Maryland Law Review

Volume 71 | Issue 3 Article 5

Sustaining Privacy and Open Justice in the Transition to Online Court Records: A Multidisciplinary Inquiry

Amanda Conley

Anupam Datta

SUSTAINING PRIVACY AND OPEN JUSTICE IN THE TRANSITION TO ONLINE COURT RECORDS: A MULTIDISCIPLINARY INQUIRY[+]

AMANDA CONLEY,[*] ANUPAM DATTA,[**]
HELEN NISSENBAUM[***] & DIVYA SHARMA[****]

I. INTRODUCTION 773
II. COURT RECORDS AND THE PRESUMPTION OF ACCESS 778
A. Definitions 778
B. Legal Basis for the Right of Access 784
1. Sources of the Right of Access 784
2. Contours of the Right of Access 787
a. Who May Access: The Actors 788
b. How Records Can Be Accessed at the Courthouse: Transmission Principles 789
c. What May Be Accessed: Information Type 790
C. Balancing the Presumption of Access Against Other Considerations 797
III. CONTEXTUAL INTEGRITY 803

[+] Acknowledgments: We owe the greatest debt to Grayson Barber, Steve Schultze, and Tim Lee, who generously shared their considerable expertise over the long haul of this project. Jim Rebo and Frank Hoeber of the New Jersey Administrative Office of the Courts graciously opened their doors and provided invaluable guidance, not least by detailing for us the incredible intricacies of the inner workings of a courthouse. We are grateful to Bob Deyling and Thomas Clarke, whose feedback at various stages of the project much improved our arguments. We are also grateful for opportunities to present this work to the NYU Privacy Research Group, Privacy Law Scholars Conference 2011, and the Conference on Privacy and Public Access to Court Records (2008 and 2011). Support for this project came from the National Science Foundation, Cyber-Trust, Award No. CNS-0831124.

[*] Associate, O'Melveny & Myers, LLP, San Francisco. M.A. 2008, University of Colorado (sociology); J.D. 2011, New York University School of Law. The opinions in this Article are solely those of the authors and do not reflect the opinions of O'Melveny & Myers, LLP, or its clients.

[**] Assistant Research Professor, CyLab, Electrical & Computer Engineering, and (by courtesy) Computer Science Department, Carnegie Mellon University.

[***] Professor, Media, Culture, and Communication, and Computer Science, New York University.

[****] Ph.D. candidate, Electrical & Computer Engineering, Carnegie Mellon University.

IV. COMPARING PATTERNS OF INFORMATION FLOW 808
A. *Information Retrieval Model* 810
B. *Overview of Empirical Study*.....814
1. *Search Using Online Systems* 814
a. *Systems Used*.....814
b. *Observations from Search Experience*.....815
2. *Search Using Physical Systems*.....818
a. *Systems Used*.....818
b. *Observations from Search Experience*.....820
C. *Root Causes of Cost Differences in Online and Local Access to Court Records*.....821
V. EVALUATING ACCESS PRACTICES AND POLICIES.....824
A. *Impacts of Information Flows in General Moral and Political Terms*.....824
B. *Values and Purposes Internal to Courts and the Justice System* ...832
1. *Ends and Purposes*.....832
2. *Values* 833
3. *Function of Court Records* 835
4. *Disruptive Information Flows*.....836
VI. RECOMMENDATIONS FOR ACCESS 839
VII. CONCLUSION: A HIDDEN VARIABLE 845

I. INTRODUCTION

Courts, like other institutions, are undergoing a transformation: from largely paper-based systems of processing and record keeping to digital records, and from primarily locally accessible records to records accessible online via the Internet. The adoption of new media, however, is not a new experience for courts. They have adapted to video and audio recording, to microfilm and computer tape, and, in the more distant past, to novel indexing schemes like citation tables and legal citation indexes.[1] Members of the public wishing to consult court records increasingly find a point of entry in computer terminals located in courthouses rather than in the traditional manila

1. *See generally* Patti Ogden, "*Mastering the Lawless Science of Our Law": A Story of Legal Citation Indexes*, 85 LAW LIBR. J. 1 (1993) (discussing the chronology and reasoning behind the evolution of the legal citation index, from the first comprehensive index of overruled cases published in 1821 to online citation systems, first developed in the 1960s).

folders handed over the counter by a court clerk.[2] Yet while the prospect of a move toward fully digitized records accessible online with no gatekeepers has found no shortage of eager supporters and enthusiasts, it has also stirred a chorus of concern among academics and court administrators.[3] Why is this? Of what interest, beyond its purely technological implications, can a change in medium be? Why should anyone outside of a court bureaucracy worry about what goes into records and whether they can be accessed via a doorway to the courthouse or a portal on the Web? These are the questions we address in this Article: not only why we should pay attention to these questions and concerns, but how to go about answering them and crafting access policies accordingly.

Court records exist at the confluence of two strong currents in liberal democratic societies. One current is the demand for openness. Because records provide an essential window into the functioning of one of the three pillars of government—the courts—citizens are presumed to have a right to inspect them to ensure that courts are exercising their powers not only competently and fairly but also within the limits of their mandate.[4] The other current is privacy.[5] The courts are a stage where many of life's dramas are performed, where people may be shamed, vindicated, compensated, punished, judged, or exposed. These human dramas are chronicled through court records, which include volumes of information about the various people involved in a given dispute. It is only natural, and to be ex-

2. *See* Nancy S. Marder, *From "Practical Obscurity" to Web Disclosure: A New Understanding of Public Information*, 59 SYRACUSE L. REV. 441, 444 (2009) (explaining that, before the Internet, an individual wishing to consult court records "would have to go down to the courthouse, request the document, examine it at the courthouse or perhaps copy each page, and then examine the copy at home").

3. *See* Daniel J. Capra et al., *The Philip D. Reed Lecture Series, Conference on Privacy and Internet Access to Court Files, Panel One: General Discussion on Privacy and Public Access to Court Files*, 79 FORDHAM L. REV. 1, 1–22 (2010) (providing the views of certain academics and court professionals attending a conference on public access to court files and the associated privacy concerns).

4. *See* Peter W. Martin, *Online Access to Court Records—From Documents to Data, Particulars to Patterns*, 53 VILL. L. REV. 855, 857–58 (2008) (noting several reasons that justify public access to court records, including enforcing public confidence in the judicial system, assuring the fairness of judicial proceedings, and permitting a public check on the courts).

5. *See* Reena Raggi, *The Philip D. Reed Lecture Series, Conference on Privacy and Internet Access to Court Files, Welcome and Opening Remarks*, 79 FORDHAM L. REV. 1, 2 (2010) (speaking about the traditional concerns that have been raised about disclosure of personal information in court files).

pected, that the creation and exposure of these accumulated volumes raise privacy concerns.[6]

Striving to reconcile these crosscurrents, a complex body of rules, regulations, principles, and policies govern the creation of court records and access to them.[7] In the United States, a range of sources, from constitutional principles[8] to the material facts of particular cases,[9] shape these constantly evolving sets of constraints at the federal, state, and local levels.[10] Some commentators warn that rules and procedures developed for locally accessed or hardcopy records cannot directly be transferred over to electronic records. They also argue that changes brought by new media disrupt a delicate balance by tipping the scale in favor of openness and unacceptably against privacy.[11] Because of these substantive ramifications, the debate over a medium becomes one about core societal values, and it spills beyond court bureaucracies. Hence we find that academics, court administrators, judges, public interest advocates, and citizens are among those calling for discretion and study.[12]

6. *See* Capra, *supra* note 3, at 4–5 (Joel Reidenberg saying that the Internet removed much of the obscurity surrounding court records and, as such, that this open access has raised privacy and public safety implications).

7. *See, e.g.*, Copeland v. Copeland, 966 So.2d 1040, 1050–51 (La. 2007) (noting the state legislature's power to make exceptions to the right of public access to court records, the individual's power to challenge the disclosure of specific court records, and the court's power to determine whether a privacy interest exists in what those records contain, and noting that most states can close or limit access to court records in specific instances).

8. *See infra* notes 58–63 and accompanying text (discussing the lack of clarity surrounding a constitutional right to access court records).

9. *See, e.g.*, *infra* text accompanying note 84 (listing certain kinds of cases where, in New Jersey, records are excluded from public access).

10. *See infra* Part II.B.2.

11. *See* Capra, *supra* note 3, at 4–6 (Joel Reidenberg arguing that the procedure for accessing physical records effectively protected privacy but that electronic records implicate privacy concerns, and suggesting changes that could protect privacy in an online court record system).

12. A yearly conference held in Williamsburg, Virginia, for example, offers scholars, practitioners, judges, and policy advocates the opportunity to discuss the move from paper to online court records. *See* Eighth Conference on Privacy & Public Access to Court Records, held by Center for Legal & Court Technology (Nov. 3-4, 2011), http://www.legaltechcenter.net/education/conferences/8th-conference-on-privacy-public-access-to-court-records/ (providing the proposed itinerary and topics for discussion for the November 2011 conference). It bears noting that not all participants at the conference are necessarily against full and open access to court records on the Internet, and not all favor privacy over judicial openness. Members of the Reporters Committee for the Freedom of the Press, for example, strongly advocate for entirely open and accessible court records. *See* Capra, *supra* note 3, at 12–16 (Lucy Dalglish discussing how open court records are beneficial to journalists).

Taking on questions about the disruptive potential of electronic media and digital networks, this Article aims to contribute to the debate over what courts ought to do. At this time, avoiding electronic media and digital networks entirely is not a serious option, nor, in our view, is it defensible. Some readers, however, might wonder whether anything less than full adoption is even worth considering, given that the federal courts have already committed to the Public Access to Court Electronic Records ("PACER")[13] online record system. In our view, there remain issues of principle both in PACER and other electronic access systems that have not been adequately addressed.[14] Our Article and its concluding recommendations focus primarily on state and local, rather than federal, courts. We do this in part because public and internal deliberations over state access policies have remained actively in progress in the period during which we conducted research and wrote this Article.[15] But, just as importantly, we focus on state court records, particularly at the trial level, because they contain an abundance of personal information, some of which may drop away as cases move from trial courts to appellate courts.[16] Because of the

13. PACER is an electronic public access service provided by the Administrative Office of the U.S. Courts that contains "case and docket information from federal appellate, district and bankruptcy courts, and the PACER Case locator via the Internet." PUBLIC ACCESS TO COURT ELECTRONIC RECORDS, http://www.pacer.gov (last visited Mar. 8, 2012).

14. For example, Peter Martin makes the important and often overlooked point that while one purported purpose of increased access to court records is to enhance judicial accountability, PACER—a system developed for use by lawyers and judges, rather than by the general public—does not allow users to search cases by judge. *See* Martin, *supra* note 4, at 871 ("[T]here is no search feature comparable to PACER's party name search that would allow a user to gather and inspect judge or attorney actions across multiple cases. Of course, the system holds this data, but it does not permit the data fields for judges and attorneys to be the subjects of search." (footnote omitted)). This is not PACER's only limitation. *See, e.g.*, John Schwartz, *An Effort to Upgrade a Court Archive System to Free and Easy*, N.Y. TIMES, Feb. 12, 2009, at A16 ("'Pacer is just so awful,' said Carl Malamud, the leader of the effort and founder of a nonprofit group, Public.Resource.org. 'The system is 15 to 20 years out of date.' Worse, Mr. Malamud said, PACER takes information that he believes should be free—government-produced documents are not covered by copyright—and charges 8 cents a page.").

15. New York, New Jersey, California, Arkansas, Tennessee, and Wyoming, among others, have all recently or are currently in the process of evaluating logistical, accessibility, and privacy concerns in the move from paper to online court records. *See, e.g.*, BARRY T. ALBIN, SUPREME COURT OF N.J., REPORT OF THE SUPREME COURT SPECIAL COMMITTEE ON PUBLIC ACCESS TO COURT RECORDS i–iii (2007) [hereinafter REPORT ON PUBLIC ACCESS], *available at* http://www.judiciary.state.nj.us/publicaccess/publicaccess.pdf (explaining that "the Committee debated the delicate balance between the public's general right to know and the individual's limited right of privacy within our court system, and [considered] how placing court records on the Internet will alter exponentially the calculus between those competing rights," and outlining the Committee's recommendations).

16. In New Jersey, for example, the record on appeal need only contain "all papers on file in the court or courts or agencies below, with all entries as to matters made on the

enormous variability across jurisdictions, we used the New Jersey courts as a reference point for honing our ideas about the transition from paper to electronic records. New Jersey presented a particularly rich environment because it is in the process of transforming to digital court filing and record keeping, and has facilitated deliberation over the ramifications of this transformation, including for privacy.[17]

The conclusion of our inquiry is that courts have an obligation to rewrite rules governing the creation of, and access to, public court records in light of substantive changes that online access augurs.[18] Although it lies outside the scope of this Article to specify these rules, we aim to trace a line of analysis toward developing such policy and regulation, whether through case law or administrative rules. Our analysis integrates the diverse disciplinary expertise of its authors: traditional legal sources have informed our grasp of the status quo;[19] a conception of a right to privacy, drawn from the philosophical theory of contextual integrity, has guided our evaluation of access policies;[20] and a grasp of underlying technical capacities of information retrieval systems has shaped the comparative study of information flows, bringing rigor to and generalizing prior work on contextual integrity.[21]

How to read this article: What follows is a product of distinct multidisciplinary perspectives—legal, philosophical, and technical. Because each has influenced the others and has shaped our findings and recommendations, we chose to present an amalgamation of all three even though the resulting paper does not embody the traditional form of any one of them. Readers who prefer to focus on arguments relevant to their particular disciplines may wish to read only those parts that pique their interest and skip those that are less pertinent. A

records of such courts and agencies, the stenographic transcript or statement of the proceedings therein, and all papers filed with or entries made on the records of the appellate court." N.J. R. 2:5-4(a).

17. *See* REPORT ON PUBLIC ACCESS, *supra* note 15, at 9 (discussing the need to balance the tradition of public access to records with the concern for privacy, and noting that a comprehensive listing of exceptions to public access is necessary to help reach that balance).

18. *See infra* Part V (suggesting that if the rules governing online access to personal information do not change, then the role of the courts may diminish and parties might choose other avenues to settle disputes).

19. *See infra* Part II.B.1 (noting the common law right of access and referencing the constitutional and common law bases for making court documents open to the public).

20. *See infra* Part III (discussing the theory of contextual integrity).

21. *See infra* Part IV.A–B (discussing a model information retrieval system and detailing the results of actual searches using PACER and Google Scholar); *see generally* HELEN NISSENBAUM, PRIVACY IN CONTEXT: TECHNOLOGY, POLICY, AND THE INTEGRITY OF SOCIAL LIFE (2010) [hereinafter NISSENBAUM, PRIVACY IN CONTEXT] (setting out the theory and framework of contextual integrity).

quick overview will help in this determination. Part II draws mostly on legal discussions. In particular, Part II.A presents an integral definition of the term "court record," Part II.B discusses the sources of a legal right of access to court records, and Part II.C introduces key mechanisms by which courts determine whether to grant requests to seal or redact information contained in their records. Part III provides a brief philosophical overview of contextual integrity—a theory of information privacy that guides the inquiry in Part IV and analysis in Part V. Part IV develops a model of information retrieval for court records and uses it to compare typical or expected information flows for in-person access to court records at the courthouse with online access to court records via PACER and Google Scholar. Focusing on the transformation from in-person to online access, Part V presents a comparative normative analysis of information flows, drawing together threads from Parts II, III, and IV. The Article concludes in Part VI with general recommendations for addressing privacy concerns.

II. COURT RECORDS AND THE PRESUMPTION OF ACCESS

A. Definitions

Court records are a subset of government records to which the public has access. Because the term "court records" is imprecise, we take this opportunity to explain how we will use this term throughout the paper. Consider a plaintiff who brings a lawsuit against his boss for employment discrimination.[22] In order to bring this civil claim, the employee will need to fill out a complaint form, which details his claims against his boss, the defendant, and file it with the court. He must also serve the boss with a summons to appear in court. In reply, his boss will file an answer to the complaint, responding to the employee's allegations and possibly bringing counterclaims of her own. Through a procedure known as pre-trial discovery, each party's attorney (if they are both represented) will then request documents and other evidence from the opposing party, and possibly from relevant third parties. Before the trial has even begun, each side will likely have amassed hundreds or even thousands of documents full of information about the people, places, and events involved in the dispute. The parties will take depositions, which will be recorded or

22. For an explanation of the application of the Federal Rules of Civil Procedure, which would govern such an action in the federal courts, see generally A.J. STEPHANI & GLEN WEISSENBERGER, FEDERAL CIVIL PROCEDURE LITIGATION MANUAL (Matthew Bender ed., 3d ed. 2011). For state procedural rules specifically applicable to this New Jersey case study, see N.J. R. 4:1–4:101.

transcribed. Each side will file a variety of additional motions, often with exhibits and declarations attached. At this point the court may dismiss the case by summary judgment, the parties may settle out of court, or the case may go to trial.[23] If the case goes to trial, an even larger case file will be created that will likely include court transcripts in writing, audio, or video form, or a combination of the three;[24] witness testimony; evidence and exhibits; information about jury selection;[25] additional motions; and, ultimately, a judicial opinion setting out the final decision in the case.

23. Less than 2 percent of federal civil cases go to trial. ROBERT P. BURNS, THE DEATH OF THE AMERICAN TRIAL 2 (2009). The pattern is similar for state cases. *Id.*

24. Sworn officers of the court create transcripts. There are guidelines about what is and is not included in the transcript, and a delay period before the transcript is placed on PACER. Parties are notified in advance that the transcript is going "live" and can move to not have it published or make specific redaction requests. *See, e.g.*, U.S. DIST. COURT DIST. OF S.C., TRANSCRIPT FILING INSTRUCTIONS FOR ATTORNEYS (2011), *available at* http://www.scd.uscourts.gov/CMECF/DOCS/Transcript_Filing_Instructions_Attorneys.pdf (describing how a transcript is made available through PACER after a ninety-day restricted review period, a Notice of Electronic Filing, and the opportunity to redact certain personal identifiers).

25. In New Jersey, courts do not release a list of juror names at any point during or after a trial, but names may nevertheless be available in the *voir dire* transcript. Telephone Interview with Gina Fe' Whittaker, Jury Manager, Burlington County Superior Court (Apr. 26, 2011). In high-profile cases, juror names are sometimes kept secret as a precautionary measure. *See* KIMBERLEY KEYES, REPORTERS COMM. FOR FREEDOM OF THE PRESS, SECRET JUSTICE: SECRET JURIES (2005), *available at* http://www.rcfp.org/secret-justice-secret-juries/secret-juries (noting that in cases involving terrorism or organized crime, a judge might choose to use an anonymous jury to protect the jurors). Other state courts, including Ohio and Michigan, have found a qualified First Amendment right of access to juror names and addresses. *See* State *ex rel.* Beacon Journal Publ'g Co. v. Bond, 781 N.E.2d 180, 194 (Ohio 2002) ("[W]e hold that the First Amendment qualified right of access extends to juror names and addresses, thereby creating a presumption of openness that may be overcome 'only by an overriding interest . . . .'") (citation omitted)); *In re* Disclosure of Juror Names and Addresses, 592 N.W.2d 798, 809 (Mich. Ct. App. 1999) ("We therefore hold that the press has a qualified right of postverdict access to juror names and addresses, subject to the trial court's discretion to fashion an order that takes into account the competing interest of juror safety . . . ."). The Ohio Supreme Court has also recognized a qualified First Amendment right to access prospective juror questionnaires, which are used to determine whether potential jurors are suitable for service. *Beacon Journal*, 781 N.E.2d at 188. In federal courts, however, "documents containing identifying information about jurors or potential jurors are no longer included in the public case file and are unavailable to the public, either electronically *or* at the courthouse." KEYES, *supra*. In March 2004, the Judicial Conference of the United States, the principal policy-making body of the federal court system, formally adopted this restriction as part of its guidelines about public access to electronic criminal case files. *See* JUDICIAL CONFERENCE OF THE UNITED STATES, REPORT OF THE PROCEEDINGS OF THE JUDICIAL CONFERENCE OF THE UNITED STATES 10 (2004) [hereinafter PROCEEDINGS OF THE JUDICIAL CONFERENCE 2004], *available at* http://www.uscourts.gov/FederalCourts/JudicialConference/Proceedings/Proceedings.aspx?doc=/uscourts/FederalCourts/judconf/proceedings/2004-03.pdf (adopting a policy that required, in part, the redaction of personal data identifiers from all court documents that could be accessed by the public). *But see* Op. Fla. Att'y Gen. AGO 2005-61, *available at*

Once all the relevant claims have been resolved, whether on summary judgment or at trial (settlement agreements are typically not available for public inspection),[26] some number of the above-listed documents and files will become a permanent part of the court record[27] for this case, for example by being introduced into evidence as exhibits or attached to motions or other papers filed with the court. Depending on a number of factors, every document and media file created during this process may become part of the final court record, or only certain documents may be entered into the record while other information-rich documents are excluded.[28] We will discuss some of the privacy-related justifications for keeping a file out of the final court record below,[29] but other reasons may be at play as well, including those related to ease of storage,[30] file or software compatibility,[31] or simply local rule or custom.[32] Of course, not all documents produced during the pre-trial discovery period will ultimately become part of the court record, but a fair number may be included, particularly those that the parties rely on in their moving papers.

In this Article, we will use the term "case file" to refer to the sum total of documents, media files, and exhibits that are produced and collected by the parties and/or the court as a case makes its way through the system. We will use "court record" to refer to the subset of these documents that remain after a case has been resolved and

http://myfloridalegal.com/ago.nsf/Opinions/4A410DA4389543CC852570C000750D51 (stating that juror names and addresses are not exempt from public disclosure pursuant to a Florida state public records statute). For a general discussion of the risks and benefits of juror anonymity, see Christopher Keleher, *The Repercussions of Anonymous Juries*, 44 U.S.F. L. REV. 531, 547–59 (2010) (discussing how juror anonymity affects such elements as juror privacy and safety, *voir dire*, summation, and right of access to trial proceedings).

26. *See* Sharona Hoffman, *Settling the Matter: Does Title I of the ADA Work?*, 59 ALA. L. REV. 305, 313 (2008) (noting that "settlement amounts are generally not recorded in publicly available court documents").

27. This term should be distinguished from the official Appellate Record (or the Record on Appeal)—the record an appellate court has access to and reviews in an appeal—which may or may not be identical to the publicly available record at the trial court. *See, e.g.*, *supra* note 16 and accompanying text.

28. *See infra* Part II.B.2.c (discussing how certain documents, or parts of documents, may be let in or kept out of court records by various means, including state rules, sealing of records, or redaction of personal information).

29. *See infra* Part II.B.2.c.

30. *See infra* text accompanying notes 111–113.

31. *Cf.* Robert A. Guy Jr., *A Model Protocol for Electronic Filing: Best Practices for Law Firms Making the Transition to Case Management/Electronic Case Files*, 53 FED. LAW. 38, 45, Aug. 2006 (noting some potential errors, including software errors, that may occur when an attorney is filing court documents electronically into an electronic case management system).

32. *See infra* Part IV.C.2.b-c.

become part of the permanent, *public* record.[33] If a court chooses to allow remote electronic access to its records, the entire court record may be placed online, or only a subset or summary of the record may be available. As we discuss at length in Part IV, online access may be limited by logistic constraints, may be available only to certain categories of persons, may be located behind a paywall, or some combination of the three.

As we have defined it, a case file may include thousands of documents full of information about the parties and the dispute. It may contain motions; pleadings; briefs; attachments; dockets; transcripts in textual, audio,[34] or video form; exhibits entered into evidence; and records and responses to interrogatories and requests for admission produced during pre-trial discovery. Each and every form filled out by the parties, their lawyers, or by related third parties (witnesses, jurors, etc.) potentially contains vast amounts of personal data including home or school addresses, places of employment, birthdates, and, in many cases, Social Security numbers.[35] While much of this information may appear innocuous and uninteresting to anyone but the parties themselves, some of it may be quite revealing or even embarrassing for the individuals and organizations involved in a dispute.[36] Exhibits entered into evidence may include bank statements, medical records, psychological evaluations, personal and business emails, and other intimate details about each party and their interactions with one another or with others. All or a portion of this file may become part of the public record, available in paper form at the courthouse

33. It bears noting that while we are making this analytical distinction—between the case file and the court record—it is not necessarily accurate in practice. Documents in the case file may become public at the time of trial, for example, complicating our categorization. Thus, we make this distinction here only for analytical purposes.

34. *See, e.g.*, JUDICIAL CONFERENCE OF THE UNITED STATES, REPORT OF THE PROCEEDINGS OF THE JUDICIAL CONFERENCE OF THE UNITED STATES 9–10 (2010) [hereinafter PROCEEDINGS OF THE JUDICIAL CONFERENCE 2010], *available at* http://www.uscourts.gov/FederalCourts/JudicialConference/Proceedings/Proceedings.aspx?doc=/uscourts/FederalCourts/judconf/proceedings/2010-03.pdf (stating that the Judicial Conference adopted a recommendation that digital audio files be available to the public through the PACER system, especially because the availability of such recordings improved public access).

35. For example, a person filing a complaint with the New Jersey courts is required to provide the following information: name, address, phone number, name of defendant, defendant's address, summary of issue that prompted filing of complaint, summary of harm caused by defendant, date, and signature. *See* N.J. Civil Compl. Form, *available at* http://www.judiciary.state.nj.us/civil/forms/11210_civil_action_complaint.pdf.

36. *See* Marder, *supra* note 2, at 445–47 (noting that the personal disclosures that occur in litigation may include the names of rape victims, evidence of marital infidelity, personal assets and debts, and other information that the parties involved wish to keep private).

and, in some states and in the federal court system,[37] via remote access in whole or in part on the Internet.

Even if the information in the case file does not itself immediately appear revealing, when combined with other publicly available data, such as phone and property records, it may provide ample information for identity thieves.[38] Especially (but not exclusively) if the dispute took place several years ago, before the widespread use of Social Security numbers in identity theft became a major concern, documents may be riddled with this information.[39] In most states, the burden is on the lawyers and their clients to redact sensitive information like Social Security numbers before filing a document with the court.[40] Even the most well intentioned counsel, however, may let a few items slip, particularly in files cataloged by hand which may include written Social Security numbers in the upper right corner of the file.[41] When these documents are scanned into portable document format ("PDF"), hand-written Social Security numbers will be nearly impossible to locate and remove through any software that a court might use to detect and black out sensitive information.[42] Signatures

37. For an in-depth discussion of the PACER system, see *infra* Part IV.

38. *See* Marder, *supra* note 2, at 447 (noting that personal information in case files can lead to other personal information available elsewhere on the Internet, which can then be compiled and used to commit identity theft).

39. For an extensive discussion of the use of Social Security numbers in certain documents, the associated risk of identity theft, and legislation and court cases addressing Social Security numbers, see *Social Security Numbers*, ELECTRONIC PRIVACY INFORMATION CENTER, http://epic.org/privacy/ssn/ (last visited Mar. 5, 2012).

40. New Jersey, for example, "requires the parties to redact any confidential personal identifiers from documents submitted to the court, but . . . fails to provide an enforcement mechanism to punish a party's lack of compliance." Kristin M. Makar, Comment, *Taming Technology in the Context of the Public Access Doctrine: New Jersey's Amended Rule 1:38*, 41 SETON HALL L. REV 1071, 1073 (2011).

41. *See* Andy Opsahl, *Privacy: Agencies Struggle to Redact Personal Data from Online Public Documents*, GOV'T TECH. (July 8, 2008), http://www.govtech.com/gt/375540 ("These documents frequently contain Social Security numbers (SSNs), mothers' maiden names, signatures, minors' names and other red meat for identity thieves and stalkers. . . . To redact SSNs, states use software to black out the section of the document where the number appears. Redaction software vendors typically guarantee 98 percent accuracy, but SSNs tend to crop up everywhere in documents, increasing the difficulty of redacting them. For example, some lenders required borrowers to write their SSNs beneath their signatures. Other lenders organized the documents by SSN within their office filing systems. To do that, they wrote the SSNs on the upper right corners of the documents."); *see also* Jacob Ogles, *Court Documents Not Fit For Web?*, WIRED.COM (Nov. 23, 2004), http://www.wired.com/politics/security/news/2004/11/65703 ("Though most states require information like Social Security numbers or dates of birth to be concealed before documents are posted on a web server, it is impractical and sometimes impossible for clerks to catch every instance.").

42. *See* Timothy B. Lee, *Studying the Frequency of Redaction Failures in PACER*, FREEDOM TO TINKER (May 25, 2011, 1:52 PM), https://freedom-to-tinker.com/blog/tblee/studying-

on forms may be scanned as well, and if these scans are ultimately placed on the web, their presence further increases the risk of identity theft.[43]

Case files may contain personally identifiable or otherwise revealing information not only about the parties in a dispute, but also family members, colleagues, witnesses, jurors, and even victims or wronged parties in criminal or civil cases, respectively.[44] Laws regulating when and how such information may be disclosed, particularly with respect to victim and juror information, vary by state. For example, the website of a Massachusetts victims' rights group warns sexual assault victims that while their names will be blacked out in all police reports and court records, they must file a formal request with the judge to have their address, telephone number, place of employment, or school location redacted.[45]

While the focus of this Article is on court records of civil cases, it bears noting that when an individual is arrested or indicted, this information will most likely become part of the publicly available court record *even if the charges are ultimately dismissed.*[46] Thus, even in cases where charges were wrongfully brought—a case of mistaken identity, perhaps, or simply a misunderstanding—the record of that individual's history in the criminal justice system will remain. Rarely will this

frequency-redaction-failures-pacer; Timothy B. Lee, *What Gets Redacted in Pacer*, FREEDOM TO TINKER (June 16, 2011, 1:48 PM), https://freedom-to-tinker.com/blog/tblee/what-gets-redacted-pacer.

43. Opsahl, *supra* note 41.

44. *See infra* Part II.C (discussing the privacy concerns of unrepresented third parties).

45. The website reads:

> In Massachusetts, if you make a report to the police about sexual assault, [Massachusetts law] makes sexual assault reports to the police unavailable to the general public. They are kept in separate, confidential files. However, if your case becomes active in the criminal courts, your police report will be made public. Another Massachusetts law . . . requires that your name should automatically be blacked-out ("redacted") in all public police records and court documents. This means that your name should not be released to the public. You also have the additional option to ask the Judge to keep your address, telephone number, place of employment, and/or location of your school private ("impounded."). If the Judge agrees, this personal information will be redacted from the record and not be stated in open court. Please note, in order to do this, you must request this from the Judge as soon as possible.

Privacy, VICTIM RIGHTS LAW CENTER, http://www.victimrights.org/node/109 (last visited Mar. 5, 2012).

46. *See, e.g.*, COMMONWEALTH OF MASSACHUSETTS, DISTRICT COURT DEPARTMENT OF THE TRIAL COURT, A GUIDE TO PUBLIC ACCESS, SEALING & EXPUNGEMENT OF DISTRICT COURT RECORDS 11–13 (2010), *available at* http://www.mass.gov/courts/courtsandjudges/courts/districtcourt/pubaccesscourtrecords.pdf (listing scenarios in criminal cases where information remains public despite the case being dismissed, the prosecution being abandoned, or the verdict being not guilty).

record note that the charges were dismissed, or that the individual was found to be innocent. Instead, the record will likely end with the arrest or arraignment, providing no further details. A potential employer can easily access this information and may reasonably—though mistakenly—assume that the applicant has a criminal record.[47] This is particularly disconcerting given that low-income individuals and people of color are considerably more likely to be wrongfully accused or arrested.[48]

B. Legal Basis for the Right of Access

1. Sources of the Right of Access

Under the Freedom of Information Act ("FOIA"), passed in 1966, the majority of government records in the United States are open to public inspection.[49] If you are curious what the Food and Drug Administration or the Department of Homeland Security are up to, a FOIA request can provide that information (including that which is not already made available on agencies' respective web-

47. This type of unfair discrimination is not limited to employment opportunities. In Minnesota, "banks used court records to pore through financial and other records of minority applicants and then decided to deny home loans to some of those individuals, according to a Minnesota court advisory committee report issued in September." Ogles, *supra* note 41.

48. This assumes that white and African-American individuals commit crimes at equivalent rates. *See* Vincent Schiraldi & Jason Ziedenberg, *Race and Incarceration in Maryland*, JUST. POL'Y INST. 6 (Oct. 23, 2003), http://www.justicepolicy.org/uploads/justicepolicy/documents/03-10_rep_mdraceincarceration_ac-md-rd.pdf (stating that the likelihood of being incarcerated is seven to eight times greater for African-Americans than for whites). Harvey Grossman, the Legal Director for the ACLU of Illinois, stated:

> In looking at this data [from the Illinois Traffic Stops Study], it is reasonable to conclude that for police in Illinois, a driver's race is a proxy for suspicion in deciding whether to request a consent search. . . . It is clear that police view drivers of color with far more suspicion than their white counterparts.

ROGER BALDWIN FOUNDATION OF ACLU, ANNUAL REPORT 2008-2009: BUILDING AN AMERICA WE CAN BE PROUD OF 9, *available at* http://www.aclu-il.org/wp-content/uploads/2011/01/Annual-Report-2009-webversion.pdf. Racial profiling is not limited to traffic stops. The Justice Policy Institute found that predominantly African-American neighborhoods are targeted more often for violations of drug laws. Schiraldi & Ziedenberg, *supra*, at 16. Despite the fact that "African Americans and whites use drugs at a comparable rate, [African Americans] represent 68% of those arrested for drug offenses, and 90% of those incarcerated for drug offenses" nationwide. Ngozi Caleb Kamalu et al., *Racial Disparities in Sentencing: Implications for the Criminal Justice System and the African American Community*, 4 AFR. J. CRIMINOLOGY & JUST. STUD. 1, 15 (2010).

49. Freedom of Information Act, 5 U.S.C. § 552 (2006). Most states have an analogous state statute providing access to state government records. *See generally* ANNE WELLS BRANSCOMB, WHO OWNS INFORMATION? FROM PRIVACY TO PUBLIC ACCESS 8 (1994) (providing an overview of FOIA and privacy legislation in the United States).

sites).[50] But requests pursuant to this act cannot reach records of the judicial branch—namely, court records—because FOIA applies by its terms only to *agency* records (records of the executive branch).[51]

The tradition of providing public access to court records, however, pre-dates FOIA.[52] Given all the privacy concerns noted above, one reasonably might ask why the public is given a right to access such records in the first place.[53] But the right to open courts and their records is actually as longstanding as our right to the courts and to justice itself: it is based on the widely held belief that for a justice system to function successfully and consistently, it must be accountable to its citizens.[54] In theory, open access to everything that happens within this system, from the first filing of a lawsuit or of criminal charges, to the final hearing for damages or sentencing, provides journalists and concerned citizens alike with the opportunity to ensure that justice is being apportioned fairly and consistently throughout the country.[55] Judicial openness has two components: the right of access to trials themselves, and the right of access to judicial documents for inspection and copying.[56] The right of public access to

50. *See generally* REPORTERS COMM. FOR FREEDOM OF THE PRESS, FEDERAL OPEN GOVERNMENT GUIDE (10th ed. 2009), *available at* http://www.rcfp.org/federal-open-government-guide (describing the various components of FOIA, including how to make a request, what agencies are subject to those requests, and how to track those requests).

51. The E-Government Act of 2002 does provide for public access to federal records across all three branches of government. Under this Act, federal courts are required to establish and maintain websites with docket information for pending cases and closed cases (up to one year old), electronically filed court documents, judicial opinions in text searchable format "regardless of whether such opinions are to be published in the official court reporter," and "[a]ny other information . . . that the court determines useful to the public." E-Government Act of 2002, Pub. L. No. 107-347, § 205(a)–(b), 116 Stat. 2910, 2913–14 (codified at 44 U.S.C. § 3501 (2006)). It should be noted that not all electronically filed court documents are subject to disclosure. *See id.* at § 205(c).

52. *See, e.g.*, UNITED STATES DEP'T OF JUSTICE, ATTORNEY GENERAL'S MANUAL ON THE ADMINISTRATIVE PROCEDURE ACT 24 (1947) (discussing the Administrative Procedure Act, which pre-dated FOIA, and when records could be made available under the APA).

53. One may also ask why we do not litigate anonymously. *See generally* Lior Jacob Strahilevitz, *Pseudonymous Litigation*, 77 U. CHI. L. REV. 1239 (2010) (providing further discussion of the potential merits of anonymous pseudonymous litigation).

54. *See* Globe Newspaper Co. v. Superior Court, 457 U.S. 596, 606 (1982) ("Public scrutiny of a criminal trial enhances the quality and safeguards the integrity of the factfinding process, with benefits to both the defendant and to society as a whole.").

55. *See id.* ("[P]ublic access to the criminal trial fosters an appearance of fairness, thereby heightening public respect for the judicial process.").

56. *See* Richmond Newspapers, Inc. v. Virginia, 448 U.S. 555, 580 (1980) (plurality opinion) ("We hold that the right to attend criminal trials is implicit in the guarantees of the First Amendment . . . ." (footnote omitted)); *see also* United States v. Gotti, 322 F. Supp. 2d 230, 239 (E.D.N.Y. 2004) (discussing both the right of access to judicial proceedings and right of access to judicial documents).

judicial documents predates the Constitution and may be traced as far back as English common law, the predecessor to the U.S. legal system.[57]

While the common-law right of access to judicial documents has been well recognized, whether such a right is separately rooted in the Constitution remains somewhat of an open question.[58] In *Richmond Newspapers, Inc. v. Virginia*, the Supreme Court held that both the First and Fourteenth Amendments grant the press and the public the right to attend criminal trials.[59] But the right to attend judicial proceedings does not necessarily give one access to any judicial documents used in or produced by such proceedings.[60] To wit, in *Zenith Radio Corp. v. Matsushita Electric Industrial Co.* a federal district court held that while the Supreme Court has recognized First and Fourteenth Amendment rights of access to criminal trials themselves,[61] the second component of judicial openness—access to all types of judicial records for inspection and copying—is not constitutionally protected.[62] Thus, while the

57. United States v. Amodeo, 44 F.3d 141, 145 (2d Cir. 1995) (citing Leucadia, Inc. v. Applied Extrusion Techs., Inc., 998 F.2d 157, 161 (3d Cir. 1993)); *see also* United States v. Criden, 648 F.2d 814, 819 (3d Cir. 1981) (further supporting that the right to inspect public records predates the Constitution); *Gotti*, 322 F. Supp. 2d at 239 ("The common law right of access to judicial documents under American jurisprudence traces its origin to the general English common law right of access to public records, but has a broader reach.").

58. The Supreme Court has laid out a two-part test for determining whether the public has a First Amendment right of access to judicial proceedings: (1) Has the proceeding historically been open? (2) Does the right of access play an essential role in the functioning of the judicial process and the government as a whole? *See In re* Reporters Comm. for Freedom of the Press, 773 F.2d 1325, 1331–32 (D.C. Cir. 1985) ("Apparently, both these questions must be answered affirmatively before a constitutional requirement of access can be imposed."); *see also Globe Newspaper Co.*, 457 U.S. at 606 (noting that the right of access to court records is based on "constitutional stature," but is "not absolute"); *Richmond Newspapers*, 448 U.S. at 569–71 (discussing how historical evidence supports that the right to open access to trials stems from common law).

59. *Richmond Newspapers*, 448 U.S. at 580.

60. Zenith Radio Corp. v. Matsushita Elec. Indus. Co., 529 F. Supp. 866, 897 (E.D. Pa. 1981).

61. *Id.* at 895 (citing *Richmond Newspapers*, 448 U.S. 555).

62. *Id.* at 908 ("With respect to the question whether the common law right to inspect and copy has a constitutional dimension, we conclude that it does not."); *see also* Brown & Williamson Tobacco Corp. v. FTC, 710 F.2d 1165, 1176–81 (6th Cir. 1983) (holding that the First Amendment limits judicial discretion to seal court documents but also stating that the right of access is still not absolute); *Gotti*, 322 F. Supp. 2d at 243–50 (noting a qualified First Amendment right of access to both judicial proceedings and judicial documents). *But see* Republican Co. v. Appeals Court, 812 N.E.2d 887, 892 (Mass. 2004) ("The exercise of the power to restrict access, however, must recognize that impoundment is always the exception to the rule, and the power to deny public access to judicial records is to be 'strictly construed in favor of the general principle of publicity.'" (citation omitted)).

right of access remains available under common law, the First Amendment may not add any additional protection.[63]

2. Contours of the Right of Access

In 1978, the Supreme Court began to clearly articulate the contours of the common-law right of access to court records and the justifications for this right.[64] Federal and state courts and legislatures have picked up this project as well, outlining the precise shape of this right and enshrining it in both common and state statutory law.[65]

As part of a broader discussion about the purposes of public access, federal and state courts have addressed such questions as: *Who* is entitled to access judicial documents?[66] *What* types of documents are included in this right?[67] And *when* should access be allowed or denied based on concerns about secrecy or control over information?[68] While the presumption of access to court records is longstanding, it has never been interpreted to allow access to everything in the case file.[69] These restrictions on access trickle down from state and federal appellate courts to the local courthouses themselves, where state and local law, custom, and in some cases simply the whims of court clerks determine which information in the court record will actually be made available to the public, and how.

Within the broad mandate to provide access, courts are given some leeway in terms of how they implement the public's right to inspect and copy judicial records.[70] In the following section, we provide a basic outline of the restrictions on access to court records as well as the balancing tests judges employ to determine when and why some court records, in whole or in part, will not be available for public inspection. At each point, it is important to remember that the law remains unclear at the level of individual access in part because the sta-

63. *Zenith Radio Corp.*, 529 F. Supp. at 913 (concluding "that a constitutional right of access to judicial records does not exist").

64. Nixon v. Warner Commc'ns, Inc., 435 U.S. 589, 597, 607 (1978).

65. Daniel J. Solove, *Access and Aggregation: Public Records, Privacy and the Constitution*, 86 MINN. L. REV. 1137, 1158–63 (2002).

66. *See infra* Part II.B.2.a.

67. *See infra* Part II.B.2.c.

68. *See infra* Part II.B.2.c.

69. Zenith Radio Corp. v. Matsushita Elec. Indus. Co., 529 F. Supp. 866, 897–901 (E.D. Pa. 1981).

70. For an extensive discussion of the legal basis of the right to access court records, as well as the specific contours of this right in each U.S. state, see REPORTERS COMM. FOR FREEDOM OF THE PRESS, ELECTRONIC ACCESS TO COURT RECORDS, SPRING 2007, *available at* http://www.rcfp.org/rcfp/orders/docs/EACR.pdf.

tutory language is not itself precise, and in part because many of the decisions about access are made on an ad hoc basis by individual court clerks or other employees.[71]

a. Who May Access: The Actors

We all have the image in our minds of a concerned citizen going down to the local courthouse to page through old court records, perhaps looking for information on land title or family history. But the details of this scenario may be a bit fuzzy in our heads. To which records does she have access? May she remove them from the courthouse? Make copies of them at a local terminal? Does she need to prove that she has some connection to these records, some valid interest in acquiring them? Is there a restriction on what she may do with the information in the records once she obtains it?

Generally speaking, an individual need only show a "legitimate interest" in the public record requested in order to gain access.[72] As early as 1908, the Supreme Court of Alabama held that the clerk of the court was compelled by statute to provide copies of papers held by the court to any person requesting them, and had no power to determine whether the requester should be granted the copies sought.[73] In 1995, the New Jersey Supreme Court determined that even a purely for-profit interest in judicial records counts as "legitimate," and therefore allowed a seller of municipal tax-assessment data and a commercial real estate appraiser to obtain copies of computerized records of municipal tax assessments for use in their business enterprises.[74]

71. *See infra* Part II.B.2.c & note 107. This is not to suggest that all courts do not have clear access policies. Many do, and many others are in the process of constructing them.

72. *See, e.g.*, Nixon v. Warner Commc'ns, Inc., 435 U.S. 589, 597–98 (1978) ("American decisions generally do not condition enforcement of this right on a proprietary interest in the document or upon a need for it as evidence in a lawsuit. The interest necessary to support the issuance of a writ compelling access has been found, for example, in the citizen's desire to keep a watchful eye on the workings of public agencies . . . and in a newspaper publisher's intention to publish information concerning the operation of government." (citation omitted)); Irval Realty Inc. v. Bd. of Pub. Util. Comm'rs, 294 A.2d 425, 428 (N.J. 1972) (a citizen requesting records must "show an interest in the subject matter of the material he sought to scrutinize. Such interest need not have been purely personal. As one citizen or taxpayer out of many, concerned with a public problem or issue, he might demand and be accorded access to public records bearing upon the problem, even though his individual interest may have been slight.").

73. Jackson v. Mobley, 47 So. 590, 593 (Ala. 1908). *But see Nixon*, 435 U.S. at 598 ("Every court has supervisory power over its own records and files, and access has been denied where court files might have become a vehicle for improper purposes.").

74. Higg-A-Rella, Inc. v. County of Essex, 660 A.2d 1163, 1169 (N.J. 1995). *But see* Burnett v. County of Bergen, 968 A.2d 1151, 1164–66 (N.J. 2009) (stating that while a com-

b. How Records Can Be Accessed at the Courthouse: Transmission Principles

Whether a record-seeker needs to show identification, pass through a metal detector, or sign in to access records depends entirely on the particular court she is visiting.[75] At the New Jersey Supreme Court, she would need to provide identification and receive a name badge with her photo on it in order to enter the courthouse; if she has already visited before, her name will be stored in the courthouse security database, and her original name badge will be printed upon providing identification. To access the records maintained at the Supreme Court, she will need to sign in once she reaches the records room.

Most courts provide copies of records for a nominal fee. The New Jersey state courts will provide records "only in the form in which they are maintained or indexed by the Judiciary" and state in their official rules that "[r]equests by private individuals or entities for programming, searching, or compilation of records in a form other than as used for the Judiciary's purposes will not be granted."[76] Thus, in New Jersey our hypothetical record-seeker could, for a small fee,[77] obtain a copy of any publicly available court record in the format in which it is kept. She could not, however, request that the clerk provide her with a compilation of all records involving civil suits with damage awards greater than $50,000, or all records from cases involving noise ordinances.[78] Once she obtains a copy of the publicly available records, she is free to use them however she chooses, provided that she does not violate any state or federal laws.[79]

mercial entity has the same rights to judicial records as anyone else, "the compilation of otherwise hard-to-obtain information alters the privacy interest implicated by disclosure of that information" (citation and internal quotation marks omitted); *see also infra* Part V (discussing balancing tests and electronic access).

75. *See infra* Part IV (comparing in-person access to court records in New Jersey to electronic access).

76. N.J. R. 1:38-13.

77. *Id.* New Jersey Courts Rule 1:38-9 states that "[t]he Supreme Court shall establish a schedule of fees for copies of records." *Id.* at 1:38-9.

78. *See id.* at 1:38-13 (stating that records are only available in the form in which they are "maintained or indexed" by the judiciary, which implies that compilations of records from multiple cases of the same type are not available for public access).

79. Certain uses of information could, for example, result in the commission of a privacy tort. *See* Cox Broad. Corp. v. Cohn, 420 U.S. 469, 474, 487 (1975) (discussing the individual right of privacy and court actions associated with that right with respect to information contained in publicly accessible records).

c. What May Be Accessed: Information Type

Explicit Inclusions: Each state court may indicate via statute what material is and is not included in the publicly accessible court record.[80] New Jersey Courts Rule 1:38, "Public Access to Court Records and Administrative Records," provides a definition of the term "court record" based on what it does and does not include. As stated at the beginning of this rule, New Jersey court records include:

> (1) any information maintained by a court in any form in connection with a case or judicial proceeding, including but not limited to pleadings, motions, briefs and their respective attachments, evidentiary exhibits, indices, calendars, and dockets;
> (2) any order, judgment, opinion, or decree related to a judicial proceeding;
> (3) any official transcript or recording of a public judicial proceeding, in any form;
> (4) any information in a computerized case management system created or prepared by the court in connection with a case or judicial proceeding;
> (5) any record made or maintained by a Surrogate as a judicial officer.[81]

The mandate for federal courts is considerably more vague. Federal Rule of Civil Procedure 79 requires the clerk of each federal court to "keep a record known as the 'civil docket,'" which includes "papers filed with the clerk" as well as "appearances, orders, verdicts, and judgments," and "any other records required by the Director of the Administrative Office of the United States Courts with the approval of the Judicial Conference of the United States."[82] This rule is implemented via the PACER system, and discussed in Part IV.

Explicit Exclusions: Some information is excluded from the court record by definition. According to Rule 1:38-2(b), New Jersey court records do not include "information gathered, maintained or stored by a governmental agency or . . . unfiled discovery materials."[83] The rule also lists specific types of records that are excluded from public access, including, but not limited to, records pertaining to: juvenile

80. *Ex parte* Capital U-Drive-It, Inc., 630 S.E.2d 464, 469 (S.C. 2006) (discussing how public access "[r]estrictions may be based on a statute or the court's inherent power to control its own records and supervise the functioning of the judicial system").

81. N.J. R. 1:38-2(a).

82. FED. R. CIV. P. 79.

83. N.J. R. 1:38-2(b).

delinquency, victims of abuse or sexual assault, child placement, and state mandated drug treatment programs.[84] Other states, however, such as Massachusetts, do not provide for the automatic exclusion of personally identifying information of victims of sexual assault, so the burden is on individuals concerned about their privacy or the privacy of their family members to be aware of state law and, if necessary, to request that sensitive records be excluded with a showing of good cause.[85]

Some records that would otherwise be available to the public may instead be "sealed." Even before a document ever becomes part of the court record, a judge has the option to seal that document.[86] Thus, a judge may seal entire case files under certain circumstances[87] or only certain documents in the file, such as discovery materials[88] or financial records.[89] A judge may decide to seal a record on her own,[90] at the request of one party,[91] or at the request of both parties.[92] Even if a record is not sealed, some information may be redacted before filing.[93] Each party's lawyer is responsible for redacting sensitive material, such as Social Security and bank account numbers, from her party's filings *before* they become public.[94] In some cases, parties may request that the judge redact certain otherwise public information

84. *Id.* at 1:38-3(c)–(d)

85. *See supra* note 45 (noting that individuals must file formal requests with the judge in order to have certain information redacted from their record).

86. FED R. CIV. P. 5.2(d).

87. *See, e.g.*, Joy v. North, 692 F.2d 880, 893 (2d Cir. 1982) (finding that adjudications can be sealed, but only under "exceptional circumstances").

88. Seattle Times Co. v. Rhinehart, 467 U.S. 20, 37 (1984) (affirming the issuance of a protective order limited to pretrial civil discovery).

89. *See* United States v. Hickey, 997 F. Supp. 1206, 1208–09 (N.D. Cal. 1998) (noting that unsealing defendants' financial affidavits would result in "substantial and real hazards of incrimination").

90. *See In re* Knight Publ'g Co., 743 F.2d 231, 235 (4th Cir. 1984) ("The trial court . . . may, in its discretion, seal documents if the public's right of access is outweighed by competing interests.").

91. FED. R. CIV. P. 26(c)(1).

92. *See* Laurie Kratky Doré, *Secrecy by Consent: The Use and Limits of Confidentiality in the Pursuit of Settlement*, 74 NOTRE DAME L. REV. 283, 385–86 (1999) (explaining that "litigants might further condition their compromise on the sealing . . . of particular documents").

93. *See* Katie Mulvaney, *ACLU Objects to Redaction of Federal Court Records*, PROVIDENCE J. BULL., Dec. 7, 2010 (reporting that "[w]hile all federal courts are required to hide an individual's Social Security number, birth dates and minors' names from publicly accessible court records," the United States District Court for the District of Rhode Island considered a proposal that would allow parties "to ask the court to have information other than personal identifiers . . . redacted from court transcripts").

94. *See* FED. R. CIV. P. 5.2(a) (stating that "a party . . . making the filing may include only" certain information in the document).

based on privacy interests or concerns about trade secrets.[95] Depending on the court, federal or state law controls whether a record will be sealed or redacted.[96]

In federal court, the Federal Rules of Civil Procedure require parties to include only: the last four digits of their Social Security number, taxpayer identification number, or financial account number; the year of their birth rather than the entire birth date; and the initials of any minor identified in the case.[97] When other privacy concerns arise, the court may allow individuals to file certain documents under seal, but without redaction.[98] "The court may later unseal the filing or order [the party] to file a redacted version for the public record."[99] Alternatively, the court may grant a party's request that additional information be redacted from the record, or be made unavailable via remote access, upon a showing of good cause.[100]

Generally speaking, most discovery materials collected by the parties are excluded from the publicly available record, on the theory that providing access to them does not further the goal of ensuring honesty and respect in the judicial system.[101] "Discovery is traditionally a private aspect of litigation"—it happens not in a public forum but behind the scenes and before the trial—therefore, "public scrutiny would have little value."[102] Further, providing access to discovery materials may make an already cumbersome and expensive process even more so.[103] Of course, some discovery materials are admitted as evi-

95. *See* Peter A. Winn, *Judicial Information Management in an Electronic Age: Old Standards, New Challenges*, 3 FED. CTS. L. REV. 135, 138, 142 (2009) (noting that courts may use redaction to prevent access to trade secrets and "other sensitive information").

96. *See infra* Part II.C.

97. FED. R. CIV. P. 5.2(a). *But see id.* at 5.2(b) (discussing exemptions).

98. *Id.* at 5.2(d).

99. *Id.*

100. *Id.* at 5.2(e). Courts have interpreted "good cause" in a variety of ways. *See infra* Part II.C and text accompanying notes 144–159.

101. *See* Leucadia, Inc. v. Applied Extrusion Techs., Inc., 998 F.2d 157, 164–65 (3d Cir. 1993) (noting that "[t]he public policy implications of" affording a presumptive right of access to discovery materials "are unclear"); Anderson v. Cryovac, Inc., 805 F.2d 1, 12 (1st Cir. 1986) (finding that public access to discovery does not "play a significant role in the administration of justice").

102. Diane Apa, *Common Law Right of Public Access—The Third Circuit Limits Its Expansive Approach to the Common-Law Right of Public Access to Judicial Records:* Leucadia Inc. v. Applied Extrusion Techs., Inc., 39 VILL. L. REV. 981, 1000 (1994).

103. *See Anderson*, 805 F.2d at 12 ("Indeed if such cases were to be mandated, the civil discovery process might actually be made more complicated and burdensome than it already is.").

dence at the trial, at which point the presumption of access to these materials applies.[104]

Inclusions by omission: Even after taking account of procedural rules or state statutes regarding information that may be sealed, redacted, and excluded from the record, precisely what the general public is entitled to access in each court record remains somewhat of an open question. Generally speaking, courts have held that once something has become "public," that is, made available for public viewing or listening in some form, even if only to a portion of the public, it remains public for the purposes of the court record.[105] For example, the United States Court of Appeals for the Second Circuit has stated:

> Once the evidence has become known to the members of the public, including representatives of the press, through their attendance at a public session of court, it would take the most extraordinary circumstances to justify restrictions on the opportunity of those not physically in attendance at the courtroom to see and hear the evidence, when it is in a form that readily permits sight and sound reproduction.[106]

104. *See* Littlejohn v. Bic Corp., 851 F.2d 673, 683 (3d Cir. 1988) (holding that the "right of access [applies] to items that properly remained part of the judicial record, such as the deposition testimony read into evidence at trial").

105. *See* United States v. Martin, 746 F.2d 964, 967–69 (3d Cir. 1984) (finding that "[t]he public interest can best be vindicated by the release of complete and accurate transcriptions" of audiotapes played for the jury at trial and not simply by reporters' presence in the courtroom during the trial to take notes on the recordings). Note that with this move, courts are vastly expanding the meaning of "public" as it relates to court records. It may not be feasible or even desirable for the majority of the "public" to attend or otherwise listen to a trial, yet the record becomes available to everyone, including those who live far from the courthouse, especially if the record is later placed online. *See, e.g.*, United States v. Criden, 648 F.2d 814, 815 (3d Cir. 1981) (granting television networks' request "'for permission to copy, for the purpose of broadcasting to the public, those video and audio tapes admitted into evidence and played to the jury in open court'"); Lynn E. Sudbeck, *Placing Court Records Online: Balancing Judicial Accountability with Public Trust and Confidence: An Analysis of State Court Electronic Access Policies and a Proposal for South Dakota Court Records*, 51 S.D. L. REV. 81, 91 (2006) (noting that a "frequently mentioned benefit" of electronic access to court records is that it "is appropriate to the needs of South Dakota's rural court users, that is, [it] levels the geographic playing field" (citation and internal quotation marks omitted)).

106. *In re* Nat'l Broad. Co., 635 F.2d 945, 952 (2d Cir. 1980); *see also* Cox Broad. Corp. v. Cohn, 420 U.S. 469, 494–95 (1975) ("[T]he prevailing law of invasion of privacy generally recognizes that the interests in privacy fade when the information involved already appears on the public record. The conclusion is compelling when viewed in terms of the First and Fourteenth Amendments and in light of the public interest in a vigorous press.").

Thus, trial transcripts, most of the evidence introduced at trial, and, in some cases, settlement agreements,[107] have been found to qualify as part of the court record.[108] Even documents filed in conjunction with a denied motion for summary judgment may be subject to public access on the basis that denial of the motion "shaped the scope and substance of the litigation."[109] But once again, this all varies by jurisdiction and by individual court.

Exclusions by omission: In some instances, documents that might technically be considered part of the record because they were relied upon by the judge or jury, or introduced at trial, might nevertheless be excluded from the official court record to which the public has access.[110] Generally speaking, exhibits that are too cumbersome to be

107. *See, e.g.*, SEC v. Van Waeyenberghe, 990 F.2d 845, 849 (5th Cir. 1993) (holding that a settlement agreement filed in a district court is a judicial record to which the right of access applies); *see also In re* September 11 Litig., 723 F. Supp. 2d 526, 528, 531–33 (S.D.N.Y. 2010) (same); Bank of Am. Nat'l Trust & Sav. Ass'n v. Hotel Rittenhouse Assocs., 800 F.2d 339, 345 (3d Cir. 1986) (granting in part motion to unseal settlement agreement, reasoning that "[t]here is no persuasive reason to believe that the public or the press will mischaracterize the settlement information and construe it as an admission of liability" and that "[t]he Aviation Defendants' fear of public misperception is speculative and unlikely, and does not outweigh the presumption of public access to the settlement information"). *But see* Jessup v. Luther, 277 F.3d 926, 928–30 (7th Cir. 2002) (finding that settlement agreements are not judicial records and therefore not presumptively public, but that the particular settlement agreement at issue was a public document because it involved input from a federal judge and had already been made publicly available). The court held that the presumption of access generally does not apply to settlement agreements, pointing out that "[p]arties who settle a legal dispute rather than pressing it to resolution by the court often do so, in part anyway, because they do not want the terms of the resolution to be made public." *Id.* at 928. The court further noted that defendants may be particularly reluctant to disclose the terms of a settlement for fear that doing so might "encourage others to sue." *Id.* Because settlement agreements are not judicial records, "the issue of balancing the interest in promoting settlements by preserving secrecy against the interest in making public materials upon which judicial decisions are based does not arise—there is no judicial decision." *Id.*

108. *See, e.g.*, FTC v. Standard Fin. Mgmt. Corp., 830 F.2d 404, 409 (1st Cir. 1987) ("[W]e rule that relevant documents which are submitted to, and accepted by, a court of competent jurisdiction in the course of adjudicatory proceedings, become documents to which the presumption of public access applies"); *Bank of Am. Nat'l Trust & Sav. Ass'n*, 800 F.2d at 343 (finding that the common law presumption of access applied to a settlement agreement between the bank and a hotel developer); *Martin*, 746 F.2d at 968 (acknowledging a common law right of access to evidence introduced at trial, but noting that the right "is not limited to evidence" (citation omitted)). *But see Littlejohn*, 851 F.2d at 682 (holding that because exhibits would be destroyed by the clerk if not returned to the parties, they are no longer a part of the judicial record subject to public access).

109. Republic of the Philippines v. Westinghouse Elec. Corp., 949 F.2d 653, 660 (3d Cir. 1991).

110. *See* Poliquin v. Garden Way, Inc., 989 F.2d 527, 532–33 (1st Cir. 1993) (noting that post-trial access to documents introduced at trial can be restricted "only [upon] the most compelling showing").

kept at the clerk's office, or that are not in a form readily available for copying, might be omitted from the record.[111] In what is perhaps an extreme example, the Third Circuit noted in *Littlejohn v. Bic Corp.*[112] that "an animate object, such as a member of an endangered species," presents problems for the clerks, since "[i]n such circumstances, a continuing duty to retain trial exhibits is untenable because it would require the Clerk or the owner to maintain the endangered creature for the remainder of its life."[113]

A list of juror names will likely be omitted from the record, though this does necessarily mean that a curious member of the press or public could not find out who served.[114] In New Jersey, while a list of juror names is not disclosed,[115] the *voir dire* (jury selection) transcript is publicly available.[116] In most cases, the jurors will be called by their full names during *voir dire*, prior to receiving juror numbers.[117] Unless the transcript is sealed, juror names are publicly available in New Jersey, although these names may be difficult to ascertain if the final panel is not clearly identified in the transcript.

No matter how indirect, the inclusion of jurors' names in the record is particularly disconcerting.[118] As Peter Winn has pointed out, "Court records contain sensitive information about crime victims, about witnesses, about jurors, and about a host of other persons[, and] [t]he parties to a dispute may ignore the information privacy interests of third parties . . . ."[119] As explained above, the burden is almost always on the parties or their lawyers to redact sensitive or per-

111. *E.g.*, NEB. R. 6-501, 6-502; WASH. R. APP. P. 9.8(b).

112. 851 F.2d 673 (3d Cir. 1988).

113. *Id.* at 682 n.23.

114. *See* KEYES, *supra* note 25 (discussing a federal court policy that "mandates that documents containing identifying information about jurors . . . are no longer included in the public case file," but noting that jurors could be identified "through jury selection" or "simply by digging through the case file").

115. *Id. But see* N.J. STAT. ANN. § 2B:20-5 (West 2011) ("The list of names randomly selected from the juror source list shall be filed and publicly posted . . . .").

116. United States v. Antar, 38 F.3d 1348, 1351 (3d Cir. 1994); Barber v. Shop-Rite of Englewood & Assocs., Inc., 923 A.2d 286, 292–93 (N.J. Super. Ct. App. Div. 2007).

117. *See* N.J. STAT. ANN. § 2B:23-10 (West 2011) (stating that parties "may question any person summoned as a juror after the name is drawn and before the swearing"); *Antar*, 38 F.3d at 1351 (noting that during *voir dire*, "the members of the petit jury stated their names and hometowns on the record").

118. *See* Nancy J. King, *Nameless Justice: The Case for the Routine Use of Anonymous Juries in Criminal Trials*, 49 VAND. L. REV. 123, 127–28 (1996) (supporting the argument in favor of having jurors serve anonymously by noting that "enough harassment by opponents of verdicts takes place to keep many jurors worried").

119. Winn, *supra* note 95, at 151 (footnotes omitted).

sonal information.[120] For those individuals who are involved in a lawsuit but are not represented—either because they are acting *pro se* or because they are only collaterally involved in the action—it may be the case that no one remembers or bothers to redact their information.[121]

Before addressing the role of courts in making "good cause" determinations and balancing the public's right of access to court records against individual privacy concerns,[122] it bears noting that these carefully crafted judicial rules and state statutes, directing which parts of the case file become part of the publicly accessible court record, may at times, in practice, recede into the background.[123] Clerks and courthouse employees, overwhelmed with paperwork and record requests, are likely to rely just as much on custom and convenience as on rules handed down from higher courts, particularly when only a few records are requested at a time.[124] Often, it is not practicable for one or two employees to sift through hundreds or even thousands of records, removing Social Security numbers, financial data, or other information that has been left in the record by the parties or their lawyers.[125] When these records are not already in electronic

120. FED. R. CIV. P. 5.2(a).

121. *See* Peter A. Winn, *Online Court Records: Balancing Judicial Accountability and Privacy in an Age of Electronic Information*, 79 WASH. L. REV. 307, 312, 321 (2004) (noting that "courts are sensitive to protect . . . the harm that can come to . . . third parties, who may have no control over the information so disclosed[,]" and who may have "never intended" that their information would be easily accessible in an electronic record).

122. *See infra* Part II.C.

123. *See* Sudbeck, *supra* note 105, at 89–90 (explaining that "[t]raditionally, access issues have been determined by judges . . . or by the control the clerk exercises, . . . as custodian of these records" and that without an "extensive body of case law to draw from, court administrators, not judges, find they must develop policies for access to records that will protect private and sensitive information").

124. *See* Martin, *supra* note 4, at 874–76 (explaining that a statewide uniform system of case management "is a distant prospect" in many states and using California as an example of the negative effects of this "decentralization" because of the vastly different approaches that exist in courts throughout the state). When bulk record access is requested, this is more likely to come to the attention of the judges at the courthouse. *See, e.g.*, Burnett v. County of Bergen, 968 A.2d 1151, 1164 (N.J. 2009) (holding that "bulk disclosure of realty records to a company planning to include them in a searchable, electronic database" was acceptable if plaintiffs paid for redaction of Social Security numbers); Higg-A-Rella, Inc. v. County of Essex, 660 A.2d 1163, 1166 (N.J. 1995) (holding that plaintiffs have a common-law right to obtain electronic tax-assessment data for "every parcel of real estate in each of the county's municipalities"). In addition, N.J. R. 1:38-7(f) states that "[a]ny request for the mass release, in bulk, of electronically stored or microfilmed records containing Social Security numbers must be submitted to the Administrative Director of the Courts. A fee may be charged for the cost of redacting Social Security numbers from such records."

125. *See* Winn, *supra* note 121, at 320–21 (arguing that "case files often contain private or sensitive personal information" for which there will be no protection when the files be-

form, the job of removing sensitive information becomes much more burdensome.[126] When they *are* in electronic format, court clerks may in some instances, without oversight, decide to simply place all the records online to avoid having to complete paper requests at the courthouse and to provide greater accessibility to interested parties.[127] Thus, court decisions addressing what parts of the case file are and are not available for public consumption should be treated as only part of the equation that will determine the actual level of access at the moment that a request is made.

C. Balancing the Presumption of Access Against Other Considerations

While "there is a 'strong presumption in favor of public access to judicial proceedings,'"[128] it may be overcome based on the following six factors:

> (1) the need for public access to the documents at issue;
> (2) the extent of previous public access to the documents;
> (3) the fact that someone has objected to disclosure, and the identity of that person;
> (4) the strength of any property and privacy interests asserted;
> (5) the possibility of prejudice to those opposing disclosure; and
> (6) the purposes for which the documents were introduced during the judicial proceedings.[129]

come accessible electronically); Schwartz, *supra* note 14 (discussing how a spokeswoman from the Administrative Office of the U.S. Courts stated that "courts comb through the documents on a regular basis" but that a search of federal court documents from Washington state found that "thousands of documents" had not been properly redacted (internal quotation marks omitted)); *see also* N.J. R. 1:38-7(c)(1) ("[P]arties shall certify in the Case Information Statement that all confidential personal identifiers have been redacted and that subsequent papers submitted to the court will not contain confidential personal identifiers in accordance with the provisions of this rule.").

126. *See supra* text accompanying notes 41–42.

127. *See* Jennifer 8. Lee, *Dirty Laundry, Online for All to See*, N.Y. TIMES, Sept. 5, 2002, at G1, *available at* http://www.nytimes.com/2002/09/05/technology/dirty-laundry-online-for-all-to-see.html (discussing how policies at the local levels for online court documents "have tended to be murky or nonexistent" and can "sometimes fall[] to a single person" and providing as an example the clerk of courts for Hamilton County, Ohio, who, with his technology staff, created a website to provide online access to documents, such as "[s]tate tax liens, arrest warrants, [and] bond postings," which were already scanned in electronic form).

128. *In re* Sealed Case, 237 F.3d 657, 666 (D.C. Cir. 2001) (quoting Johnson v. Greater Se. Cmty. Hosp. Corp., 951 F.2d 1268, 1277 (D.C. Cir. 1991)).

129. EEOC v. Nat'l Children's Ctr., Inc., 98 F.3d 1406, 1409 (D.C. Cir. 1996) (citing United States v. Hubbard, 650 F.2d 293, 317–22 (D.C. Cir. 1980)).

Each of these factors can trigger a balancing inquiry in which a court weighs the public's interest in open access against countervailing asserted interests.[130]

There appears to be a general agreement among courts that the right of access to court records ensures a well-informed public,[131] provides an "educational and informational benefit" to the citizenry,[132] and allows the public to monitor the functioning of the courts,[133] thereby providing an additional check on the judicial system and increasing judges' accountability.[134] According to the Second Circuit, the right of access to judicial documents "is based on the need for federal courts, although independent—indeed, particularly because they are independent—to have a measure of accountability and for the public to have confidence in the administration of justice."[135]

As we have explained above, items such as sealed discovery materials[136] and financial information[137] submitted by the parties are generally not included in the court record, due primarily to privacy concerns and the belief that allowing access to such documents would not serve any of the purposes of public access.[138] As an Eastern District of Pennsylvania court stated in *Zenith Radio Corp. v. Matsushita Electric Industrial Co.*:

130. *See id.* at 1408–10 (reversing the district court's decision to seal a consent decree after weighing the six factors and finding "that on balance the nature of the services provided by the Center as well as the Center's receipt of public funding cuts against rather than in favor of sealing the record").

131. United States v. Mitchell, 551 F.2d 1252, 1258 (D.C. Cir. 1976) ("[T]he right of inspection serves to produce 'an informed and enlightened public opinion'" (quoting Grosjean v. Am. Press Co., 297 U.S. 233, 247 (1936)), *rev'd*, Nixon v. Warner Commc'ns, Inc., 435 U.S. 589 (1978).

132. United States v. Criden, 648 F.2d 814, 829 (3d Cir. 1981).

133. *Mitchell*, 551 F.2d at 1258; *see also In re* Cont'l Ill. Sec. Litig., 732 F.2d 1302, 1313–14 (7th Cir. 1984) (noting that "factors weighing in favor of public disclosure of court documents" include "'the public's interest in assuring that the courts are fairly run and judges are honest'" (quoting Crystal Grower's Corp. v. Dobbins, 616 F.2d 458, 461 (10th Cir. 1980))).

134. *See* Globe Newspaper Co. v. Superior Court, 457 U.S. 596, 606 (1982) (discussing the check on accountability in the context of criminal trials).

135. United States v. Amodeo, 71 F.3d 1044, 1048 (2d Cir. 1995).

136. *See, e.g.*, Leucadia, Inc. v. Applied Extrusion Techs., Inc., 998 F.2d 157, 163–65 (3d Cir. 1993) (holding that the common-law right of access does not extend to sealed discovery materials).

137. *See, e.g.*, United States v. Lexin, 434 F. Supp. 2d 836, 849 (S.D. Cal. 2006) (finding "that the documents containing Defendants' personal financial information submitted to support their Requests for Appointment of Counsel, are not judicial documents" and therefore are not subject to public disclosure).

138. *See supra* Part II.B.2.c; *see also Leucadia*, 998 F.2d at 164 (noting that discovery was not open to the public at common law and is conducted in private as a matter of modern practice) (citing Seattle Times Co. v. Rhinehart, 467 U.S. 20, 33 (1984)).

> [W]hether various kinds of materials are part of the judicial record is a question that must be answered for each in light of the purposes served by the common law right to inspect and copy public records[,] [b]earing in mind that access rights exist to promote knowledge of and attention to the performance of the courts, for the benefit of society as a whole . . . .[139]

The *Zenith Radio* court went on to suggest that the determinations of whether particular information is considered part of the court record to which the presumption of public access applies must be made on a case-by-case basis, considering "the magnitude and imminence of the threatened harm from disclosure, including the particularized interests of the litigants against disclosure . . . and the availability of less restrictive alternatives."[140] The language courts use in balancing these competing interests is broad and malleable; judges often make statements, for instance, that a district court should deny public access only if, "after weighing the interests advanced by the parties in light of the public interest and the duty of the courts . . . justice so requires."[141] In the context of criminal trials, the United States Supreme Court has held that certain "features of the criminal justice system" weigh heavily in favor of public access to criminal trials.[142] In order to overcome this presumption of access and "inhibit the disclosure of sensitive information," the moving party must show that the denial of access "is necessitated by a compelling governmental interest, and is narrowly tailored to serve that interest."[143]

When a party moves to have all or a portion of the record sealed, the court must decide whether to grant this motion by weighing the

139. 529 F. Supp. 866, 898 (E.D. Pa. 1981); *see also Amodeo*, 44 F.3d at 145 ("[T]he mere filing of a paper or document with the court is insufficient to render that paper a judicial document subject to the right of public access. . . . [T]he item filed must be relevant to the performance of the judicial function and useful in the judicial process in order for it to be designated a judicial document.").

140. *Zenith Radio Corp.*, 529 F. Supp. at 912 (footnote omitted).

141. *In re* Nat'l Broad. Co., 653 F.2d 609, 613 (D.C. Cir. 1981) (footnote and internal quotation marks omitted).

142. Globe Newspaper Co. v. Superior Court, 457 U.S. 596, 605–06 (1982) (noting that "the right of access to criminal trials" (1) is firmly rooted in our country's legal history and (2) "plays a particularly significant role in the functioning of the judicial process").

143. *Id.* at 606–07. Courts do not necessarily follow this in practice when deciding whether to seal all or part of the court record. *See* Ken Armstrong et al., *The Cases Your Judges Are Hiding From You*, SEATTLE TIMES, Mar. 5, 2006, at A1, *available at* http://seattletimes.nwsource.com/html/yourcourtstheirsecrets/2003595519_webseal05mNEW.html (discussing a survey of civil suits filed in the King County Superior Court in Washington state and arguing that "judges have displayed an ignorance of, or indifference to, the legal requirements for sealing court records").

party's interest in confidentiality against the public interest in access to court records. In both state and federal court, a party may move to limit access to a portion of or even an entire record upon a showing of "good cause."[144] According to the New Jersey rules of court:

> Good cause to seal a record shall exist when: (1) [d]isclosure will likely cause a clearly defined and serious injury to any person or entity; and (2) [t]he person or entitys [sic] interest in privacy substantially outweighs the presumption that all court and administrative records are open for public inspection . . . .[145]

The Federal Rules of Civil Procedure, as well as many states, do not define "good cause" but have left it to the courts to interpret.[146]

State and federal courts have taken a variety of approaches in balancing "good cause" to seal court records against the presumption of public access, from requiring that court proceedings "be open to the fullest public scrutiny," unless the law provides otherwise,[147] to suggesting (rather ambiguously) that the public "right of access may be overcome when a sufficiently compelling interest for nondisclosure is identified."[148] Courts frequently emphasize the role of judicial discretion in determining on a case-by-case basis whether a particular record, or a portion thereof, is to be sealed, rather than identifying whole categories of information to exclude in advance, or laying down bright-line rules stating what is and is not to be included in the court record.[149]

144. N.J. R. 1:38-11(a) (allowing information to be sealed for good cause); FED. R. CIV. P. 5.2(e) (permitting redaction and limitations on "a nonparty's remote electronic access" for good cause).

145. N.J. R. 1:38-11(b).

146. FED. R. CIV. P. 5.2(e) (requiring good cause for redactions and limitations on nonparties' remote access); OR. REV. STAT. § 135.873(2) (2009) (sealing records when seeking to obtain protective orders); VT. STAT. ANN. tit. 33, § 5119(a)(3) (Supp. 2011) (sealing records in juvenile cases); CAL. R. APP. P. 8.46(e)(2) (sealing records on appellate review). Furthermore, the burden of persuasion is on the moving party in federal court. *See* Bank of Am. Nat'l Trust & Sav. Ass'n v. Hotel Rittenhouse Assocs., 800 F.2d 339, 344 (3d Cir. 1986) ("The burden is on the party who seeks to overcome the presumption of access to show that the interest in secrecy outweighs the presumption.").

147. N.Y. Post Corp. v. Leibowitz, 143 N.E.2d 256, 258 (N.Y. 1957).

148. Associated Press v. State, 888 A.2d 1236, 1245 (N.H. 2005).

149. *See* Nixon v. Warner Commc'ns, Inc., 435 U.S. 589, 599 (1978) ("The few cases that have recognized [the public's right to access court records] do agree that the decision as to access is one best left to the sound discretion of the trial court, a discretion to be exercised in light of the relevant facts and circumstances of the particular case."). *But see* N.J. R. 1:38-1 ("Court records and administrative records . . . within the custody and control of the judiciary are open for public inspection and copying except as otherwise provided in

By and large, courts are reluctant to seal records in the face of the strong presumption of open access. Thus, when an individual or her lawyer seeks to deny the public access to a particular record, she must demonstrate not only good cause, but "'an overriding interest based on findings that closure . . . is narrowly tailored to serve that interest.'"[150] While a court may agree to seal a particular document in the record that contains a great deal of personally identifiable information, many courts are reluctant to seal the entire record, because doing so imposes "a burden on the courts and [is] an extreme inconvenience to attorneys."[151] Even when adverse parties stipulate to the sealing of records, "the [c]ourt must weigh any interests in confidentiality against that of the public to open court records."[152] Additionally, "[a] record should be sealed in its entirety only in extremely limited situations."[153] Courts frequently cite a fear of intruding on a constitutionally protected right to access court records, extra burdens placed on the courts, and inconvenience to attorneys as reasons to seal records only in very few circumstances.[154] Claims that a court deems less than compelling will generally not provide good cause to seal otherwise open records.[155] The Colorado Court of Appeals has suggested that because injury to one's reputation is an inherent risk

this rule. Exceptions enumerated in this rule shall be narrowly construed in order to implement the policy of open access to records of the judiciary.").

150. Publicker Indus., Inc. v. Cohen, 733 F.2d 1059, 1071 (3d Cir. 1984) (quoting Press-Enterprise Co. v. Superior Court, 464 U.S. 501, 510 (1984)); *see also* Globe Newspaper Co. v. Superior Court, 457 U.S. 596, 606–07 (1982) (holding that even in criminal cases, if "the State attempts to deny the right of access in order to inhibit the disclosure of information, it must be shown that the denial is necessitated by a compelling governmental interest, and is narrowly tailored to serve that interest").

151. Davis v. Davis, 997 So.2d 149, 161 (La. Ct. App. 2008); *see also In re* Orion Pictures Corp., 21 F.3d 24, 26 (2d Cir. 1994) (finding that "on a purely practical level, the sealing of court records inflicts a costly nuisance on the judicial system . . . [and] impose[s] substantial burdens on the clerk's office and on a judge's staff" (citation omitted)).

152. Carty v. Virgin Islands, 203 F.R.D. 229, 230 (D.V.I. 2001).

153. *Davis*, 997 So.2d at 161.

154. *Id.* at 160. When the reputations of unrepresented third parties are at stake, the balancing may—or perhaps should—come out differently. *See supra* Part II.C (discussing the privacy concerns of unrepresented third parties).

155. *See, e.g.*, Kamakana v. City of Honolulu, 447 F.3d 1172, 1183 (9th Cir. 2006) (holding that claimed reliance by a non-party on a stipulated protective order was not a compelling reason sufficient to overcome the presumption of public access); Va. Dep't of State Police v. Washington Post, 386 F.3d 567, 580 (4th Cir. 2004) (holding that the State Police offered no compelling reason to keep documents under seal).

in almost every civil suit, such concerns should not overcome the strong presumption in favor of open access to court records.[156]

Some circumstances so predictably raise sufficient privacy concerns that courts will almost always grant a party's request to seal the record. For example, courts are likely to find good cause to seal records when children or victims of sexual assault are involved,[157] or when the records in question were previously sealed and used only during discovery.[158] In other words, a judicial finding of good cause often turns on the very same factors that the New Jersey court accounts for statutorily in Rule 1:38. The reasons for this are relatively straightforward: While encouraging people to use the justice system is frequently cited as a reason for keeping court records *open*, in cases where good cause has been found to seal records, judges often note that sealing the record will actually make people more willing and likely to use the system by allaying their fears about the exposure of sensitive personal information or business trade secrets.[159]

It bears noting that when courts weigh the public's interest in access against an individual's interest in privacy, they tend to focus on issues of defamation,[160] disclosure of trade secrets,[161] and, in the case

156. Doe v. Heitler, 26 P.3d 539, 544 (Colo. App. 2001) (finding "that a court file contain[ing] extremely personal, private, and confidential matters is generally insufficient to constitute a privacy interest warranting the sealing of the file").

157. Application of Lascaris, 319 N.Y.S.2d 60, 62–63 (N.Y. Sup. Ct. 1971) (holding that court files relating to a child neglect case were to be kept confidential to protect the parties and to encourage resolution of family problems through legal means).

158. Leucadia, Inc. v. Applied Extrusion Techs., Inc., 998 F.2d 157, 163 (3d Cir. 1993) ("'[T]here is no tradition of public access to discovery, and requiring a trial court to scrutinize carefully public claims of access would be incongruous with the goals of the discovery process.'" (quoting Anderson v. Cryovac, Inc., 805 F.2d 1, 13 (1st Cir. 1986))).

159. *See Lascaris*, 319 N.Y.S.2d at 63 (sealing the record in a child-neglect case "to encourage the family resolution of problems of the family through legal means," which is "not likely to happen unless there is assurance that confidences will be kept"). *But see* White v. Worthington Indus., Inc. Long Term Disability Income Plan, 266 F.R.D. 178, 196 (S.D. Ohio 2010) (denying the plaintiff's motion to seal records containing medical information because the court "relied on portions of [these records] in making its decision [and] the public's entitlement to view court documents upon which a court's decision rests ordinarily outweighs any privacy interest in such records"). In *White*, the court noted that "[t]he plaintiff who seeks such redress ordinarily understands that in order to do so, he or she may be waiving the right to keep his or her medical history out of the public domain, and Ms. White's opposing memorandum acknowledges that she appreciates this eventuality." *Id.* Thus, while recognizing that this principle undermined the plaintiff's interest in privacy, the court nonetheless found in favor of public access based, at least in part, on the fact that the plaintiff chose to undertake this risk when she began the litigation. *Id.*

160. *See In re* Food Mgmt. Grp., LLC, 359 B.R. 543, 561 (Bankr. S.D.N.Y. 2007) (stating that the court must seal court records if the interested party "can demonstrate that the allegations in the complaint are in fact scandalous and defamatory"); *cf.* Brown & William-

of special victims, physical or emotional safety.[162] What is rarely, if ever, included in this balancing are concerns about the kind of personal information discussed in Part II.A. This information can be easily aggregated and linked to particular individuals by companies that traffic in the free collection and highly profitable sale of large amounts of personal data,[163] or worse, by identity thieves. Since the scale already tips so far in favor of open access, it would be hard to imagine a judge finding that the risk of disclosing an unredacted Social Security number outweighs the strong presumption of public access.[164] This compounds the risk posed by the vast amount of personal information readily available to anyone with an interest in finding it—whatever their intentions.

III. CONTEXTUAL INTEGRITY

In the preceding Part, we provided an overview of the legal right of access to court records, as well as recognized grounds, such as privacy, for curtailing access. This landscape is where we begin our analysis of whether and how these rights and corresponding obligations

son Tobacco Corp. v. FTC, 710 F.2d 1165, 1179 (6th Cir. 1983) (finding that harm to a company's reputation does not by itself justify sealing the record); Nicklasch v. JLG Indus., Inc., 193 F.R.D. 570, 574 (S.D. Ind. 1999) (finding that possible embarrassment to a company is not enough to seal court records).

161. *See* Valley Broad. Co. v. U.S. Dist. Ct., 798 F.2d 1289, 1294 (9th Cir. 1986) ("Counseling against [public] access would be the likelihood of an improper use, 'including publication of scandalous, libelous, pornographic, or trade secret materials . . . .'" (quoting United States v. Criden, 648 F.2d 814, 830 (3d Cir. 1981) (Weis, J., concurring and dissenting))); *In re* Iowa Freedom of Info. Council, 724 F.2d 658, 664 (8th Cir. 1983) (holding that a private party's property interests in its trade secrets sufficed to override the public's right of access because "trade secrets partake of the nature of property, the value of which is completely destroyed by disclosure").

162. *See* United States v. Cojab, 996 F.2d 1404, 1408 (2d Cir. 1993) ("[W]e have recognized as additional sufficient reasons for closure and sealing those occasions where . . . publicity might put at risk the lives or safety of government agents engaged in undercover activities."); Gannett Co. v. Burke, 551 F.2d 916, 916 (2d Cir. 1977) ("[I]t may be proper for a court to seal certain records or papers, the revelation of which might, for example, endanger a witness's safety[.]").

163. Helen Nissenbaum, *Privacy as Contextual Integrity*, 79 WASH. L. REV. 119, 121 (2004) [hereinafter Nissenbaum, *Contextual Integrity*] ("Personal data is the 'gold' of a new category of companies, like Axciom, that sell this information, sometimes organized by individual profiles, to a variety of parties . . . .").

164. *See, e.g.*, Ostergren v. Cuccinelli, 615 F.3d 263, 286–87 (4th Cir. 2010) (holding that Virginia could not punish, under its privacy statute, an individual who obtained land records with unredacted Social Security numbers and published the records online).

are affected by the placement of digitized records online, in particular, through the lens of contextual integrity.[165]

The theory of contextual integrity accounts for a right to privacy in personal information (that is, information about persons) in terms of appropriate flow.[166] Instead of characterizing privacy as control over personal information, or as the limitation of access to information, it characterizes privacy as conformance with appropriate flows of information, in turn modeled by the theoretical construct of context-relative (or context-specific) informational norms. When information is captured or disseminated in ways that violate informational norms, privacy as contextual integrity is violated.

The theory is contextual because social contexts are taken as organizing principles of social life. As such, people act and interact not simply as individuals in an undifferentiated social world but as individuals in certain capacities in a plurality of distinct social realms. These realms, or contexts, are structured social settings characterized by distinct configurations of roles, activities and practices, purposes and values, and context-specific norms prescribing expected behaviors.[167] Living in modern industrial societies, familiar contexts include healthcare, the marketplace, finance, politics, religion, education, friends, and home life.

Of special relevance to contextual integrity is a subset of norms—those that govern the flows of personal information and shape our expectations of privacy. The theory labels these context-relative informational norms and claims that their typical form is characterized by three key variables: actors (subjects, senders, and recipients),[168] information types,[169] and transmission principles.[170] Accordingly, con-

165. *See generally* NISSENBAUM, PRIVACY IN CONTEXT, *supra* note 21 (setting out the theory and framework of contextual integrity).

166. The "flow" of information is defined as the "transmission, communication, transfer, distribution, and dissemination—from one party to another, or others." *Id.* at 140.

167. The idea of social life as constructed from differentiated realms is not invented in the theory of contextual integrity. Rather, it is drawn and generalized from rigorous work by social theorists and philosophers who have put forth similar ideas, including domains, institutions, and fields in their respective theories.

168. Information *senders* and *recipients* form the two poles of information transmission—one sends, the other receives. These may be individuals, groups, or entities like organizations or committees. NISSENBAUM, PRIVACY IN CONTEXT, *supra* note 21, at 141. Information *subjects*, however, are generally individuals. *Id.* They are the actors to whom the information refers, quite literally, the *subjects* of the information transferred.

169. Information types, or *attributes*, refer to the nature of the information in question: not only who it was about, and to whom and from whom it was shared, but what it was about. *Id.* at 143.

text-relative information norms prescribe flows of personal information based on the type of information in question, the capacities in which subjects, senders, and recipients of information are acting, and the terms under which information is conveyed—also called "transmission principles." Transmission principles define constraints on information flows.[171] Although not typically explicated in other approaches to privacy, this variable is significant in dictating whether information may be shared only voluntarily; with due notice; bought, sold, or shared without cost; shared in confidence; entrusted to a third party for use only in an information subject's interest; extracted under duress; shared only pursuant to a warrant; and so on.[172]

It is important to note as a point of comparison between contextual integrity and other privacy theories that all the variables matter. This means, for example, that no information is totally private or totally public but is always defined relative to the actors and transmission principles.[173] (Whenever we omit mention of any of the variables, it is usually possible to spell these out; unfortunately elliptical speech is sometimes the root of much misunderstanding over the nature of privacy expectations.) When norms are respected, privacy as contextual integrity is preserved. It is when an action or practice contravenes a norm in one or more of the variables that our privacy expectations are violated. Fears and complaints that privacy has been violated frequently follow the deployment of novel socio-technical systems, particularly those involving information technology and digital media, because these have been frequent sources of radical disruptions in information flows.[174]

If we conclude that particular actions or practices violate rights to privacy based solely on whether they contravene entrenched norms,

170. Transmission principles function to constrain "the flow (distribution, dissemination, transmission) of information from party to party in a context. The . . . terms and conditions under which such transfers ought (or ought not) to occur." *Id.* at 145.

171. For a discussion of transmission principles, see *id.* at ch. 7.

172. *Id.*

173. Nissenbaum, *Contextual Integrity*, *supra* note 163, at 139 (adding that even in the most public of places, there may be an appropriate expectation of privacy relative to informational norms).

174. Information technology in particular is inherently "socio-technical" and is guided by the social and cultural norms of society. *See* Lauren Gelman, *Privacy, Free Speech, and "Blurry-Edged" Social Networks*, 50 B.C. L. REV. 1315, 1319 (2009) (discussing the problems of protecting privacy in the age of social networks and the prevalence of online divulgence of personal information); Katherine J. Strandburg, *Home, Home on the Web and Other Fourth Amendment Implications of Technosocial Change*, 70 MD. L. REV. 614, 673 n.290 (2011) (describing, for example, privacy disputes arising from changes to Facebook and Google settings as socio-technical changes).

the resulting account would be unhelpfully conservative. Although it would provide a refined way to detect and identify change, the account would offer no way to distinguish good change from bad, or desirable, legitimate change from undesirable, unjustifiable change. Such an account would have limited explanatory and normative power. The theory of contextual integrity addresses this problem with an evaluative layer that allows an action or practice to be judged on moral and political grounds, or, in the case of disruptive technologies and other engines of change that challenge norms, a comparison in moral and political terms between entrenched norms and disruptive practices.[175]

The evaluative layer of contextual integrity itself involves two analyses. The first calls for an inquiry into implications of a given information flow with particular attention to those with moral or political significance. A flow in question might cause harm, might unbalance a desirable power equilibrium, result in unfair discrimination, suppress liberty and autonomy, and so on.[176] Alternatively, it might do all these things to a greater or lesser extent than a novel flow or a flow prescribed by entrenched norms with which it is compared. Here, we draw on a great body of scholarly work discussing the value of privacy to individuals and societies.[177]

A second analysis considers the comparative impacts of novel and entrenched flows in light of context-specific ends, purposes, and values. For example, if one of the purposes of a healthcare context is to reduce physical suffering and its value is to achieve this purpose without regard for wealth, norms would require that rich and poor alike be relieved of pain. Additionally, to reduce the chance of prejudicial treatment, norms might even require that information about wealth status not be shared with caregivers. The second analysis can also serve to break ties or resolve conflicts that occur in the first layer of analysis. Thus, while property rights of air travelers are curtailed by

175. NISSENBAUM, PRIVACY IN CONTENT, *supra* note 21, at 140 ("The framework of contextual integrity maintains that the indignation, protest, discomfit, and resistance to technology-based information systems and practices . . . invariably can be traced to breaches of context-relative informational norms. Accordingly, contextual integrity is proposed as a benchmark for privacy.").

176. For example, an unrestricted or readily available flow of information could lead to discrimination on the basis of sexual orientation, religious beliefs, and political affiliations, to name a few.

177. *See* NISSENBAUM, PRIVACY IN CONTEXT, *supra* note 21, at ch. 4 (describing ways in which to determine the value in privacy); *see also* Nissenbaum, *Contextual Integrity*, *supra* note 163, at 150 (noting that privacy has great importance to both individuals and society, serving common and collective, as well as individual, purposes).

luggage searches to protect the security of other travelers, also at stake is the smooth functioning of air travel overall. All three factors, among others, are relevant to the evaluation of norms that, in this case, are embodied in explicit policy and regulation.

Applying the framework of contextual integrity to online placement of court records means compiling an account of changes in information flows and evaluating them in moral and political terms. Few would claim that online access to court records would have no impact on information flows—indeed, it is precisely the prospect of such changes that accounts for much of the enthusiastic support for online access.[178] Some of the support seems bluntly to deny any such change in flow, asserting that since "public is public," a transformation from local access to online access is merely doing the same thing more efficiently.[179] But, the "thing" in question that stays the same is not the way information flows, as such; it is the normative commitment to transparency of government functioning through open access to court records.[180]

Detractors, too, recognize that online placement creates changes in flow. Although, to the best of our knowledge, no one has argued flatly against the integration of digital media into court records systems, many, in writing and in numerous public discussions including those convened by state courts, have expressed concerns over problems arising from increased exposure due to the transition from locally accessible records (some still in hardcopy form) to digital records accessible online.[181] In a New Jersey report, for example, this increased exposure was labeled "hyper-dissemination," implying that while dissemination is clearly a good thing, because court records often contain sensitive information, hyper-dissemination requires closer examination.[182] Similarly, others have cited the loss of practical ob-

178. *See* Martin, *supra* note 4, at 860–61 (detailing several developments that led to widespread support for online access to court records). One could imagine other reasons, for example, cost savings and the environment.

179. *See* Nissenbaum, *Contextual Intergity*, *supra* note 163, at 120–21 (asserting that placing previously available records online "is merely an administrative move towards greater efficiency. Nothing has changed, fundamentally").

180. *See* Martin, *supra* note 4, at 861 (acknowledging "the courts' historic commitment to transparency").

181. *See id.* at 882–84 (describing the public's concerns about privacy and security of court records "especially during a period of transition"); *see also* Michael Caughey, Comment, *Keeping Attorneys from Trashing Identities: Malpractice as Backstop Protection for Clients Under the United States Judicial Conference's Policy on Electronic Court Records*, 79 WASH. L. REV. 407, 407 (2004) (arguing that clients should be allowed to bring a malpractice claim against their attorney who fails to redact sensitive personal information).

182. REPORT ON PUBLIC ACCESS, *supra* note 15, at 14.

scurity as a cause for caution, pointing out that the obscurity of local records in practice has protected against unjustifiable privacy violations; they have suggested that policy and regulation may need to address this inadvertent loss.[183]

In a similar vein, our work seeks to understand and evaluate the changes brought about by online access. Where others have invoked concepts of hyper-dissemination and loss of practical obscurity to signal potential sources of trouble spots, the framework of contextual integrity offers a richer, more rigorous way of characterizing and evaluating the changes. In Part IV, we use context-relative informational norms to characterize the nature of change brought about by digitizing records and making them available online for public access. This involves modeling differences in flow in terms of the key parameters: actors and information type. In Part V, we evaluate the significance of these differences.

IV. COMPARING PATTERNS OF INFORMATION FLOW

In this Part, we systematically examine how digitizing court records and placing them online affects flows of personal information contained in these records. Following the framework of contextual integrity, we model flows in terms of two key parameters: actors and information type.[184] The relevant actors in this context are the *recipient*, who is searching for information in court records; the *subject*, whose personal information flows to the recipient; and the *sender*, the information retrieval system that stores and provides access to the court records.[185] We associate *costs* with flows of personal information, where cost is an abstraction of the time, money, and effort that the recipient has to expend to cause a flow of information. We characterize the effect of digitizing court records and placing them online in terms of the difference in cost of flows of personal information in the resulting systems as compared to paper and digital records available at courthouses. In prior work on contextual integrity, differences in flows caused by the introduction of new technology were characterized using a binary model of flows—in effect, identifying new flows

183. *Id.* at 34 ("While such personal identifiers in the past remained in the 'practical obscurity' of the clerk's office, with the advent of the Internet, records that are placed online are now available in an instant in one's office or home, anywhere in the world." (footnote omitted)); Martin, *supra* note 4, at 863 (describing the charge to governing bodies to ensure that privacy and security are maintained for online records).

184. For a description of transmission principles, see NISSENBAUM, PRIVACY IN CONTEXT, *supra* note 21, at ch. 7.

185. *Id.* at 141–42.

that arose and old flows that were no longer possible as a consequence of the introduction of the technology.[186] In contrast, here we develop the more general notion of cost differences in flows as a basis for characterizing the effect of the introduction of new technology. The cost model subsumes the binary model of flows: a flow that is *not* possible can be modeled by assigning to it an infinite cost, and a flow that *is* possible can be modeled by assigning to it zero cost.[187] We need this additional generality in modeling flows for our study of in-person court record access at the courthouse and online access to court records.

We begin in Part IV.A by presenting a model that captures key elements of information retrieval from court records. Each step in this model has associated costs that vary based on the characteristics of the information retrieval system (including the medium) and the behavior of the user (recipient) who is conducting the search. We arrive at this model by augmenting previously developed information-retrieval models with additional features that were needed to model actual searches of court records that we conducted.

Part IV.B provides an overview of our empirical study. Specifically, we searched court records using two systems that provide online access to court records: PACER and Google Scholar.[188] In addition, we conducted searches at two physical courthouses: the Superior Court County Clerk's office in Trenton, New Jersey, and the Superior Court of New Jersey, Hudson Vicinage. Each search followed the steps of our information-retrieval model,[189] thus providing evidence that the model is a suitable abstraction of the process of retrieving information from court records. The example searches capture realistic scenarios in which the user (recipient) is acting in a certain capacity and seeks to acquire a specific type of information about the subject. For example, in one scenario, the recipient is a potential employer of the subject and may be interested in knowing whether the subject was

186. *See id.* at 143–45 (explaining how some theorists have split information types into dichotomies).

187. *Id.*

188. GOOGLE SCHOLAR, http://scholar.google.com (last visited Mar. 7, 2012). Google Scholar provides a way to perform a broad search on all the legal opinions present in its database. "Currently, Google Scholar allows [users] to search and read published opinions of US state appellate and supreme court cases since 1950, US federal district, appellate, tax and bankruptcy courts since 1923 and US Supreme Court cases since 1791." GOOGLE SCHOLAR, SCHOLAR HELP, http://scholar.google.com/intl/en/scholar/help.html (last visited Mar. 7, 2012). In addition, it allows users "to find influential cases (usually older or international) which are not yet online or publicly available." *Id.*

189. For a discussion of the information-retrieval model, see *infra* Part IV.A.

involved in criminal cases in the last ten years and, if so, learning about the exact nature of the crime.[190] In two other scenarios, we consider a lawyer who seeks detailed information about a case,[191] and a data aggregator who is interested in building dossiers containing personal information about individuals.[192] Our findings support the hypothesis that the costs of retrieving various types of personal information differ significantly in the online and local access systems, provided that the user possesses certain types of background knowledge about the data subject.[193]

Finally, in Part IV.C, we use the model and the lessons learned from the empirical study to identify the root causes of the difference in cost of flows of personal information between online court records and records available at courthouses.

A. Information Retrieval Model

The goal of the eight-step model described below is to capture key elements of information retrieval from court records. The associated costs of each step vary according to characteristics of the information retrieval system and the behavior of users conducting the search. Later in this Part, we discuss in detail the differences in cost.[194]

1. *Information need:* Imagine a user who has an information-need about a data subject. The user seeks this information in a certain context in which she is acting in a certain capacity or role relative to the data subject. As a running example, consider a user who is interested in knowing whether a potential employee has been involved in criminal cases and, if so, learning about the exact nature of the crime.

190. *See infra* Part IV.B.1.a.

191. *See infra* Part IV.B.1.b.

192. *See infra* Part IV.B.1.b.

193. Caveat: We used PACER and Google Scholar in our empirical study of online systems to search for federal court records, but the physical courthouses in which we conducted our searches contained only court records for the state of New Jersey. Thus, we cannot directly compare the cost of the same search online and at the physical courthouse. This limitation is a consequence of the fact that online systems to access court records for the state of New Jersey, as well as many other states, are still in the process of being designed and are not yet available for public use. *See* Martin, *supra* note 4, at 872 (stating that "state court systems have been far slower and less coordinated in making th[e] transition" from paper records to electronic filing). Thus, "state courts seriously lag behind the federal courts in this area." *Id.* The searches that we conducted online and at the physical courthouses, respectively, are, nevertheless, still comparable in many respects, for example in how the background knowledge of users affects their findings. We hypothesize, therefore, that the comparison of costs remains meaningful despite this limitation.

194. *See infra* Part IV.C.

2. *Selecting information systems:* The user chooses a set of information systems to conduct her search based on her information need. For example, she may decide to use PACER or Google Scholar or to use a catalog at a physical courthouse or some combination of these systems to conduct her search. The system might require the user to provide information about the context, that is, the purpose of her search and her relation to the data subject.

3. *Getting to the information system:* The user gets to the location of the information system (online or at the physical location).

4. *Querying the system:* The user then formulates and issues a *query* to the system based on her *knowledge* and the *interface* presented to her. For example, she may search by name if she knows the name of the potential employee and the interface supports searches by the names of parties involved in court cases (as indeed PACER's interface does). In addition to restrictions on the nature of queries imposed by the interface (possibly based on the role of the user), social factors may impose additional constraints. A curious user may be more willing to search online through divorce records for details about a neighbor or acquaintance than to get this information from a courthouse, where she has to confront a clerk.[195]

5. *System's response to query:* The system processes the query over the indexed database of court records and returns a set of results. The model of a system, including the content and format of its database and data, determines how the system responds to various types of queries. For example, a system might index only a few types of information, such as party name and year, or it might index each term in the full text of the documents in the court record, allowing a user to search by any word present in the documents. In presenting certain types of information, the system may also take the relation of the user to the data subject into account. Confidential juvenile records, for instance, may be released only to the child or a named parent, and possibly to a caseworker if she can show cause (as determined by a judge).

6. *Evaluation of results:* The user evaluates the results to determine whether her information need has been met. Several factors affect the cost of this step, including the size of the set of results and the manner in which it is displayed. If the information need is indeed met, the search is over and the user moves on to Step 8. The user's

195. *See* Winn, *supra* note 95, at 152, 155, 157 (describing how the "practical obscurity" of a paper-based system of judicial records can present more obstacles and therefore discourage people without substantial interest from accessing information).

information need may not be met, for example, because the set of results for a search by a common name is too large and the user cannot figure out which records are about the person of interest to her. Although the user's information need is not met at this point, her knowledge may have increased, thus enabling her to formulate new queries. Further, at this point, the user's information need could change based on what she has learned thus far. In our running example, the user might discover that the potential employee was involved in a civil case and decide that information in that record might be relevant for her hiring decision.

7. *Query reformulation:* At this point, the user might *reformulate* her query based on her knowledge and choose a system to issue the new query. The user's decision may depend on the search interface offered by the system. For example, if the user wants to search by party name and city, she may want to use Google Scholar instead of PACER, since PACER does not allow queries of this form.[196] The control flow now goes back to Step 3 above and the search proceeds iteratively. The user can choose to terminate the search at any point, either because it seems that the information need will not be satisfied or because the cost of retrieving the desired information is too high.

8. *Subsequent use and dissemination:* The user uses the information acquired for a purpose and possibly disseminates it further. In our example, the user decides whether to offer the job to the applicant based on the information retrieved from court records. In other situations, the information acquired from court records could be disseminated further on the Web, via social networks, or through data aggregation services, leading to the creation of other sources that a future user could turn to in Step 2 while searching for personal information contained in court records. For example, data aggregators might build up new sources of information by accessing court records and sell this information at a cost that is much lower than a user might incur if she were to try to retrieve the information herself from court records. As another example, if court records were to be made available online in a form accessible to indexing by Web search engines (such as Google or Bing), future users could simply use search engines to find personal information in records at a very low cost.[197]

196. *See* PUBLIC ACCESS TO COURT ELECTRONIC RECORDS, http://www.pacer.gov/ (last visited Mar. 8, 2012).

197. In a recent French case, a man successfully sued Google after terms related to his past criminal record arose in a name search based on Google's Suggest feature. *See* Seth Weintraub, *French Court Convicts Google CEO Eric Schmidt of Defamation*, CNNMONEY (Sept. 26, 2010, 11:54 AM), http://tech.fortune.cnn.com/2010/09/26/french-court-convicts-

The model of information retrieval we describe above builds on prior work in the information retrieval and library-science literature, including approaches proposed by Bates;[198] Marchionini;[199] Belkin, Oddy, and Brooks;[200] and Hearst.[201] Although all of these models consider an iterative process similar to the one we have devised, our model considers three additional factors that are essential for the court-records setting and contribute to our empirical investigations and final analysis. First, our model allows that an information retrieval system might provide different views of the court records to users in different roles, for example, because of confidentiality requirements for certain types of information (such as the juvenile records discussed in Step 5 of our model). Indeed, unlike systems studied in the information retrieval literature whose goal is to make *all* types of information easily accessible, systems for retrieving information from court records have to be designed to make certain types of information easily accessible while ensuring that other types of information are not. Second, in Steps 2 and 7 of our model, we allow for the possibility that the user may acquire information from multiple information systems during the search process. While information retrieval systems are typically studied independently, court records are, in fact, accessible via multiple systems (PACER, Google Scholar, and emerging systems from various states). An analysis of these systems in isolation is woefully inadequate since a user can easily access all the systems simultaneously in the search process. Finally, by considering the

google-ceo-eric-schmidt-of-defamation. Google has stated in its policies that individuals concerned about inappropriate information in government records being made publicly available should contact the respective government agency or policymakers, and that it is the government's responsibility to ensure that records "are free of information that infringes an individual's right to privacy." GOOGLE, Webmaster Tools Help, About Privacy, http://support.google.com/webmasters/bin/answer.py?hl=en&answer=82301 (last visited Jan. 10, 2012). Google has further stated, "If information is made publicly available, we want to help our users discover and access it." *Id.* Consequently, if court records were to be made available online in an easily indexable format, the move would likely have a strong impact on dissemination of information and create new information flows.

198. Marcia J. Bates, *The Design of Browsing and Berrypicking Techniques for the Online Search Interface*, 13 ONLINE REV. 407 (1989).

199. GARY MARCHIONINI, INFORMATION SEEKING IN ELECTRONIC ENVIRONMENTS (1995).

200. N.J. Belkin, R.N. Oddy & H.M. Brooks, *ASK for Information Retrieval. Part I. Background and Theory*, *in* READINGS IN INFORMATION RETRIEVAL 299, 299–304 (Karen Sparck Jones & Peter Willett eds., 1997).

201. Marti A. Hearst, *User Interfaces and Visualization*, *in* MODERN INFORMATION RETRIEVAL 257–323 (Ricardo Baeza-Yates & Berthier Ribeiro-Neto eds., 1999); *see also* CHRISTOPHER D. MANNING ET AL., INTRODUCTION TO INFORMATION RETRIEVAL xiv (2008) (giving "an up-to-date treatment of all aspects of the design and implementation of systems for gathering, indexing, and searching documents and of methods for evaluating systems, along with an introduction to the use of machine learning methods on text collections").

potential to create new sources of information from the onward dissemination of acquired information, our model differs from many other information retrieval models that typically do not consider what happens to information once a user retrieves it. Yet this step is critical for an analysis of the cost of information flows from court records as discussed above in Step 8.

B. Overview of Empirical Study

To understand first-hand the cost of retrieving information from court records, we conducted an empirical study. Specifically, we searched court records using two systems that provide online access to court records—PACER and Google Scholar. In addition, we searched records at two physical courthouses—the Superior Court Clerk's office in Trenton, New Jersey, and the Superior Court of New Jersey, Hudson Vicinage, a trial court. We describe below these systems and the methodology we followed in conducting searches, noting that the steps in all these searches followed our eight-part information retrieval model. Our main finding is that between online and local access, there are significant differences in the cost of retrieving various types of personal information about a data subject. We note that in order to successfully retrieve personal information in either medium the user has to possess background knowledge relevant to the search. We elaborate on exactly what types of background knowledge are useful for low cost retrieval of personal information using court record systems and describe scenarios in which it is reasonable to expect that users will possess them.

1. Search Using Online Systems

a. Systems Used

We conducted a search of nineteen federal appellate court cases using PACER. In other words, we were searching for cases that had been appealed from a federal district court.[202] PACER, an electronic public access service provided by the United States judiciary, contains "case and docket information from federal appellate, district and bankruptcy courts and the PACER case locator via the Internet." The PACER case-locator interface supports search for records by various

202. The record in appellate cases contains only that on which the trial court (judge or jury) based its decision. FED. R. APP. P. 10(a). This may or may not contain everything in the court record that is publicly available at the trial court. FED. R. APP. P. 10(c)–(d).

fields, namely region, nature of suit[203] (in case of civil records), date when the case was filed, case title, party name, and case docket number. The advanced search interface additionally allows search by case title or range of dates for when the case was filed or closed. Once an initial query has been issued and the results are presented, a user can filter results by court type (for example, appellate or bankruptcy), court name (for example, Southern District of New York or Eastern District of Pennsylvania), year in which the case was filed, nature of suit (for example, contract, antitrust, labor, civil rights), or case title. A user can also browse through records based on the nature of the suit or look at all cases that have been filed or decided within a certain time period. In order to access the search interface, a user has to register with PACER and pay for searches. In addition to a name and address, a user must provide a credit card number at the time of registration or billing.

In conducting these searches, we used Google Scholar as an additional resource. Google Scholar allows users to perform a broad search across all the legal opinions archived in its database. "Currently, Google Scholar allows you to search and read published opinions of US state appellate and Supreme Court cases since 1950, US federal district, appellate, tax and bankruptcy courts since 1923 and US Supreme Court cases since 1791."[204] A user can retrieve information through a simple keyword search. There is no fee for this service. Users can search for opinions from state or federal courts. Within these opinions, users can search the full text of the document or filter the cases by year. It is also possible to see citation history. While Google Scholar is useful for locating judicial opinions and searching within them, unlike PACER it does not provide detailed information about each case and does not include a docket.

b. Observations from Search Experience

Set up for searches: In conducting our searches, we needed a "ground truth oracle" for deciding whether our search was successful. In other words, if we were to pick party names at random (for example, John Smith) and searched for information about them in court records, we would have no way of knowing for sure whether we found the correct record (that is, whether the record we found was about

203. For a list of the available "nature of suit" search categories, see http://www.pacer.gov/documents/natsuit.pdf.

204. *See* GOOGLE SCHOLAR, *supra* note 188.

the John Smith we searched for or some other person with the same name).

We used court opinions in Google Scholar as the ground truth oracle for all nineteen cases. Specifically, we retrieved fifteen court opinions directly from Google Scholar (chosen at random). We also retrieved two opinions each directly from Justia[205] and Westlaw and subsequently found those opinions also on Google Scholar. We used different types of information available in these court opinions as our background knowledge in conducting the searches using PACER (we elaborate on the exact types of background knowledge used in the next subsection). To verify whether a search on PACER was successful, we compared the party names, the case title, case number, and the court name in the final retrieved record on PACER with the information in the original court opinion in Google Scholar.

Search results: We conducted searches assuming that the user had different types of background knowledge about the data subject. Average generalized costs for these searches were very low: approximately $3 on PACER and under ten minutes each. We found court documents with various types of personal information, including names, complete addresses of parties, photographs, individuals' signatures, personal allegations, and medical information. Operationally, the goal of each search had been to retrieve the complaint form associated with each case.

In one set of eleven searches, we assumed that the user knew the complete name of the data subject. This level of knowledge is reasonable in many realistic scenarios: A potential employer, curious neighbor, and a housing application processor are but a few examples. In four of these eleven searches, the results set from the search contained exactly one court record and included an opinion that matched the Google Scholar opinion from which we had drawn the name, thus indicating that we found the correct record.

Upon revisiting the seven unsuccessful searches, we expanded the user's background knowledge to include not only the complete name of the data subject but also the type of case and the court where the complaint was first filed. It is reasonable to assume this level of knowledge in several scenarios, such as a potential employer (or an apartment owner) conducting a background check on an applicant. Since, for example, civil and appellate courts differ by geographic region, an employer can make a reasonable guess about the court where such cases involving the applicant were filed, if she knows the

205. JUSTIA, http://www.justia.com/ (last visited Mar. 8, 2012).

cities that the applicant has resided in;[206] information about where the applicant has lived could be inferred from her job or housing application. The court type to use in the search can be decided based on the type of case a user is interested in. For example, if an employer is checking whether an applicant has a criminal record, then the court type would be criminal, and if an employer or an apartment owner wants to check an applicant's financial background, then she would look at bankruptcy cases. In four out of these seven searches, the result set from the search contained only one record.

Finally, we revisited the three searches that were unsuccessful in the previous rounds and this time assumed that the user knew the complete name of the data subject, the court type, the court where the case was filed, and the range of years when the case could have been filed. The additional background knowledge about the range of years could be determined by the date of birth or the user may only be interested in the subject's criminal or bankruptcy records in the recent past.

However, the usefulness of this information depends on whether other records exist that correspond to the range of years in which a user is interested. For instance in one search, we were not able to narrow our results to find the desired record on PACER. For the other two searches, the result set was reduced to less than twenty-five records. To narrow the search results to exactly one record, we would have needed additional information about the data subject, such as the exact year that the case was filed.[207]

In another set of nineteen searches, we assumed that the user knew the case title and the case number corresponding to a record. In one case, we assumed knowledge of the case title alone. The case title contains the names of the parties involved[208] along with the name of the court and the decision year. This level of knowledge is reasonable in several realistic scenarios, including a lawyer who is searching

206. An exception to this line of reasoning may occur when, for example, the potential employee was involved in out-of-state litigation that did not involve any state where she has resided.

207. This additional information could also be in the form of Social Security numbers ("SSN") for bankruptcy cases. In the case of an owner checking a tenant's financial credibility or an employer looking up an employee's record, it is reasonable to assume that the user would know the SSN for the data subject being searched.

208. If the parties are individuals, the case usually contains all parties' first and last names. *See supra* Part II (describing exceptions to this and providing information on anonymous litigation). If one or more parties is a business or a government entity, the case title will list the name of the business or entity, and may also list the name of one or more associated individuals. FED. R. CIV. P. 17.

for details of a specific case, or a data aggregator who is browsing through court opinions on Google Scholar and using information retrieved from there to search for additional details about the parties on PACER. In eleven out of these nineteen cases, the last name in the case title was not sufficient to narrow down the results to one record. We then used the case title to search Google Scholar and learn the complete names of all the parties. The complete names helped us narrow our results to one record in eighteen out of our nineteen cases.

Our searches demonstrate that it is feasible to recover information at low cost if the user possesses a reasonable level of background knowledge.

2. Search Using Physical Systems

a. Systems Used

We conducted searches at two physical courthouses—the Superior Court Clerk's office in Trenton, New Jersey, and the Hudson Vicinage, New Jersey, trial court.

The Superior Court Clerk's office houses various documents including judgments, foreclosures, and liens for all trial courts in the state. To obtain copies of records, a user has to interact with a clerk and provide identification as well as a reason for the visit.[209] By using a computer terminal provided at the clerk's office, users can look for different types of documents by accessing several programs, such as Automated Case Management System ("ACMS")[210] to search for judgments by party name, docket number, venue, or judgment amount; Judiciary Electronic Filing System ("JEFIS"),[211] to search for foreclosures by docket number only; and "Appellate Division Search,"[212] to search for filings and decisions by party name, case number, organization name, phone number, zip code, motion num-

209. *Location of Court Records*, NEW JERSEY COURTS, http://www.judiciary.state.nj.us/superior/copies_court_rec.htm (last visited Mar. 8, 2012).

210. ACMS-AUTOMATED CASE MANAGEMENT SYSTEM, http://www.judiciary.state.nj.us/isd/acms.htm (last visited Mar. 8, 2012).

211. JUDICIARY ELECTRONIC FILING SYSTEM, http://www.judiciary.state.nj.us/jefis/index.htm (last visited Mar. 8, 2012).

212. *New Jersey Courts Search Page*, RUTGERS SCHOOL OF LAW, CAMDEN, http://lawlibrary.rutgers.edu/new-jersey-courts-search-page (last visited Mar. 8, 2012). The New Jersey Courts website only keeps the cases for ten business days and then links its users to the Rutgers School of Law site, which serves as a source for appellate cases dating back to 1995.

ber, type of case (criminal, civil, family, agency), and category (for example, accusation, indictment, and grand jury for criminal cases).

At the Hudson Vicinage trial court there are three different divisions: civil, family, and the Children in Court ("CIC") Unit. In the civil division, a user can request records only by name or docket number, not by date of birth or address. The clerks control the process. There are copies of documents like the complaint from a landlord-tenant dispute, a summons, and some affidavits signed by the attorneys. If a user wants to request a record, she must do so through the window where the clerk sits, either after searching the terminal for the docket number, or simply by knowing the party names or docket number when she arrives. The clerk then retrieves the paper record and makes a copy. The user pays for the copy at the "fee station" where the rate is seventy-five cents per page, up to a certain number of pages, with a slightly reduced per-page fee after that.[213]

In the family division, similar to the civil division, a user has to know either a party name or docket number to perform a search. Because of the confidential nature of records available in this division, the clerk determines whether the user may have access to a requested record before it is made available.[214] If the record is confidential with regard to the public but available to particular individuals (for example, the parties), the user must show identification. But if it is a public record (custody dispute, divorce case), no identification is needed. The clerk retrieves a paper copy of the record and makes a copy for the user. This division only contains records from Hudson County, and the cost per-page is the same as in the civil division.

The CIC Unit handles presumptively confidential juvenile records and Division of Youth & Family Services ("DYFS") adoptions. These records are released only to the child or named parent, and possibly to a caseworker if she can show cause. A judge makes the "for

213. These are the rates provided to the authors when they visited the courthouse. However, the Superior Court website provides a different rate schedule of five cents per page for paper copies, with certified copies costing $5 for the first five pages and seventy-five cents per page thereafter. *See Copy and Authentication Fees*, NEW JERSEY COURTS, http://www.judiciary.state.nj.us/superior/copies_court_rec.htm (last visited Mar. 8, 2012).

214. *See, e.g.*, N.J. R. 5:17-4(b) ("Social, medical, psychological, legal, and other records of the court or family intake services, and records of law enforcement agencies, found to be part of a juvenile-family crisis matter, shall be strictly safeguarded from public inspection and shall be made available only pursuant to N.J.S.A. 2A:4A-60 to -62. Any application for such records shall be made by motion to the court.").

cause" determination.[215] However, a release of authorization may be signed by the child or parent to release the records to a third party. Unlike in the family and civil divisions, there is no charge for these records.

b. Observations from Search Experience

In order to get to the search interface, we had to travel to the physical location of the courthouse and pass through courthouse security.[216] We conducted searches assuming that the user knew the party name. This level of background knowledge is reasonable in a scenario where a data aggregator is looking for information about an individual and is aware only of the individual's name. The entire retrieval process took about an hour and a half and cost about $10.

At the Superior Court Clerk's office in Trenton,[217] we performed a search using the party name in ACMS and were able to get one record in the result set. We were only able to retrieve the party names, attorney names, docket number, and judgment amount from the record.

We also used the JEFIS program to access foreclosures. Because JEFIS only allowed searches by docket number, date, or filing or transaction ID, it limited the background knowledge that would be useful for our searches. The JEFIS search ultimately produced documents related to some cases, but these documents only listed the plaintiff's address (and, very rarely, the defendant's as well), the date of foreclosure, and the various amounts of money owed. The documents did not provide any further information.

We also performed a party-name search using the "Appellate Division Search" feature. No documents were available, however, apart from docket entries. Addresses and occasionally phone numbers of the parties were available, though in most cases only the attorneys' addresses were listed.

The process of getting to the interface at the trial court in Hudson Vicinage, similar to the court in Trenton, was time-consuming. Further, as mentioned earlier, it was necessary to interact with clerks at every step. Although we were able to retrieve a record in the family

215. *See* N.J. STAT. ANN. § 2A.82-46 (West 2011) (stating that all information in child assault or abuse cases is "confidential and unavailable to the public" unless judge finds good cause at a good cause hearing and after notification to victim or victim's family).

216. The Hudson Vicinage William Brennan Courthouse is located at 583 Newark Avenue, Jersey City, NJ 07306.

217. The office of the Superior Court Clerk is at 25 W. Market Street, Trenton, NJ 08625.

division, it was only two pages long and appeared to have very little information. We could not access the records in the CIC Unit because we failed to show cause. When we tried to retrieve criminal court records at the same location, we were directed to another building referred to simply as "the Justice Center," about a mile away. After passing a security check at the Justice Center and stating the reason for our visit, we were redirected to yet another building. At this location we were asked to go back to the Justice Center, where we were told that the criminal records were at the courthouse where we had looked at civil, family, and CIC unit records. We could not locate the criminal records room to make a request. Practical obscurity prevailed at this courthouse.

C. Root Causes of Cost Differences in Online and Local Access to Court Records

In this section, we identify six root causes for the difference in cost of flows of personal information between online and local access systems for court records. We draw on our general understanding of searching court records as captured in the eight-part information-retrieval model[218] and observations from the empirical study to arrive at these results.

1. *Getting to the location of the information system.* The cost difference, while unsurprising, was significantly greater than we had anticipated. A user can search online if she has access to a terminal with an Internet connection and knowledge of the URL or name of the online system (PACER, Google Scholar, etc.). In contrast, the cost of a physical search is greater because of the time required to get to a courthouse and the expense of the commute. Moreover, while records of different types are easily identifiable and searchable online, our search for different types of cases at physical courthouses involved hunting for the right room or building, thus adding to the cost.

2. *Query interface and indexing mechanism.* The query interface determines what types of background knowledge can be used in the search process.[219] Online systems, such as Google Scholar, which support keyword search and full-text indexing, significantly reduce the cost of information retrieval for users who may not have the exact knowledge required to formulate queries on systems, such as PACER,

218. *See supra* Part IV.A.

219. *See supra* Part IV.A.4.

which allow only fixed format queries.[220] In the physical courthouses we visited, the query interfaces were much more restrictive;[221] it was difficult to access records for which we did not know exactly what we were looking.

3. *Linking multiple information sources.* Multiple online information sources can be queried during a search. In our experiments with PACER, for example, we used Google Scholar as an auxiliary source to drastically cut down on the search costs. In one test search, where we assumed knowledge of the case title alone, a search by only the party's last name (as obtained from the case title) yielded more than 200 records in PACER. Searching on the case title on Google Scholar also yielded multiple results, but when including the case year given in the case title, we were able to pinpoint the correct record. From this record, we obtained the case number as well as the complete party names. Returning to query PACER with the case number we were able to find the correct court record. These searches also illustrate that restrictions in the query interface of one search system may be circumvented by the use of another search system. In contrast, shuttling between different physical courthouses to execute such linking incurs a significantly greater cost.

4. *Access restrictions.* For certain types of court records, the relation between the user and the data subject determines whether access is permitted. As discussed previously, the CIC Unit releases juvenile records only to the child or named parent, and possibly to a caseworker if she can show cause. It is unclear how to enforce such access restrictions online without an expressive identification and authentication infrastructure—that is, a system that enables users to demonstrate who they are and how they are related to the data subject. It is technically feasible to construct such a system (for example, by using so-called "trust management systems"[222]), although the cost of deploying and maintaining such a system could be quite high. Currently, Google Scholar does not require any form of identification and is incapable of providing access to such records. While PACER does require identification and authentication, it is not clear whether this in-

220. Depending on the type of case, PACER allows users to search "by case number, party name, complete or last four digits of a social security number, case filing dates, and much more." PACER USER MANUAL FOR ECF COURTS, at 18 (June 2010), *available at* http://www.pacer.gov/documents/pacermanual.pdf.

221. *See supra* Part IV.B.2.a–b.

222. *See, e.g.*, Matt Blaze et al., *Decentralized Trust Management*, *in* 1996 IEEE SYMPOSIUM ON SECURITY AND PRIVACY, at 164–73.

formation is used to restrict access to records.[223] In addition, PACER does not allow users to state their relationship to the data subject or to provide evidence to support those claims, in contrast with courthouse access where presenting and verifying such credentials are possible. Thus, at this point in time, certain flows of information that are possible at physical courthouses are not possible online (at least, in the systems we studied).

5. *Format of records.* The format of the records affects the cost of subsequent use and the dissemination of information in the records. For example, searching a native PDF document[224] is significantly less costly than searching a scanned handwritten record. Our online searches often yielded portable documents for which the first copy was available for a nominal fee;[225] once obtained, these documents could be subsequently used and disseminated without incurring any additional cost. In contrast, in some physical courthouses, only paper documents were available.[226] Making additional copies, at five cents per page, increases the cost. Although one could disseminate scanned paper copies without incurring additional cost, the resulting documents could only be searched at significantly higher cost than native PDF documents.[227]

6. *Human factors.* Finally, human factors affect the cost of retrieving information from court records. In particular, the user-interface for searching records affects the cost in at least two ways. First, much work in search engines has gone into figuring out how to present results so that users can quickly identify the results of greatest interest. In Google Scholar, the feature of displaying snippets of text in addition to the title of returned result was particularly helpful. In contrast, searching at the physical courthouse involved using dated sys-

223. In one of the searches we tried, we were denied access to the documents related to the court record, including the complaint form, summons orders, answer to the complaint form, and the Social Security transcript filed. We could not tell from the information provided by PACER why access was denied to this particular court record. Further research on this particular case suggests it may have been related to privacy concerns over medical records or Social Security information or both.

224. A native PDF is one that is digitally converted from an electronic text format into the PDF format. While converted in this manner, the text of the document may be searched for specific terms and specific passages may be copied and pasted out of the PDF.

225. *See supra* Part IV.B.1.a. There is no fee for Google Scholar, but PACER charges a fee for documents retrieved.

226. *See supra* Part IV.B.2.a.

227. Unlike native PDFs, scanned PDF documents are created from the scanned image rather than the original metadata. Thus, the ability to search for specific words or phrases is lost. Some software, which is capable of recognizing individual letters, can be used to maintain the ability to search the document for specific words or phrases.

tems with poorly designed interfaces and interacting directly with clerks. Second, a human interface in the clerk's office provides a level of informal social protection;[228] a user might be hesitant to ask for information that would indicate improper use of court records (for example, a juror accessing a case file), but a user who is searching anonymously online is not similarly inhibited.

V. EVALUATING ACCESS PRACTICES AND POLICIES

In Part IV, we revealed how the flows of information in court records, shaped by explicit policy that has been enacted through the media of hardcopy and local access terminals, may systematically be altered by the adoption of Web-based publication and access.[229] Our analysis makes clear, however, that the precise contours of the alterations in flow are not determined by the choice of medium alone; rather, several specific further choices are crucial to the outcome, such as search interfaces, indexing characteristics, and linkages to other information sources.[230] Since our inquiry is normative—what courts *ought* to do—it must look beyond merely revealing differences to evaluating them. Guided by the framework of contextual integrity, as outlined in Part III, this involves a two-layer evaluation: (a) taking stock of the impacts of respective information flows in general moral and political terms, and (b) establishing the significance of these impacts in terms of values and purposes internal to the justice system and courts, specifically. We follow in Part VI with recommendations drawn from our two-layer analysis.

A. Impacts of Information Flows in General Moral and Political Terms

There are various sources from which to draw insight into the impacts of information flows into and out of court records. For example, the reasoning of judges as they decide whether to seal parts of a record often reveals their assessments of impacts, as do rules and policies explicitly adopted by respective courts. Finally, insights drawn from privacy scholarship, in general, may inform our assessment of information practices followed by courts.

228. *See* Winn, *supra* note 95, at 157 ("Perhaps the most significant change to the judicial information ecosystem was the elimination of a human interface in the clerk's office, ending the informal social protections which formerly existed to control access to the case file.").

229. *See supra* Part IV.B.1.

230. *See supra* Part IV.C.

While judges have cited shame, stigma, the negative judgment of their communities, and chilling effects as reasons for granting anonymity to rape victims, rape is not the only context in which such chilling effects might come into play. A different type of chilling argument is found in *In re Twentieth Century Fox Film Corp.*,[231] where a New York court approved the request of Macaulay Culkin and the film studio to seal a court-approved performance contract so as not to discourage others from entering into contracts with children; the judge also determined that the public's right to know these details was weak.[232] Adverse publicity is not in itself sufficient; for example, in *In re Azabu Buildings Co.*,[233] which involved a debtor and mortgage company, a federal court held that the threat of negative publicity was not a compelling reason to seal a record.[234] In a case involving Bob Dylan, a judge closed portions of the record deemed immaterial to the case both to prevent unfair use of it by the plaintiff and to protect Dylan's privacy and reputation.[235]

The unfair advantage one side may gain over the other has, on other occasions, too, been found sufficient to restrict access to parts of a record.[236] To protect the identity of an informant in an ongoing investigation and assure the safety of her family, government requests to seal records have been honored.[237] And although medical information might be sealed or redacted if it is deemed stigmatizing and immaterial to the case, judges have refused to do the same in cases where it is material, such as personal injury litigation.[238]

231. 190 A.D.2d 483 (N.Y. App. Div. 1993) (per curiam).

232. *Id.* at 486–87.

233. No. 05-50011, 2007 WL 461300 (Bankr. D. Haw. Feb. 7, 2007).

234. *Id.* at *2.

235. Damiano v. Sony Music Entm't Inc., 168 F.R.D. 485, 491–93 (D.N.J. 1996) (finding that the material was unrelated to public safety and that Dylan's personal business affairs are not a legitimate public concern, and suggesting that fairness and efficiency are promoted among litigants as long as they share these materials with each other, and that this is not contingent on public accessibility).

236. *Cf.* In re *Twentieth Century Fox Film Corp.*, 190 A.D.2d at 488 (noting that Twentieth Century Fox's "relationship with its competitors, as well as with other artists in its employ, could be compromised by the disclosure of the details of the contracts, which include information as to how it has marketed the subject motion picture").

237. *See, e.g.*, United States v. McVeigh, 918 F. Supp. 1452, 1458, 1466 (W.D. Okla. 1996) (noting that statutory laws allows records to be sealed when they "'could reasonably be expected to endanger the life or physical safety of any individual'" (quoting 5 U.S.C. § 552(b)(7))).

238. *See, e.g.*, White v. Worthington Indus., Inc. Long Term Disability Income Plan, 266 F.R.D. 178, 196 (S.D. Ohio 2010) (requiring redacted version of record be publicly filed because of its importance to the court's decision and denying motion to file record under seal).

Although courts handle access restrictions in a variety of ways, there is substantial, if not universal, consistency surrounding certain types of information, including trade secrets, other confidential commercial research, national security concerns, and wholly private family matters such as child custody or adoption.[239] Information that may promote scandal, defamation, or unnecessary embarrassment is handled similarly.[240] Information that "poses a serious threat of harassment, exploitation, physical intrusion . . . or the potential for harm to third persons not parties to the litigation" also may be generally protected.[241] Certain categories of people, such as minors,[242] victims of rape and abuse,[243] and celebrities,[244] are also generally given special consideration.

We do not mean to set much store by the particulars of any one of these cases, except to illustrate the *types* of considerations that have been and might be brought to bear in determining limitations on access to information in a court record. To summarize: courts have weighed the commitment to public access to records against considerations of safety, stigma, shame, unfair disadvantage, and reputational damage to concerned parties. At times, these factors are considered on a case-by-case basis and at times systematically through principles and rules developed and adopted by court administrators. It is worth noting that there are several ways to restrict information flows into and out of records: redaction, sealing, and selective disclosure to specific personnel.

While enthusiastic supporters of a transformation from local to online access admit that limited but legitimate reasons for restricting access exist, they deny that the medium or mode of access bears a sys-

239. *See supra* Part II.C; *see also* Doe v. Blue Cross & Blue Shield United, 112 F.3d 869, 872 (7th Cir. 1997) (stating that "[r]ecords or parts of records are sometimes sealed for good reasons, including the protection of state secrets, trade secrets, and informers"); Holland v. Eads, 614 So.2d 1012, 1016 (Ala. 1993) (listing scenarios where most courts presume that the court documents should remain sealed).

240. *See, e.g.*, James v. Jacobson, 6 F.3d 233, 238–39 (4th Cir. 1993) (noting that "privacy or confidentiality concerns are sometimes sufficiently critical that parties or witnesses should be allowed" anonymity, and collecting cases).

241. *Holland*, 614 So.2d at 1016.

242. *See, e.g.*, N.J. R. 1.38-3(d)(17) (protecting records of hearings on welfare or status of child from public access); *Blue Cross & Blue Shield United*, 112 F.3d at 872 ("[F]ictitious names are allowed when necessary to protect the privacy of children, rape victims, and other particularly vulnerable parties or witnesses.").

243. *See, e.g.*, N.J. R. 1.38-3(d)(9)–(10) (protecting records related to domestic violence and sexual offenses).

244. *See generally* John Gibeaut, *Celebrity Justice: The Rich and Famous Get Star Treatment, Creating the Appearance of a Two-Tiered Court System*, 91 A.B.A.J. 43, 45–46 (2005) (discussing how trial courts have protected celebrities' information with increasing frequency).

tematic relationship to these restrictions. Arminda Bepko, for example, in pressing for online access to all records currently available in courthouses, "argues that the constitutional and common law presumption in favor of public access to court documents should not shift depending on the medium."[245] Other advocates of unfettered Internet access offer a more pragmatic rationale. Associate Justice Albin of the Supreme Court of New Jersey, for example, a supporter of greater public access to court records through the Internet, states: "The information genie already has been released from the lamp, and we cannot return to a simpler time when court records, although open to the public, were stored in the practical obscurity of the clerk's office in the county courthouse."[246]

Given the widespread agreement that the choice of medium makes a difference to degree of access, why is there disagreement over policies that should govern publication and access? To answer this question, it is useful to situate court records within the landscape of government records more generally.

Government agencies amass vast quantities of information in countless databases, making them available to citizens and other residents in highly variable arrangements. At one extreme, for records such as geological surveys, FDA drug testing, and macroeconomic indicators, unrestricted public access seems unquestionably warranted. Let us call these "green records." At the other extreme, for records pertaining to "top secret" national security, law enforcement, and individual income tax, no access is given beyond a handful of authorized personnel. Let us call these "red records." In between, lies a huge range of what we will call "orange records," which are publicly accessible with varying restrictions that may foreclose scrutiny of all or parts of the records. In the case of the Census, for example, while access to raw data is highly restricted,[247] public access is freely given to anonymized aggregations.[248] Restrictions might be imposed through

245. Arminda Bradford Bepko, Note, *Public Availability or Practical Obscurity: The Debate Over Public Access to Court Records on the Internet*, 49 N.Y.L. SCH. L. REV. 967, 968 (2005); *see also* David Robinson et al., *Government Data and the Invisible Hand*, 11 YALE J.L. & TECH. 160, 166–67 (2009) (suggesting government open access to its data for public use).

246. REPORT ON PUBLIC ACCESS, *supra* note 15, at ii.

247. *See* 13 U.S.C. § 9(a) (2000) (restricting use, publication, and access of individual data collected for the Census); Douglas Kysar, *Kids & Cul-de-Sacs:* Census 2000 *and the Reproduction of Consumer Culture*, 87 CORNELL L. REV. 853, 870 (2002) (reviewing U.S. CENSUS BUREAU, U.S. DEP'T OF COMMERCE, CENSUS 2000: CENSUS OF POPULATION AND HOUSING (2001)) ("[T]he Bureau has maintained a policy of strict confidentiality with respect to individual census returns since at least 1840.").

248. *See* 13 U.S.C. § 8(b) (2000) (stating that "tabulations and other statistical materials" may be obtained); U.S. CENSUS BUREAU, DATA PROTECTION, http://www.census.gov/pri-

conditions on access, such as credentials or authorization, or by imposing limits on use. Court and DMV records (particularly following the Drivers Privacy Protection Act) are also examples of orange records.[249]

When considering how to adapt existing policies to the new media of digital file storage and network access, the key challenges might seem to be economic and engineering. For green records, for example, this seems straightforwardly to be the case, namely, how to design a new system that provides the most effective, efficient, and useful access. As media for information delivery have progressed, and, with the advent of digital electronic information technologies, improved by orders of magnitude, it is reasonable to expect that socio-technical systems for providing access will improve accordingly. The only countervailing consideration might be expense, as governing authorities juggle diverse claims on limited resources. Because less-than-optimal access increases the generalized costs (as defined in Part IV) of access for citizens, resisting improvements might easily be read as opposition to government openness.

Orange records—those governed by varying ensembles of rules—present a greater challenge. The records administrator must engineer access constraints not only to embody the ensemble of rules accurately but—here we state the obvious—to adapt them appropriately to characteristics of the medium; thus, the proverbial thick black marker may be effective for selectively redacting print on paper; locked filing cabinets effective for providing selective access to those with keys; and an obliging, non-intrusive clerk effective for approximating open access to information held in manila folders stored in filing cabinets. When the recording medium changes, and the task of implementing policies are revisited, indeterminacies may be encountered when a new medium allows new options or shuts down existing ones. To give a simple example, redaction in electronic media could be modeled in a variety of ways.[250] Although permanent erasure of words within a document might most closely mimic the action of an indelible black marker, an administrator, confronted with a new technology allowing for reversible obfuscation, might find explicit policy silent on the question of permanence or reversibility. New op-

vacy/data_protection/ (last revised Feb. 16, 2012) (explaining the obligation of Census Bureau to provide statistics for public and private use).

249. Drivers Privacy Protection Act, 18 U.S.C. § 2721 (West Supp. 2011) (limiting ability of state department of motor vehicles to disclose information collected in connection to motor vehicle exception, subject to some exceptions).

250. *See* Lee, *supra* note 42.

tions can be liberating, but they also reveal incompleteness in explicit rules in relation to actions that simply were not possible before; what might seem at first to be merely an engineering question turns out, after all, to be a policy question.[251] Although some of the new options opened by reengineering are neutral with respect to policy, others might render policy ambiguous or, worse, may actually undermine unarticulated policy expectations. It is important for policymakers and administrators at least to recognize these possibilities and be ready to make determinations, where necessary.

If the purpose of Part IV was to expose how a change in technical systems may result in altered flows and novel access options, this Part's purpose is to expose where these options reveal policy ambiguities and even undermine unexpressed policy expectations, particularly those relating to privacy interests. Because migrating to the new medium results in a lifting of constraints that had "naturally" been embodied in prior media, it exposes a need to locate resulting flows that contravene policy or expectation. Some of the trouble surrounding Social Security numbers illustrates the need to be alert to inadvertent gaps opened by technological choices. Against the backdrop of widely available personal information, Social Security numbers emerged as valuable keys to identity theft.[252] To counteract the lifting of natural barriers inherent in prior dominant media, as well as novel threats posed by networked data sources, lawmakers and regulators have acknowledged the need to limit the availability of Social Security numbers, both retroactively and proactively, in court and other public records.[253]

In light of transformations in flow due to the lowered costs of access,[254] let us start by considering the criteria that courts—including judges and administrators—have consistently recognized as legitimate for restricting access to records or parts of records, whether in regulations or in case-by-case assessments. Whatever the case was favoring restricting access (for example, by sealing) to specific types of information, such as financial, medical, and adoption, it is surely strengthened by the prospect of radically lowered costs of access and ought to add significant weight in case-by-case balancing of potential harms

251. *See, e.g.*, LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE 111–18 (1999) (discussing the development of wiretapping technology and the Court's approach to its use under the Fourth Amendment).

252. *See* Bepko, *supra* note 245, at 998 (suggesting that Social Security numbers stolen from court documents can be used to commit fraud).

253. *See, e.g.*, N.J. R. 1:38-7(a)–(b) (prohibiting the disclosure Social Security numbers and other personal identifiers from any court document or pleading).

254. *See supra* Part IV.

of access against benefits. An argument may also be made, however, for imposing restrictions beyond those criteria readily acknowledged by judges and court administrators.[255] Powerful new capacities to infer financial, family, or medical information, for example, from combinations of other more innocuous information types suggest a need to expand the class of potentially stigmatizing information that can justify sealing.[256] Under the "Mosaic Theory," an equivalent argument is offered to justify limits on access to a wider swathe of information in government hands on grounds of potential threats to national security.[257] Even information not previously thought to be risky, by itself, can become risky when linked to additional publicly available information.[258]

Out of concern for the intimately personal and sometimes even embarrassing information revealed in many cases, courts have been prepared to seal or redact information in records that could cause scandal, defamation, harassment, ridicule, or unnecessary attention and embarrassment, giving special consideration to minors, victims of rape and domestic abuse, third persons not party to the litigation, and, in some instances, celebrities. The point highlighted here is that given the magnitude of impact when information is posted online, the chances of scandal, defamation, embarrassment, and unnecessary attention are increased significantly, particularly when nosy neighbors, celebrity followers, or anyone with an interest in a particular individual is able to search online and, at low cost, associate that individual with information from a court record. An extreme example of the impact of linking publicly available information is a case in France, where Eric Schmidt, the former CEO of Google, was found guilty of defamation because Google Suggest brought up terms such as "Satanist," "Rape," "prison," and "rapist" when the name of a man identified in court records only as Mr. X was entered into a search box.[259] Mr. X had been convicted of "corruption of a minor" in a

255. For example, the Judicial Conference Committee on the Court Administration and Case Management adopted a set of guidelines that require sensitive information to be redacted in civil and bankruptcy cases. Bepko, *supra* note 245, at 977. Remote electronic access is also denied to any documents in criminal cases. *Id.*

256. David E. Pozen, Note, *The Mosaic Theory, National Security, and the Freedom of Information Act*, 115 YALE L.J. 628, 630 (2005).

257. *Id.* at 645–47 (noting that the mosaic theory has lead to heightened protection of national security information by the Executive Branch, including several new limits to the Freedom of Information Act).

258. *Id.* at 630 ("In the context of national security, the mosaic theory suggests the potential for an adversary to deduce from independently innocuous facts a strategic vulnerability, exploitable for malevolent ends.").

259. Weintraub, *supra* note 197.

French court.[260] His conviction triggered the terms returned by Google Suggest, but the associated terms were clearly far more damning than the charges leveled against the man, who upon appeal was given merely a three-year suspended sentence and a fine.[261] A more typical case, however, might involve web searches that potentially reveal involvement in legal action.[262]

It may be difficult to sympathize with criminals or those who have deservedly been sued wishing to avoid scandal and embarrassment. There are, however, many non-criminals identified in court records who are subject to the same treatment, including persons found innocent after having been arrested and charged, those who have sued with justification, those implicated as third-parties, and witnesses, to name but a few. It is worth weighing the added burden to these individuals, at least those who may suffer embarrassment, ridicule, prejudicial discrimination in seeking employment and housing, and even simply uninvited attention as a result of linkable information.

The potential for harassment, however, deserves special mention. There have been chilling stories of people suffering grossly disproportionate harassment for misdemeanors that happen to have captured the attention of the online masses, such as the infamous "dog poop girl" in Korea.[263] In China, the Internet has also facilitated so-called "human flesh search engines," a form of vigilantism aimed at punishing accused wrongdoers through coordinated, massive online and offline reaction.[264] While a single reproach, deserved or not, can be unpleasant, the same reproach coming from thousands is surely harassment that no civilized society should encourage. The capacity of the Internet to unleash collective retribution in this manner deserves to be among the factors considered when balanced against the posting online of personal information in a court record. This, along with other considerations suggests that with higher stakes, the bar for restricting access should be lowered.

260. *Id.*

261. *Id.*

262. *See, e.g.*, David Kravets, *Mug Shot Industry Will Dig Up Your Past, Charge You to Bury It Again*, WIRED.COM (Aug. 2, 2011, 1:52 PM), http://www.wired.com/threatlevel/2011/08/mugshots/.

263. *See* Jonathan Krim, *Subway Fracas Escalates Into Test of the Internet's Power to Shame*, WASH. POST, July 7, 2005, at D1 (describing the impact of blogs in forcing a South Korean student to quit college after she refused to clean up after her dog on a subway, pictures surfaced of her online, and people discovered her identity and vilified her in public).

264. Ariana Eunjung Cha & Jill Drew, *New Freedom, and Peril, in Online Criticism of China*, WASH. POST, Apr. 17, 2008, at A1.

B. Values and Purposes Internal to Courts and the Justice System

Part V.A addressed the first layer of evaluation prescribed by the framework of contextual integrity, taking stock of impacts of new patterns of information flow with ethical import. This section addresses the second layer of evaluation by considering the significance of these impacts for the advancement of context-specific values, ends, and purposes, particularly as compared with prior practice. In the cases of the U.S. Census Bureau or Internal Revenue Service, for example, both known for highly restrictive access policies for raw data, the impacts of more relaxed privacy policies might include harms to individuals, an unbalancing of power relations between citizens and government, and so forth. Yet, in both cases, an argument that has held sway highlights deleterious consequences for the effective functioning of respective agencies. Because individuals will be less likely to participate and less likely to do so honestly without appropriate restrictions on flows, the agencies' goals of universal and accurate reach would be undermined.[265]

So, too, must we investigate how changes to constraints on access to courts records, including constraints motivated by privacy concerns, affect not only individual data subjects but also the degree of success courts may have in advancing context-specific goals and values. What precisely are the goals and values of courts within the justice systems of liberal democracies is likely to be debated and discussed, even in well-ordered, free societies, and this Article is no place for a comprehensive account of these. Instead, we draw from a set of goals and values that are widely cited, solidly robust, and relatively uncontroversial, which can be found in scholarly articles and formal and informal accounts of individual commentators and professional organizations. These serve as a basis for our argument, which proceeds as follows: brief discussions of (1) ends and purposes (or goals); (2) values served by courts; (3) how court records function to promote goals and values; and (4) how deployment of networked access that disrupts information flows may undermine the attainment of these goals and values.

1. Ends and Purposes

Most would agree that courts serve to establish that a crime has occurred and pass judgment on criminal guilt. They adjudicate, or resolve conflicts and disputes among public and private parties in ac-

265. IRS, U.S DEP'T OF THE TREASURY, DISCLOSURE & PRIVACY LAW REFERENCE GUIDE 1-1 to 1-9 (2007).

cordance with prevailing law. They officiate key societal relationships, such as marriage and divorce, and establish child custody. Courts are empowered not only to assess crime and adjudicate disputes but also to attribute blame, assess liability, and determine punishment, remedies, compensation, fines, and injunctive relief. According to the National Association of Court Management:

> Courts reinforce the authority of the state and the legitimate use of force and protect individuals against the arbitrary use of governmental power.
>
> . . .
>
> . . . Courts exist to do justice, to guarantee liberty, to enhance social order, to resolve disputes, to maintain rule of law, to provide for equal protection, and to ensure due process of law.[266]

The association further states:

> Only the judiciary can definitively determine who is to prevail in the inevitable conflicts that arise between individuals; between government and the governed, including those accused by the state of violating the law; between individuals and corporations; and between organizations, both public and private. . . . They resolve disputes by applying the law to the facts of particular cases independently and impartially.[267]

Similarly, in Sharon Rodrick's words, "Judicial power is concerned with the determination of disputes and the making of orders concerning the existing rights, duties and liabilities of persons involved in proceedings before the courts."[268] Courts are not merely applying law but, more often than not, they are creating law in their interpretations and judgments and in setting precedents for future courts. Within these broad categories, there are, of course, myriad activities and practices that support them.

2. Values

As one of the pillars of the tripartite system of democratic governance, courts absorb core values of the overarching political system while paying special heed to values that are tied to their specific

266. NATIONAL ASS'N FOR COURT MGMT., CORE COMPETENCY CURRICULUM GUIDELINES: PURPOSES AND RESPONSIBILITIES OF COURTS, 1 (2003) [hereinafter PURPOSES], *available at* http://www.nacmnet.org/sites/default/files/images/1purposes_0.pdf.

267. *Id.*

268. Sharon Rodrick, *Open Justice and Suppressing Evidence of Police Methods: The Positions in Canada and Australia (Part One)*, 31 MELB. U. L. REV. 171, 185 (2007).

mandate.[269] In listing the courts' purposes as adjudicating disputes, establishing guilt, meting out punishment, and so forth, it is essential to pair these with the values to which courts subscribe: values of justice, fairness, proportionality, impartiality, predictability, freedom from bias ("equal protection"), independence from inappropriate influence, and due process.[270] As a wing of liberal democracy, supported by and serving citizens, courts strive for effectiveness, efficiency, competence, trustworthiness, accountability, service excellence, openness, and transparency.[271] Courts must not only administer justice but also do so with the appearance of justice. Rodrick writes:

> Much has been written about the purpose and value of open justice. Of prime importance is the belief that open justice enhances the integrity, accountability and performance of those who are involved in the administration of justice. For example, it is supposed that openness makes judges more accountable for the manner in which they exercise the judicial power that is vested in them; secret courts are regarded as having a propensity to spawn corruption. Witnesses are thought to be more likely to tell the truth if they have to testify orally in open court. For their part, the public, having observed responsible and truthful behaviour on the part of the judges and witnesses, will have increased confidence in the operation of the courts and a greater understanding of society's laws and legal system. Any perceived shortcomings in the behaviour of particular individuals or in the substance of the law or its application in a particular case can be publicly scrutinised and, perhaps, corrected.[272]

One final remark on ends, purposes, and values: courts might do all of the above and yet still fall short of their responsibilities to society. It is not sufficient for courts, *if* employed by members of society, to mete out justice and perform their tasks fairly, transparently, and so forth, if they are under-utilized. Although there are means outside of the courts to deal with bad actors and alternative institutions for re-

269. *See* John F. Manning, *Separation of Powers as Ordinary Interpretation*, 124 HARV. L. REV. 1939, 1942–43 (2011) (recognizing two approaches to judicial separation of powers: (1) a functionalist approach that eschews the separation of powers notion; and (2) a formalist approach in which courts are bound by the U.S. Constitution).

270. *See* PURPOSES, *supra* note 266, at 1–2.

271. Natalie Gomez-Velez, *Internet Access to Court Records—Balancing Public Access and Privacy*, 51 LOY. L. REV. 365, 369 (2005).

272. Rodrick, *supra* note 268, at 172 (footnote omitted).

solving disputes, courts carry the authority of the law, must act in accordance with the law, and often create law when they act.

3. Function of Court Records

Internal procedures: Records are inherent to the internal functioning of courts as a *court of record*.[273] Decisions made in court must be noted so they can be properly acted upon: perpetrators fined or imprisoned, debts paid, harms compensated, and so forth. Records not only document these decisions, they convey them to other collaborators within the justice system, such as prisons and law enforcement.[274] Court records constitute a repository of information necessary first for the processing of cases at the trial level, and later as vehicles for transmitting key facts, as cases are appealed. Good records systems can contribute to efficiency, integrity, and fidelity of court function.

Educate and inform: Records contribute to the quality of courts' functioning by educating and informing those who work in or with, or are training to work in or with, the legal system, including lawyers, judges, legislators, court administrators, and academics.[275] Records help make these crucial participants better at their jobs. Although most legal professionals pay for access to records from the dominant third party services that organize records and include citations, such as Westlaw and Lexis, full records remain freely available. By communicating process, reasoning, and precedent, court records promote competence, knowledge, consistency, and understanding.

For society at large, court records provide a window into how courts function.[276] Though most likely interpreted by the media and not read directly, court records facilitate for the public a better understanding of the judicial branch of government, specifically, the court system, how and when to use it, and what to expect if one does.[277] Court records serve the end of familiarizing citizens with the

273. Interview with Jim Rebo, Chief Security Officer, New Jersey Courts, in Trenton, N.J. (May 15, 2009).

274. Margo Schlanger & Denise Lieberman, *Using Court Records for Research, Teaching, and Policymaking: The Civil Rights Litigation Clearinghouse*, 75 UMKC L. REV. 155, 160 (2006).

275. *Id.* at 162.

276. *See* Gomez-Velez, *supra* note 271, at 402–03 ("[A]n important counterpoint to concerns about individual privacy in the context of placing court records on the Internet is the competing and legally recognized interest in government transparency—in providing public access to information about government functions.").

277. *See id.* at 406 n.136 ("'Electronic access ensures that the public and the news media can oversee how justice is administered like never before.'" (quoting REPORTERS COMM. FOR FREEDOM OF THE PRESS, ELECTRONIC ACCESS TO COURT RECORDS: ENSURING ACCESS IN THE PUBLIC INTEREST)).

workings of government; an educated and informed citizenry contributes to better government in all branches.[278]

Public oversight, open justice, and transparency: As windows to the courts, court records serve not only the educational functions noted above, which contribute to effective functioning, but also serve as internal and external checks. Whether via direct access or via interpretations by NGOs and the media, court records hold judges and other court officials accountable for the quality of their work. Calls for government transparency and accountability are, to a large extent, answered by the free availability of records, the right of any member of society, for whatever reason, to inspect records so they may judge for themselves whether the courts are functioning as they should, serving the public effectively (and cost effectively) and impartially. Transparency and openness of courts to the citizens of a democracy are as vital to this branch of government as to others, whether to educate, satisfy public curiosity, ensure oversight and accountability, protect against corruption and abuses of government power, or provide grounds for the appointment or reappointment, or election or reelection, of judges and other court officials.

4. Disruptive Information Flows

In scholarly discussions of online access, deliberations commissioned by courts themselves,[279] and advocacy rhetoric of public interest organizations, the rationale most consistently offered in favor of boosted access is transparency of court functioning to the public.[280] Our analysis, in terms of the function courts records serve, leads to a similar conclusion, for most of the other functions can be served just as well by restricting access to authorized personnel only—such as court and prison officials, judges, authorized social workers, lawyers, and bail bondsmen—or by heavily limiting what parts of records are disseminated. When judges and court administrators weigh the various and possibly conflicting interests for and against placing documents or information into the public court record, they cite most frequently this right of the governed to inspect, take stock of, and know how they are being governed and how the institutions of their government are functioning.

278. *Id.* at 402–03.

279. *See, e.g.*, *id.* at 368 n.3 (noting that New York commissioned a study on making its government documents more accessible to the public).

280. *See, e.g.*, Schwartz, *supra* note 14 (chronicling an entrepreneur attempting to start a database with free access to court documents under the auspice of openness and transparency).

Part V.A argued that, when balancing and trading off interests against one another in order to formulate explicit rules for so-called "orange" records, properties of underlying distribution media cannot be ignored because they may well make material differences to these interests. Following this logic, we revealed ways that online access can increase the hardship of participants in court cases. Prima facie conclusions drawn on grounds of this observation need to be considered in light of the second layer of analysis prescribed by the theory of contextual integrity.[281]

The array of affected interests, although an essential part, is not the whole story. How significant the disruptions are to the furtherance of context-specific ends, purposes, and values is the rest of it. Rodrick voices a similar concern in the Australian context:

> [C]ourts have taken the view that the principle of open justice is so fundamental that it can be curtailed only when necessary in the interests of the administration of justice in the particular proceeding. . . . [T]he proceeding could not effectively continue (or would be prejudiced or frustrated in some real and tangible way), or the decision of the court would be deprived of practical utility.[282]

Further, "[t]his is because open justice, which is primarily valued for its contribution to the administration of justice, must yield to the need to secure the administration of justice in the unusual event that publicity would be to its detriment."[283]

In other words, when evaluating the consequences of posting records online, although threats of harm posed to information subjects are important, our analysis highlights another set of key consequences, in particular, "the administration of justice" itself, that is, goals and values of courts within the judicial system. If, say, victims of certain types of crimes or wrongs do not bring them to the courts or choose alternative venues to settle them, in order to avoid the additional burdens of publicity, there may be cause for concern.[284] Perpetrators may go unpunished and suffering uncompensated as the

281. *See* Adam Barth, Anupam Datta, John C. Mitchell & Helen Nissenbaum, *Privacy and Contextual Integrity: Framework and Applications*, *in* 2006 IEEE SYMPOSIUM ON SECURITY AND PRIVACY 184, 184 ("Contextual integrity is a conceptual framework for understanding privacy expectations and their implications developed in the literature on law, public policy, and political philosophy.").

282. Rodrick, *supra* note 268, at 185–86 (footnote omitted).

283. *Id.* at 187.

284. *See, e.g.*, Strahilevitz, *supra* note 53, at 1247–48, 1252 (noting that people use anonymous forums online as cheaper alternatives to litigation).

court's role is diminished in areas of social life that might benefit from the justice system's intervention. Although there may be good reasons to choose, say, mediation or arbitration, society may lose out on the development of important legal precedent and remain ignorant of worrying social trends.[285] Further, the courts are important constituents of democratic governance and are accountable to the public in ways that other systems of mediation and arbitration may not be; it would be cause for concern if their range of influence were to tilt or erode because parties were choosing these other systems, possibly less exposed to public oversight.

Equally as important is the role of non-parties, who make essential contributions to court proceedings by acting as jurors, as witnesses, or in other capacities. Grayson Barber and Peter Winn have highlighted the plight of such "unrepresented" parties, including third parties mentioned in court proceedings whose privacy interests might simply be overlooked.[286] These oversights might be irrelevant where court records are accessed via traditional paper methods. In a world of open access to full and complete records, however, certain indexing choices made by court administrators would allow a Google search on any of the parties' names to reveal their involvement in a case.[287] Even a slight disincentive such as this might tip the scale against offering one's services to a court.

Two final points linking new patterns of disclosure with purposes and values: the first concerns justice and proportionality. To the extent that the placement of records online gives rise to further hardships, such as harassment, and discrimination in job and other opportunities, it raises questions about disproportionate punishment. One may have little sympathy for serious criminals, particularly for those who have sexually assaulted children, considering it their just deserts and pointing out the high rates of recidivism. But these are most likely a small percentage of those affected by, say, harassment.

285. *See id.* at 1258 (highlighting the potential loss of important precedent if a litigant had not been allowed to pursue trial under a pseudonym).

286. *See, e.g.*, Grayson Barber, *Personal Information in Government Records: Protecting the Public Interest in Privacy*, 25 ST. LOUIS U. PUB. L. REV. 63, 63–67 (2006) (arguing the need for courts to recognize their "special obligation to protect the public's interest in individual privacy" with respect to government records containing private information); Winn, *supra* note 95, at 152 ("Because of the nature of the adversarial system, the interest of the public in the transparency of judicial records and the interest of unrepresented third parties in protecting sensitive private information are equally ignored as the parties before the court pursue their own personal interests.").

287. *E.g.*, Marder, *supra* note 2, at 450 (describing how Proposition 8 opponents in California used "Eightmaps.com" and Google Maps to locate the homes and email addresses of Proposition 8 supporters and then harass and threaten them).

The fate of expunction is a second related point. Many states offer those found guilty of certain categories of offenses, duly punished and clear of any further offenses, the opportunity to have their records expunged.[288] The same is offered to persons who have been arrested but against whom no further action was taken. Online posting could nullify, or seriously limit, the chance of disassociating one's name from such convictions or arrests.[289]

VI. RECOMMENDATIONS FOR ACCESS

In defining the aims of our research, we have sought to develop a general line of reasoning for considering privacy while focusing on the transition from locally accessed court records to networked electronic records.[290] Our arguments and findings, thus far, point to various approaches that may be taken to address key concerns.[291] Given the vastness of the enterprise and variation across courts and states, the business of embodying a particular, principled approach in specific rules for specific courts will require substantial adaptation and interpretation. Thus, for now, we limit our concluding discussion with brief descriptions of alternatives we favor, as well as respective implications for the design of electronic court-record systems and management of associated information-handling practices. Although we are able to present only open-ended accounts, we aim to provide enough substance to stimulate meaningful debate.

• *Option 1: A radical departure from the status quo*

One global policy option that follows from our discussion is for courts to "sanitize" records by redacting proper names and possibly other immediately identifying information, such as Social Security numbers, bank account numbers, or coded biometric identifiers of all parties and other participants (for example, witnesses, jurors, etc.), before releasing them to the public via the Web or local portals (for example, local terminals or paper files). Given the possibility of on-

288. *See, e.g.*, N.J. STAT. ANN. § 2C:52-2 (West 2005 & Supp. 2011) (providing that "a person [who] has been convicted of a crime under the laws of this State and who has not been convicted of any prior or subsequent crime . . . and has not been adjudged a disorderly person or petty disorderly person on more than two occasions may, after the expiration of a period of 10 years from the date of his conviction, payment of fine, satisfactory completion of probation or parole, or release from incarceration . . . present a duly verified petition . . . to the Superior Court in the county in which the conviction was entered praying that such conviction and all records and information pertaining thereto be expunged.").

289. *See, e.g.*, PURPOSES, *supra* note 266.

290. *See supra* Part III (defending a theory of information privacy).

291. *See supra* Part V (describing policy based on normative information flow analysis).

line posting and the potential to increase the burden on parties as well as unrepresented participants, instead of relying on judges and administrators to make determinations on a field-by-field, case-by-case basis, this alternative would constitute a bright line policy across the board. Its purpose would be to protect threats of harms to interested parties as well as threats to the attainment of ends, purposes,[292] and values[293] served by courts posed by the far-reaching disruptions of information flows,[294] discussed above.

Although, in the United States, it would constitute a radical departure from the status quo to provide public access only to sanitized records,[295] experience with Jane Roe and John Doe cases and the practice in many European countries of maintaining strict anonymity for the parties in court cases suggests it can be done.[296] Our main criteria for evaluating this option, and the ones that follow, are drawn from the approach we have developed, generally, for addressing worrying implications for privacy of online placement.[297] This evaluation considers advantages as well as costs and drawbacks.

Redacting immediately identifying information somewhat lowers the stakes of providing free and unrestricted online access and diminishes pressure to seal or redact other information typically deemed sensitive, such as medical and financial information. It protects against Natalie Gomez-Velez's worry that changes in dissemination patterns due to online access of personal information in court records will not necessarily serve the intended values, increasing such improper uses "as identity theft, stalking, discrimination, locating domestic violence victims, and interference with business and social relationships, rather than for the appropriate oversight, educative, and accountability reasons for which court records are made public."[298] Caren Myers Morrison, in an article focusing on Internet access to

292. *See supra* Part V.B.1 (discussing the ends and purposes of the justice system).

293. *See supra* Part V.B.2 (discussing the values of the justice system).

294. *See supra* Part V.B.3 (discussing potential harm caused by information flows).

295. *See* Marder, *supra* note 2, at 453 (noting that currently neither courts nor lawyers accept responsibility for redacting sensitive information).

296. *See, e.g.*, Elena Larrauri, *Conviction Records in Spain: Obstacles to Reintegration of Offenders*, 3 EUR. J. PROBATION 50, 51–52 n.11 (2011) (noting the Spanish practice of rendering anonymous decisions, i.e., not publishing names, except for Constitutional Court decisions); *see also* James B. Jacobs & Elena Laurrauri, *Are Criminal Convictions a Public Matter? The USA and Spain*, 14 PUNISHMENT & SOC'Y 1, 3–28 (Jan. 2012).

297. *See supra* Part V.A (arguing for a balancing approach to resolve competing interests with respect to governing information flow).

298. Gomez-Velez, *supra* note 271, at 371.

criminal justice records, questions the need for proper names in records with a similar challenge:

> What is the information of value that the public needs to know? Does the public need to know that an individual indicted for distributing five kilograms of cocaine, which would ordinarily entail a mandatory minimum sentence of ten years, cooperated with the government and received a sentence of thirty-six months, or does it need to know that Billy Costigan, in particular, cooperated with the government?[299]

Following this line of reasoning, sanitized records could offer a good compromise between the respective currents of privacy and open government, on the one hand, by limiting improper and harmful uses of personal information in court records while,[300] on the other hand, continuing to support key functions of court records in the attainment of core values and goals.[301] Quality of court and professional performance, education, public oversight, transparency, and accountability[302] will not be diminished because the names of lawyers and court and other officials of the justice system (such as judges) would not be redacted under this approach. Meanwhile, the content of actual value to the public—decisions, reasoning, and other materials necessary to evaluate the quality of a judgment[303]—will be just as available as before, indeed, more available than before, because they would be posted online. In other words, sanitizing records before posting them online promises protections against threats to privacy without significantly compromising values and goals of the courts.

While this option embodies a clear and well-justified principle of action, it remains vulnerable to criticism from at least two sources, one technical and the other normative. On the technical front, the ongoing quagmire surrounding anonymization—its limits and whether it is even possible—surely must raise doubts over the possibility that

299. Caren Myers Morrison, *Privacy, Accountability and the Cooperating Defendant: Towards a New Role for Internet Access to Court Records*, 62 VAND. L. REV. 921, 971 (2009) (footnote omitted). Of course, a litigant or defendant's identity may in fact matter, for example if a particular judge is alleged to give harsher sentences to defendants of a particular race, or if the same plaintiff repeatedly brings harassing lawsuits.

300. *See supra* Part V.B.3 (discussing the moral and political impacts of information flows).

301. *See supra* Part V.B.3 (discussing the function of court records in relation to the key values of the justice system).

302. *See supra* Part V.B.3 (discussing how court records serve these key functions).

303. *See supra* Part V.B.3 (explaining the importance to the public of this information contained in court records).

removing proper names and other obvious identifiers would effectively shield the identities of those involved in a case, either parties or others, particularly if the seekers of information are armed with additional background knowledge.[304] We wonder, however, whether such doubts are sufficient to foreclose this option. Although it may be possible for a highly motivated searcher to re-identify key individuals in a case, redaction of proper names will still protect against serendipitous discovery of someone's involvement in a case as a result of an Internet search on her name. Accordingly, the technical challenges posed by anonymization, generally, must be evaluated in light of the specific challenges of sanitizing court records. Imperfect as it is, simple redaction may constitute the most practically feasible option not, perhaps, compared with perfect anonymization but compared with doing nothing at all. Going forward, therefore, we would like to see further discussion of this point considering not only, for example, what types of information might need to be redacted from a record, or what fields should and should not be indexed, in order to ensure that identifying information be shielded, but also questions economic in nature, such as, how much effort—or cost, as we have called it above—searchers or search services would likely be willing to expend to re-identify individual parties in a particular record of interest or to do the same for records in bulk.

No matter how reasonable this option might be on principled grounds, the break from tradition, the thoroughgoing changes in the way cases proceed and the ways they are recorded and even named make it an unlikely candidate in practice, not least because it may infringe on media and citizens' First Amendment rights by imposing a form of prior restraints on access. Critics might argue that by redacting names, society will lose an important dimension of answerability from the courts, for example, assurances that there is no discrimination for or against plaintiffs or defendants based on race, ethnicity, economic standing, or other inappropriate dimensions. This point is important to consider, both to weigh its value against the benefits of redaction and, in practical terms, to consider how important searchable names are to address it.

304. Linking attacks on relational databases have demonstrated this inadequacy. *See, e.g.*, Arvind Narayanan & Vitaly Shmatikov, *Robust De-anonymization of Large Sparse Datasets*, *in* 2008 IEEE SYMPOSIUM ON SECURITY AND PRIVACY 111, 111 (noting how, even in databases where names and Social Security numbers are redacted, individuals can nevertheless "use background knowledge and cross-correlation with other databases to re-identify" the anonymized individuals); Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701, 1703–04 (2010) (noting how "adversaries can often *reidentify* or *deanonymize* the people hidden in an anonymized database").

Furthermore, the change in this aspect of court administration would be costly, though already state courts have expended millions of dollars in the transformation to electronic systems.[305] Besides the general resistance one would expect from those for whom a change of this scope would be simply too risky to consider, there are many parties whose interests in fully identified open records would be placed at risk, from reporters to information brokers.

Variations: A variant on this approach would sanitize the record of identifying information of all those named in a case except the parties themselves. This alternative would address some of the normative concerns[306] and would diminish the burden on technical requirements, as non-party participants in a case, despite important contributions to the functioning of courts and the justice system, would almost always be of far lesser interest to third parties.

• *Option 2: A two-tiered alternative*

A second option would retain the status quo for local access but produce a sanitized version for the open, indexable Web. In other words, two versions of court records would be produced, one available to the public at local courthouses that includes all information currently approved and available for access, the other, posted online, scrubbed of identifying information in the manner discussed above. This would eliminate concerns about prior restraints, since public access itself would not be curtailed—only online access would be affected by this proposal. Because the First Amendment does not require courts to post their records online but only to make them accessible at the courthouse for inspection and copying, courts are free to post limited or sanitized versions of court records online as long as the full paper record is available at the courthouse.

While this option would allay major constitutional concerns, it would nevertheless continue to support the existing practice of commercial data aggregators paying employees to camp out at courthouses, manually copying records into their data repositories. Furthermore, it does not sufficiently address concerns expressed over inappropriate flows of personal information from court records, and does not stop the circulation of problematic information but rather

305. *See, e.g.*, Megan Poinski, *Electronic Court System Will Make Judiciary Paperless, Judge Says*, MARYLANDREPORTER.COM (Jan. 14, 2011), http://marylandreporter.com/2011/01/14/electronic-court-system-will-make-judiciary-paperless-judge-says/ (discussing how Maryland's transformation to electronic filing will cost between $50 million and $60 million).

306. *See supra* Part V.A.

results in uneven access to those who can afford to pay commercial data aggregators for information.

A two-tiered approach that results in uneven access may, however, yield footholds for policy by holding identifiable third parties, such as data aggregators, to higher standards of accountability. Such policies might impose constraints on how the information flows and uses to which it is put. Ideally, it would also hold third-party aggregators to higher quality standards for the data itself.[307]

• *Option 3: Fine-grained differential access*

The broad policy options we have outlined thus far approximate solutions that would take into consideration finer detail, such as the roles of parties requesting access, the roles of information subjects, the types of information, and the conditions under which information is granted. This ideal is less complex than it may seem, as it generalizes what is already common practice, for example, sealing information about minors from public scrutiny that, presumably, is provided to social workers, or redacting certain types of information from the record that, ultimately, will be provided to the public.[308] Our argument has been that a reconsideration of common practice is necessary in light of disruptions to information flows due to online access (discussed above).[309] But these same developments in information science and media technologies also offer greatly enhanced capacities for managing the additional complexity of finer-grain rules, as long as the systems being built for use in courts take advantage of them.[310]

In creating digital records, system designers have at their disposal the means of building structure into them, including the ability to tag data fields. This structure would allow constraints on selective access to various parties based on rules governing differentiated access privileges to differently tagged fields of information. Thus, for example, a policy with a default to redact from a publicly accessible record the names of all non-parties, including members of juries, witnesses, and those inadvertently implicated through the case, which may have been practically impossible in the past, becomes quite routine with

307. Interview with Grayson Barber, Grayson Barber LLC (April 2009).

308. *See supra* Part II.C (explaining the mechanisms and processes through which courts decide whether or not to redact or seal information in their records).

309. *See supra* Part V.B.4 (noting the problems with online information flows and proposing various prescriptions).

310. *See* Gomez-Velez, *supra* note 271, at 421–22 (noting those who have "argue[d] that the access/privacy tension can be resolved through the use of technology, including the use of computer programs to redact sensitive data elements or to anonymize information").

the technical tools now available. Rule sets could be finely tuned to surrounding conditions and, as prescribed by contextual integrity, drawn from prior convention honed by context specific values and goals.[311] This promising option would require that systems currently being designed for state courts be fitted with necessary components.

VII. CONCLUSION: A HIDDEN VARIABLE

There are clearly powerful voices in favor of open court records.[312] Despite this legacy, court records, according to our color-coding scheme, have never been green but always orange, accommodating various reasons for differentially restricting access to fields of information within them.[313] With the stakes raised by radical changes in access and dissemination, we have argued in favor of imposing appropriate constraints to compensate for these changes. Although additional constraints constitute a change in practice, the goal of such change is to sustain the underlying interests and values at stake. We may be defying convention but we do so for good reason, just as a good general dares not apply conventional strategy designed for a world with cannons to a world with ballistic missiles.

But the departure from convention we have suggested, particularly in a system that has evolved over centuries and whose very labeling of cases inextricably binds them to parties' names, calls for careful attention to counter arguments. We have considered the interests of parties and other participants and we have considered goals and values of the context; neither of these offers overwhelming counterweight. Although convention itself surely counts for something, resistance to change that would comprehensively block or diminish access to named individuals is rooted in something else, another factor that has not been fully acknowledged nor reckoned with.

Among the staunchest supporters of unrestricted online access are data aggregators, such as Choicepoint and LexisNexis. Although they align themselves with supporters of openness, transparency, citizen oversight, and accountability, data aggregators mine personal information in public, including court records, for different reasons en-

311. *See* THE SEDONA GUIDELINES: BEST PRACTICES ADDRESSING PROTECTIVE ORDERS, CONFIDENTIALITY & PUBLIC ACCESS IN CIVIL CASES 46–49 (Ronald J. Hedges & Kenneth J. Withers eds., 2005) (providing an overview of "Privacy and Public Access to the Courts in an Electronic World").

312. *See, e.g.*, Mulvaney, *supra* note 93 (reporting ACLU opposition to a proposed change to a rule "that would allow people to seek to block certain information from appearing in court records available over the Internet").

313. *See supra* Part V.A (discussing "green," "red," and "orange" records).

tirely.[314] Their business is assembling into dossiers as much information as possible about individuals, touting the utility of this service to all of hardworking, law-abiding society.[315] Is it not useful to know whether a neighbor has committed fraud, a job applicant has embezzled funds, a suitor has undergone a nasty divorce, a political candidate has been sued for business corruption, or an insurance applicant has been convicted of drunk driving? But it is important to notice that something new has edged into the landscape. These reasons invoke a different value in support of faster, cheaper, and unconstrained access to full court records online having little to do with openness, transparency, citizen oversight, and accountability. By their light, court records are valuable because they constitute a repository of useful information about people.

As useful as court records may be as sources for aggregated repositories of information about people, we are not convinced that an argument has yet been made that prioritizes this function in shaping access policies to court records. Indeed, explicit acknowledgment of it, evaluation of its legitimacy as a policy driver, and its relative benefit weighed against other goals and values, has been conspicuous in its absence. In bringing to light this "hidden variable," we recognize as precursor Justice Holmes's justification of open records,

> not because the controversies of one citizen with another are of public concern, but because it is of the highest moment that those who administer justice should always act under the sense of public responsibility, and that every citizen should be able to satisfy himself with his own eyes as to the mode in which a public duty is performed.[316]

In other words, satisfying with our "own eyes" that public duty is being properly performed is primary; prying into the private disputes of other citizens is a mere artifact of a system that must deliver the second with the first. The radical option might have satisfied Justice Holmes well.

Admittedly, we have not settled the question of whether court records *should* function as a repository of personal information for use by government or corporate actors as a basis for getting to know or

314. *See, e.g.*, Robert O'Harrow Jr., *In Age of Security, Firm Mines Wealth of Personal Data*, WASH. POST, Jan. 20, 2005 at A1 (reporting on Choicepoint's private intelligence activities).

315. *See* DANIEL J. SOLOVE, THE DIGITAL PERSON: TECHNOLOGY AND PRIVACY IN THE INFORMATION AGE 16–26 (2004) (providing an overview of the history of private-sector databases).

316. Cowley v. Pulsifer, 137 Mass. 392, 394 (1884).

vetting people of interest. Instead, we have endeavored only to expose the role that this function plays in influencing the shape of policy. The question deserves close and explicit scrutiny as a public matter. Should it be judged legitimate for court records to function as repositories, the next urgent matter would be to articulate the courts' responsibilities as creators of databases of personal information: Is the role curatorial? Are courts trustees and guardians? What is the extent of their obligation to assure accuracy and security? If information in records is decisively affecting people's lives during or after the completion of court cases, these matters are crucial to settle. In defining the role of courts in relation to third-party information intermediaries, it should also be discussed, as a public matter, what the respective obligations are of both parties—courts and commercial consumers of court records—in maintaining, sharing, and using information in records.