on an individual officer level. He chastised the Court for “overlook[ing] the reality that computer technology has changed the nature of threats to citizens’ privacy over the past half century.” Justice Stevens also refused to presume that the database was kept in an orderly fashion, calling the testimony “slim evidence on which to base a conclusion that computer error poses no appreciable threat to Fourth Amendment interests.” Lastly, he conjured the spectre of erroneous arrests attributable to “some bureaucrat” who “has failed to maintain an accurate computer database.”

Justice Ginsburg, dissenting separately and joined by Justice Stevens, opened by warning about the “evolving problem” of “the increasing use of computer technology in law enforcement.” She went on to discuss a theme that she picked up again in her dissent in *Herring*, noting that the lower courts found Supreme Court precedent “not helpful,” in resolving the issues. Justice Ginsburg agreed that the problem was not just “a court

---

100 514 U.S. at 21 (Stevens, J., dissenting). In fairness, Justice Stevens’ observation was more motivated by the systemic relationship between the law enforcement and judicial branches, as opposed to the many operators within a law enforcement (or database) system.
101 Id. at 22.
102 Id. at 22.
103 Id. at 23.
104 Id. at 23 (Ginsburg, J., dissenting).
105 The *Herring* dissenters largely mirrored the justices that either dissented in *Evans* or else concurred with specific reservations related to database entry. Justice O’Connor, who had joined the concurers in *Evans*, had left the Court by the time that *Herring* was decided.
106 Id. at 24.
employee’s slip” but rather the “potential for Orwellian mischief” in the government’s increasing reliance on computer technology in law enforcement.\textsuperscript{107}

In an opinion thoroughly consumed with assessing the impact of these new technologies, Justice Ginsburg asserted that “[w]idespread reliance on computers to store and convey information generates, along with manifold benefits, new possibilities of error….”\textsuperscript{108} Moreover, “computerization greatly amplifies an error’s effect,” because the error can seep out across a wide variety of systems.\textsuperscript{109} Citing examples of harassing and dangerous arrests based on erroneous computer data,\textsuperscript{110} Justice Ginsburg worried about leaving such databases constitutionally unregulated absent a showing of deliberate error.

Her dissent, this time on behalf of Justices Stevens, Souter, and Breyer, echoed the same themes. Rather than accept the database on its face and assume it optimally operative, she elaborated the many shortcomings of its current configuration.\textsuperscript{111} For example, she cited the lack of a obvious electronic connections between information terminals, the entirely uncontrolled procedures for inputting and extracting information, and the lack of routine oversight or maintenance for quality control purposes.\textsuperscript{112} By comparison, she noted that the national database system, the NCIC, had numerous safeguards in place both to control the quality of information stored and to minimize the adverse consequences of error.\textsuperscript{113} In short, Justice Ginsburg

\textsuperscript{107} Id. at 25.
\textsuperscript{108} Id. at 26.
\textsuperscript{109} Id. at 26.
\textsuperscript{110} Id. at 27
\textsuperscript{111} 129 S. Ct. at 708-09 (Ginsburg, J., dissenting).
\textsuperscript{112} Id. at 709.
\textsuperscript{113} Id. at 709.
demanded some evidence of good database management, rather than trust that such practices were in place absent evidence to the contrary.\textsuperscript{114}

In sum, what is most evident from the foregoing discussion of the Court’s treatment of cases involving computerized criminal justice records is its overall failure to articulate a consistent vision of the constitutional significance of the aggregation and easy retrieval and dissemination of otherwise lawfully held – and even publicly available – information. Many of the cases were largely driven by statutory analyses, and yet the degree to which the Court undertook its own assessment of the manner in which centralization and digitalization of records transformed the significance of those records seemed to vary case-by-case. To be sure, much of that variation may be attributable to political or ideological differences among justices and courts over time. But it may also be that databasing is sufficiently new, and sufficiently untested, that a template for a more sensitive inquiry has yet to be developed.

\textbf{III. THE SLIPPAGE BETWEEN DOCTRINE AND DATABASES.}

When considered carefully, several background presumptions emerge as recurring themes – at times gaining a majority, although at other times not -- within the Court’s cases. Although not expressly the study of this Article, these themes also arise in the lower courts that address widescale databasing. Yet courts rarely consider database cases as a specific sub-genre of case or problem requiring specialized attention. The notion that there are “database cases” or commonalities amongst database issues the way that other familiar criminal justice investigative methods or techniques collect in doctrinal groups – say, “informant cases” or “confession cases”

\textsuperscript{114} \textit{Id.} at 709-10. Indeed, she worried both that it would be difficult to identify the source of error in many cases, and that defendants (especially indigent ones) would have trouble obtaining the information necessary to make such a showing. \textit{Id.}
– has not yet taken shape.\textsuperscript{115} Instead, database cases seem to benefit from certain presumptions that help to obscure their common, often unique characteristics or any specialized scrutiny that they might receive. The Court’s failure to apprehend these presumptions, however, exacerbates the disconnect between the constitutional doctrines that regulate police behavior and the reality of how widescale databasing works.

\subsection{A. The Presumption of Ready Analogy.}

A strong inclination exists to ignore the ways in which databases transform information, and instead characterize them as identical to, or at most as simply more efficient forms of, other kinds of information collection. Typical is the assertion in \textit{Whalen} that the centralized collection of prescription records for databasing was “not significantly different” than that which existed when prescriptions were subject to other forms of limited disclosure.\textsuperscript{116} The Court in \textit{Smith} similarly analogized the collection and dissemination of conviction records online to a simply more efficient form of historical archiving.\textsuperscript{117}

I have elsewhere written at length about the dangers of viewing technological incursions onto privacy as significant only inasmuch as they mirror physical ones,\textsuperscript{118} and so will not elaborate on that assertion here. What is significant, however, is that the reliance on a non-technological precursors to legitimate technology-based regimes serves to excuse courts from having to seriously contemplate both the actual steps required to construct and operate a database, as well as the new or innovative ways in which information may be used as a result. If an electronic or online repository of convictions is considered identical to a paper one, then

\begin{thebibliography}{11}
\bibitem{115} But see \textit{Evans}, 514 U.S. at 17 (O’Connor, J., concurring) (likening desire to see standards for databases to similar approach to informants).
\bibitem{116} \textit{Whalen}, 429 U.S. at 602.
\bibitem{117} \textit{Smith}, 538 U.S. at 98-100.
\bibitem{118} See Murphy, \textit{Paradigms, supra} note X, at X.
\end{thebibliography}
courts need not inquire into the complicated questions of algorithms or file maintenance or malicious alteration or interception.

Moreover, databases may simply gain authority by association: if law enforcement has kept fingerprints of criminals for a hundred years, why should any court revisit their continued preservation? Never mind that, for most of those hundred years, the dusty and distintegrating fingerprint cards existed in thousands of drawers in hundreds of decentralized sheriff’s offices, whereas now they will be instantly accessible and electronically searchable anywhere in the world. The utility of the analogy is that it bypasses the need for careful consideration of the details of databasing – the nitty gritty operational information about inputs and outputs that would require careful reconsideration and possible intervention in areas of otherwise longstanding familiarity.

**B. The Presumption of Demonstrable Harm.**

Similarly, databases benefit from the general requirement that constitutional violations allege actual and not speculative harms. The moment of confrontation with a database system almost always is divorced from its optimal moment of control. The harm most easily articulated from databasing is by its nature speculative, and applicable to a range of individuals as innumerable as the database’s capacity itself. But it is incredibly unwieldy to recognize the mere risk of harm as a worthwhile target of constitutional attention, or to define parameters for the minimization of such risks in the absence of specific instances of injury.

In addition, when harm does result, the diffused and decentralized nature of databases may make correction difficult. Back-up records may have to be obtained and destroyed. Inaccurate or obsolete information may have spread virally throughout networks and be virtually
impossible to recall. Indeed, a database error, like the database itself, can often more conceptual
than real – it may less be about a server with a bunch of ones and zeros that require erasing than
it is about a patchwork of information let loose in a web of systems. Once the information is
released it carries its own momentum, and eradicating it from every system may simply not be
possible.

C. The Presumption of Regularity.

Databases also appear to benefit from the strong presumption of regularity that attaches
to most law enforcement actions. This presumption, as applied to the behavior of individual
officers, tends to spring partly from an institutional desire to defer to the expertise of the
executive branch in conducting investigations, and partly from the recognition that assessing the
subjective motivations of individual persons is difficult at best. But even though no equivalent
functional or institutional justification impels the same deference when it comes to databases,
there nevertheless exists the tendency to exercise it.

As regards databases, the presumption of regularity means that, absent affirmative
evidence that a database is kept in a shoddy or substandard fashion, courts will assume that the
soundness of the information generated. Notably, this presumption seems to hold even when
information about the procedures or practices governing the collection and maintenance of the
database are lacking,119 either because the record was not developed factually or because the
oversight structures for the database are entirely informal. Thus, unlike areas of constitutional
criminal procedure in which articulation of formal restraints, and compliance with them, dictates
the acceptability of a particular action,120 database cases can be decided largely ad hoc.

119 See, e.g. Herring, 129 S. Ct. at 2 (“For whatever reason, the information about the recall…”).
120 See, e.g., [warrants, wiretaps, administrative searches, etc.]
As a corollary to this presumption of regularity, databases also seem to benefit from the fact they typically spring from organic enabling statutes. Courts can therefore rely upon a statutory regime to guide and govern review of the database, rather than grasp at constitutional doctrine. At the same time, administrative compliance with statutory provisions about the parameters of the database may constitute a form of shield against claims that the database is impeding or contravening constitutional interests.

D. The Presumption of Individual Action.

Relatedly, databases can benefit from the difficulty that courts encounter in fitting the actual activity of databasing into conventional molds for police behavior. A database is typically not the product of any single individual actor; it may not even be the product of a single police force. Rather, databases are the ultimate collaborative projects. They may require myriad inputs and enable equally diverse outputs. They may cross jurisdictional borders, or require public-private sector cooperation.

By way of example, consider a database as simple even as the one raised in *Whalen*. The database software must be created by computer scientists, and the hardware kept and operated by technologists. Doctors create the source information, which must be transferred to a variety of state officials responsible for inputting it into the machine. The hard copies of the information must then be tended and eventually destroyed. The physical storage mechanism – there, magnetic tapes – must be kept and safeguarded, and then perhaps operated and searched when appropriate. They too must be regularly maintained to ensure compliance with regulatory restrictions, and destruction of data when appropriate. Throughout all of this, technology may evolve such that the magnetic tape or machines or the software becomes outdated or obsolete.
And that is just one state’s simple prescription registry, accessible to only a limited number of state officials and operated (the Court took great pains to state) out of one room in Albany. A diffused networked system of information like NCIC or the DNA databank or the sex offender registries, therefore, raise a whole new set of complicated problems. In sum, a “database” is often a misleadingly singular concept that in fact embodies multifold layers of individuals and objects, and spans both geographical and temporal boundaries.

E. The Presumption of Technological Neutrality.

As a companion to this presumption of regularity, databases also remained cloaked in a powerful technological neutrality. The symbolic and actual manifestation of the database as a computerized technology helps neutralize it. Disembodied from the human beings that define, create, realize, and benefit from its parameters, the database is easily viewed as incapable of bias in the way that human law enforcement agents might be.

That is, although courts of course recognize that a human being might create false entries in a database or intentionally manipulate information, the underlying architecture of the database is largely ignored as a human event. It has long been a familiar cautionary refrain that police officers are engaged in the “competitive enterprise of ferreting out crime” and thus might act with understandable, but impermissible, zeal. But no companion notion exists with regard to databases: we do not conceive of databases as themselves acting “zealously.”

Yet databases mirror and replicate human predispositions, and they cannot operate wholly divorced from them. What information is selected for input, how it is input, how it will be searchable – all of these values are defined by humans and therefore replicate human
judgments about the relative importance and ordering of information. The database cannot communicate wholly neutrally: it always functions in the shadow of the human hand. Indeed, the way in which information is stored and retrieved may itself communicate predilections and biases. For example, DNA match probabilities are expressly quantified in racial terms that relate to individual self-identification rather than actual biological ancestry. In short, databases are essentially human. Despite the patina of technological perfection, databases are ultimately compilations of human knowledge – created, maintained, and used by humans. It is flawed human beings that collect their information, write their operating codes, input their entries, maintain their systems, and search and retrieve their data. At base, therefore, databases constitute turbo-charged knowledge – still human, just more powerful.

IV. TOWARD A CONSTITUTIONAL CRIMINAL PROCEDURE OF DATABASES.

If these erroneous presumptions represent pits into which constitutional analysis is prone to fall when confronted with database technology, then what considerations might instead better govern or shape constitutional doctrine in this area? This Part offers five characteristics that define databases, and which merit acknowledgment in any honest constitutional inquiry. In outlining these qualities, my goal is to encourage direct thinking about databases, both in how they might dispense with some of the conventional concerns that guide constitutional doctrine as well as raise new concerns as yet unrecognized. Supplanting faulty stereotypical presumptions with descriptively accurate ones will hopefully facilitate development of doctrine that guides the formulation of appropriate constraints on database use and deployment, while eliminating constraints that unnecessarily impede it. The ideal is to acknowledge the ways in which
databases represent a new form of collection, use, and dissemination of information, and capitalize on those strengths while minimizing their weaknesses.

By way of clarification, my aim is more to think about the meaning of databases than the meaning of constitutional doctrine. That is, the qualities identified below to some extent transcend the particular constitutional doctrines at issue. Criminal justice database challenges can arise under a variety of constitutional clauses – including most obviously Fourth Amendment search and seizure, Fifth Amendment due process, and Sixth Amendment confrontation, compulsory process, and right to counsel. But for purposes of this paper, the specific contours of particular doctrine is less important than the shared features that inhere across databases generally, and which might help illuminate constitutional questions regardless of the particular claim at issue.

A. Structural (vs. individual) oversight.

Perhaps the most singular trait that differentiates databases is that they require structural, rather than individual, oversight. That is, constitutional doctrines in criminal justice tend to particularize – they assume individual actors and individual actions. The archetypal criminal investigation is of the individual criminal, committing an individual crime (or series of them), investigated by individual officers. Thus the system is structured accordingly: its restrains individual police officers through the Fourth Amendment, or respects individual rights through the Fifth and Sixth Amendments, and adjudicates individual claims through the Sixth Amendment processes.

Even to the extent that constitutional criminal procedure wrangles with entities, it does so awkwardly. Courts often reach for analogies to the individual or to personalization of a
collective rather than wholly alter the doctrinal approach. For instance, corporate liability simply substitutes a single, fictional corporate entity for an individual one. Or doctrines of racketeering liability serve to penalize groups of individuals by imputing liability for the actions taken by others with whom they associate. We still depend on the constitution to regulate investigations by regulating the actions of individual officers (through search and seizure and interrogation), and to regulate adjudication by inquiring of evidence related to an individual case (through case-specific discovery, cross-examination, assistance of counsel, and burdens of proof or evidentiary rules).

Databases, however, require a different tactic. Investigations of one person cannot be neatly disaggregated from that of others in the database; nor can adjudication be relied upon to expose shortcomings or flaws in the collection of evidence. Databases are rarely the product of one individual’s action, and rarely contain easily separable individual information. Instead, they tend to be the product of numerous actors and inputs, and collate numerous tiers of information. Think about *Herring*: an anonymous person put in the erroneous information, or else failed to remove it; then the information was accessed by one clerk, and transmitted to another who failed to undertake any steps to verify it. It is difficult, and maybe impossible, to identify the moment the error occurred, or the individual who perpetrated it. It is likewise futile to attempt to regulate databases with reference to only one constituent part, or one discrete moment of constitutional significance. The database’s very purpose is to derive information from myriad inputs over a stretch of time.

Take, for example, the DNA database. The DNA database is in itself largely a fiction; even the name of the federal database, CODIS, reveals as much. CODIS, or the Combined DNA Index System, in fact refers not to a central repository of information, but rather to the software
used by the individual law enforcement entities that have met the standards and entered into an agreement to share data. Each local or state entity uploads basic information to a centralized repository, and automated or intentional searches then indicate matches that can be pursued by contacting the uploading agency. Thus, to the extent that CODIS even exists, it incarnates as a pointer system – it tells a user where to look for the source information to which they have generated a match.

Moreover, the stored information itself is a product of a chain of information generation: the chemical and mechanical technologies required to type and analyze a genetic sample, the analyst who must interpret and enter the data, and the engineers that write the software code and maintain and superintend the databases themselves. Within each category are subcategories of data technicians, analysts, and law enforcement agents, each with a particular role to play that is essential to the creation and maintenance of the database. And this description only scratches the surface; the DNA database also depends upon other databases in order to draw meaning. Without sub-databases of profiles collected to gauge population frequencies, for instance, it is impossible to state with certainty what the significance of a DNA match is at all.

Thus, to the extent that oversight of the integrity of such databases is to occur, or that constitutional doctrine intends to regulate their formation and use, any such oversight must inevitably operate structurally, not individually. Yet doctrine is currently ill-suited to such a task. To the extent that there is any notion of database oversight, it ill-accommodates this kind of structural, overarching approach. An individual model focuses on what happens in a particular case, and which individuals interfaced with the database in that instance (who loaded a sample, or ran a search, for example), without considering that person within the larger web of inputs and outputs necessary to the databases’ continued operation.

---

121 Murphy, New Forensics, supra note X, at 739.
This kind of myopia is understandable, however, given the highly tailored and individualized notion of criminal justice that has typified the Anglo-American system. Indeed, some critics lament that American justice is so intensely myopic that it excludes unnecessarily even highly probative evidence of general propensities even limited to a particular individual.\textsuperscript{122} The current model of adjudication is ill suited to the task of reviewing and assessing evidence derived from databases. The procedural entitlements of the adversary system depend upon matched opposing forces engaging in a contest to resolve discrete disputed facts. Tools such as cross-examination or compulsory process cannot sweep broadly enough to disclose systemic failings.\textsuperscript{123} Discovery, compulsory process, or cross-examination in a single case yields little opportunity to identify and uncover, much less broadly correct, errors apt to occur (and be visible) only from scrutiny on a systemic level.

Any truly effective constitutional doctrine for databases will have to contend with their multifaceted and interdependent, rather than discrete and intimate, operation. A structural approach, rather than a model that presumes individual investigative or adjudicative oversight, is required. Such a shift requires a major reorientation in constitutional thinking. Rather than ask about individual actors specific behaviors – a substantive approach, in some respects – a procedural inquiry that asks about the existence of protocols and monitoring systems, would be preferred. In other words, it is arguably impossible to regulate databases \textit{substantively} – to truly inquire whether a particular series of tests or entries or searches were accurate, fair, and correct. But it is much easier to impose \textit{procedural} requirements upon databases – to inquire into the existence and thoroughness of protocols for those processes, and to presume inadequate or

\textsuperscript{123} See, e.g., Ken Strutin, Databases, E-Discovery, and Criminal Law, 15 Richmond J. L. & Tech. 6, 14-16 (2009) (discussing United States v. Dioguardi, 428 F.2d 1033 (2d Cir. 1970)).
defective any database system maintained without them. Certain structural devices are demonstrably effective in minimizing mistakes, and with greater attention, surely others would be uncovered. A constitutional doctrine that looks for the signs of good management – think scrutiny of policies for access controls, or regular audits, or blind tests – is far more likely to improve database deployments in society than one that attempts to determine whether a database has failed or not in a case-specific context.

B. Suspicionless (vs. suspicion-based) targeting.

Concomitant with recognition of the structural rather than individual nature of database systems is an acknowledgement that, as a result, databases can (and often do) yield information divorced from individual suspicion of a particular person. That is, databases can function both to identify and to confirm. In the confirmatory mode, the search for information is targeted – a user seeks particular information about a particular person, based on suspicion or other information. The information in the database simply confirms – perhaps rapidly or more comprehensively – the suspicions already held. For example, a known suspect’s records are searched for a criminal past, or a DNA match, or evidence that she crossed a bridge or made a call.

In the identification mode, in contrast, the user is more or less dispassionate about the ultimately identified individual – the database, in essence, chooses the suspect. The known information is the input, and its relationship to a criminal act; what the database supplies is a match or list of suspects – the building blocks, or final nail, of the investigation.

By way of example, think of a cell phone call record database. If an investigator seeks all the calls made from a particular phone, or set of phones, then the investigatory seeks to confirm or dispel suspicion related to the holders of those numbers. But an investigatory might also seek
the identity of all callers who contacted the homicide victim, say, or whose calls were routed through a particular tower at a particular time. In this usage, the investigator relies upon the database to identify the suspect, independent from the investigator’s own suspicions about the likeliest suspect.

The difference in these two modes of operation matters meaningfully. Where databases are used to confirm otherwise-held suspicions, they more traditionally resemble other traditional forms of police investigation. But the use of databases to generate suspects represents a new kind of investigation altogether – whether based on particular information (like “who called this number”) or upon predefined algorithms (like “who has traveled to these three countries and bought these two items within a one month period”).

Suspicion models that constrain investigation, therefore, will both over-regulate and under-regulate the use of databases in this context. The great fear with regard to this kind of databasing is, in Dan Solove’s formulation, the fear of the anonymous bureaucrat, not the malevolent inspector.124 The restraints called for will inevitably falter if they presume the need for the overzealous, overstepping constable. The true risk is a leaping-to-conclusions, or confirmation bias. It is the fear that the individual will be sucked into a morass of suspicion from which escape is arduous or impossible – Kafka’s The Trial, not Orwell’s Big Brother.125

Conversely, the need to constrain such searches due to lack of suspicion may in fact constitute unnecessary overreaching. The invasion of privacy occasioned by physical confrontations, or even informational inquiries, conducted by police against suspected individuals raises a different tenor of concern than that represented by search queries into established and regulated databases. Many of these database searches – for the DNA match, the

---

124 Solove, supra note X, at 36-42.
125 Solove, supra note X, at 31, 41-43.
license plate match, the cell phone callers – will be troubling most inasmuch as the composition of the database reflects some troubling inequities, or inasmuch as what follows next raises issues. The search in itself is not the problem, and questions like reasonable suspicion or probable cause carry little meaning.

A constitutional doctrine obsessed with suspicion standards, however, misses this point. Suspicion standards help winnow the number and degree of interests affected by law enforcement actions, but when a database search is at issue, those are the wrong metrics. As a result, however, there is too much effort to ignore that there is any import to a database search at all. Neither path is desirable. Database searches enhance privacy in that they are anonymous and suspicionless – those qualities increase the likelihood that they are executed fairly, unobtrusively, and neutrally. But that remains true only so long as the database itself is equitably composed and adequately monitored. Thus, it is these latter questions that should serve as the gates of entry to relatively unfettered use of databases, rather than to abandon any effort to regulate them at all or to attempt to impose suspicion-based models where they ill fit.

C. Operative opacity (vs. transparency).

Relatedly, databases do not just generate information anonymously; they also tend to operate anonymously, as well. At the most removed level, databases typically are shrouded in secrecy. And such secrecy is usually desirable – the mere effort required to compile and organize the information they contain represents a major investment. Moreover, that information often relates to sensitive topics, whether it is the criminal history of an individual or the fact that they are suspected of some illicit activity. Generally speaking, the secrecy of databases is desirable – this is precisely why the possibility of “hacking” databases generates such loud
concerns, and why privacy experts demand myriad safeguards on the integrity of the information they keep.

Databases also are often, by their nature, also secret from within. The code that makes them run, the operators that input data, and even the persons who access that data may all interact with the information in a nameless, faceless way. Unlike the beat officer or public prosecutor, the data technician need not necessarily have as personal a stake in the construction of the database or the outcome of a search. Databases can be created, maintained, and used by fungible technicians, rather than identifiable parties with vested interests. Again, these can be good qualities – it defuses some of the potency of the information held in databases if they are primarily held by neutral analysts.

But the secrecy that protects databases from abuse can also often enable abuse from within. Without transparency in the manner in which databases are created and kept, it is all too easy to presume their regularity. This is even more true because the errors or mistakes made in databasing can be hard to uncover – without the systemic or structural oversight discussed above, even recurrent errors can go undetected when viewed on an ad hoc basis. For this reason, so many database related errors involve major scandals rather than minor slip-ups.

It is too much to require courts, or to expect the Constitution, to demand full transparency in the methods of database administrators. But courts, and the Constitution, can demand transparency in the articulation and consequences of those administrators’ efforts. Courts can require that database undergo regular, demonstrably effective auditing processes, and ask to see proof of such. They can view the absence of information about the database – such as how often it is used, how often it is audited, what the results are – as a sign that the database is inadequately attended, rather than as confirmation of its reliability.
D. Use (vs. acquisition) restrictions.

Databasing also requires reconsideration of the current focus upon regulating the acquisition of information, rather than its subsequent maintenance or use. Fourth Amendment doctrine scrupulously attends to how information is required, resting largely upon a premise of a physical dimension to the moment. The “reasonable expectation of privacy” test once promised to consider “people, not places,” but that exhortation has largely reverted to a highly material standard for the constitutional threshold that little corresponds to individuals’ actual subjective expectations.126 The Sixth Amendment likewise strongly preserves the encounter between accused and evidence – the physical face-to-face confrontation embodied in cross-examination – as represented by the Supreme Court’s recent jurisprudence privileging in-court accusation over other, arguably more effective, forms of confronting evidence.127

But the import and the impact of a database occurs less with regard to the moment of the information’s acquisition than with all the moments that then may follow. Indeed, acquisition may not represent any kind of threat to individual liberty or privacy at all. Recall the criminal records database at issue in Reporters’ Committee – there, the Court acknowledged that the true significance of the database was not its contents, which were all technically a matter of public record, but the act of compiling and rendering that information accessible in a particular way.

Yet the salience of the meaning of compiling or organizing information all too often remains obscured in constitutional analysis. Consider again, by way of example, the DNA

126 See, e.g., CHRISTOPHER SLOBOGIN, PRIVACY AT RISK: THE NEW GOVERNMENT SURVEILLANCE AND THE FOURTH AMENDMENT 112.
127 David A. Sklansky, Hearsay’s Last Hurrah, ___ Supreme Ct. Rev. ____ (2010) (criticizing the romanticization of cross-examination as a tool to uncover errors especially with regard to certain kinds of evidence, such as the scientific evidence that now pervades criminal cases, without inquiring whether other methods might in fact be more effective).
databases. Almost all of the cases that have examined the constitutionality of collecting DNA samples from convicted persons, and some of the recent arrestee cases, have zeroed in on the moment of collection as the relevant point of inquiry. If collecting the sample is permissible, then the constitutional inquiry effectively ends.

A more sensitive inquiry would require a more complicated disaggregation of the steps involved in creating and maintaining a DNA database. It would care less about mere collection – an act which, in all candor, is rather easily and painlessly accomplished with or without the consent of the person sampled, and focus instead upon the remaining moments in the process. The actual typing of the samples to determine what genetic code they contained, both in terms of the information revealed and the validation of the accuracy of the methods used, might garner greater attention. The structure of the databasing of the information, including the manner of data entry, the steps taken to ensure accuracy, the approved duration of retention, and the means through which information can be retrieved would likewise merit scrutiny. Lastly, the purpose to which the database could be put – not as a function of the kinds of searches undertaken but rather as a function of the ends such searches intend to serve, would become a critical aspect of any review.

128 Murphy, Paradigms, supra note X, at 1358-62.
129 In the DNA cases, pains are occasionally taken to characterized the typed information as “junk” as a means of assessing the privacy invasion, but even this inquiry is rolled into the assessment of whether the acquisition of the material infringes any constitutionally protected interest. As it is, the Supreme Court has suggested that the testing of biological material constitutes a “search” for Fourth Amendment purposes. Ferguson v. City of Charleston, 532 U.S. 67, 76 (2001). However, even in Ferguson, three justices, including currently sitting Justices Scalia and Thomas, deemed the testing of the urine at issue to not even constitute a search. Id. (Scalia, J., dissenting) at 92 (“I suppose the testing of that urine for traces of unlawful drugs could be considered a search of sorts, but the Fourth Amendment protects only against searches of citizens' ‘persons, houses, papers, and effects’; and it is entirely unrealistic to regard urine as one of the ‘effects’ (i.e., part of the property) of the person who has passed and abandoned.”); see also id. at 92-93 (“Some would argue . . . that testing of the urine is prohibited by some generalized privacy right “emanating” from the ‘penumbras’ of the Constitution (a question that is not before us); but it is not even arguable that the testing of urine that has been lawfully obtained is a Fourth Amendment search. (I may add that, even if it were, the factors legitimizing the taking of the sample, which I discuss below, would likewise legitimize the testing of it.”).
In short, rather than follow an industrial age model reliant upon physical acquisition, constitutional doctrine would transition to an information age approach based on knowledge dissemination. Such attentiveness would offer more effective safeguards around the creation and utilization of databases, and be responsive to concerns about insufficient auditing structures and function creep. Viewed as living, evolving organisms rather than as static repositories of discrete bits of information, the lawfulness and constitutionality of a database would more closely correspond to its actual accuracy and deployment.

E. Benign neglect (vs. deliberate misdeeds).

Finally, regulation of databases require constitutional criminal procedure to focus less upon deliberate or intentional abuses of power than upon unintentional omissions, or mere benign neglect. There is always the risk that a malfeasant actor will corrupt or exploit a database system, to be sure. But constitutional regulation of databases aimed at ferreting out intentional harms will be very thin indeed; it is far easier to do harm, and far greater harm can be done, through mere benign neglect of database systems than through intentional manipulation.

The split among the justices in Herring starkly illustrate this distinction. The majority viewed the sole purpose of the exclusionary rule to be deterrence, and concluded that applying the rule yielded little deterrent benefit with regard to a negligent recordkeeping error.130 In contrast, Justice Ginsburg in dissent argued that more than marginal deterrence was possible, in the specific context of database entry, even when the complained of error constituted mere negligence in care.131

130 129 S. Ct. at 701, 703-04.
131 Id. at 708 (Ginsburg, J., dissenting) (noting that the tort system is premised on deterrence for negligent actions).
Most tellingly, however, the majority of the court essentially presumed the regular operation of the database – even in the face of a factual dispute about the record in this regard\textsuperscript{132} -- whereas the dissenters questioned whether such evidence could in fact ever be effectively adduced.\textsuperscript{133} The dissenters in \textit{Herring} talked about structural reforms and best practices – they could not point to one operator that acted wrongfully or one rule that was flouted. The dissenters, in other words, viewed deterrence as systemic because the database operated systemically, rather than view it as a question of individual deterrence, related to one particular operator’s actions.

To be sure, there are political and ideological fault lines that likely separate the majority and the dissenters in \textit{Herring}; their differences rest on more than mere perspective shifts in the meaning of databasing generally. But the debate is nonetheless illuminating in that it demonstrates how the very nature of databases – their opacity, anonymity, and systematic qualities – can in turn enhance their invisibility to conventional constitutional doctrine.

\textbf{CONCLUSION}

Information databases are an enduring part of the landscape of criminal justice: that much is obvious. They are simply too valuable and too essential to the project of law enforcement to imagine discarding them altogether. But the database model – what goes in them, how they are used, and what comes out – little corresponds to the models of criminal justice that have operated through the ages. Attempts to shoehorn databases into current doctrine have thus largely failed, and simply ignoring the differences risks leaving an important source for investigations and evidence wholly unregulated. By identifying some of the shared, and inaccurate, presumptions

\textsuperscript{132} \textit{Id.} at 706 & n.2, 709.
\textsuperscript{133} \textit{Id.} at 708-09.
that tend to shield databases from closer constitutional scrutiny, and then outlining some of their unique characteristics that merit special attention, this Article underscores the need for deliberate conversation about the significance of this new technology to current criminal justice frameworks.