



2024

Access to Justice as Access to Data

Tanina Rostain

Georgetown University Law Center, tr238@law.georgetown.edu


This paper can be downloaded free of charge from:

<https://scholarship.law.georgetown.edu/facpub/2619>

<https://ssrn.com/abstract=4905900>

Northwestern Law Review, Volume 119, Issue 1, Pp. 5-22.

This open-access article is brought to you by the Georgetown Law Library. Posted with permission of the author.
Follow this and additional works at: <https://scholarship.law.georgetown.edu/facpub>

 Part of the [Courts Commons](#), [Housing Law Commons](#), [Intellectual Property Law Commons](#), and the [State and Local Government Law Commons](#)

Keynote

ACCESS TO JUSTICE AS ACCESS TO DATA

Tanina Rostain

ABSTRACT—This Keynote Address, delivered in celebration of the launch of SCALES, discusses the importance of making local and state court data available for research on the functioning of the American civil justice system. It describes the regulatory and administrative challenges of obtaining high-quality data from courts. It calls for a concerted effort among researchers and policymakers to develop open-source technologies for the development of case management systems and data infrastructure. And it urges researchers to foster a collaborative research ecosystem based on broadly sharing court data.

AUTHOR—Agnes C. Williams Sesquicentennial Professor of Justice Innovation, Georgetown University Law Center. I am grateful to the organizers of the *Northwestern University Law Review’s* Data Justice Symposium for inviting me to give the Keynote Address. Thanks are also due to Erika Rickard, Neel Sukhatme, Daniel Wilf-Townsend, and Zachary Zarnow for their helpful comments. Special thanks to Amy O’Hara, my collaborator on the Georgetown Civil Data Commons, and Jason Tashea, my partner on the Judicial Innovation Fellowship program, for many helpful conversations.

- INTRODUCTION 6
- I. WHY NOT A DATA COMMONS FOR CIVIL COURT DATA?..... 11
- II. INSTITUTIONAL BARRIERS TO OBTAINING DATA..... 13
- III. DIRTY, INCOMPLETE, AND DIFFICULT TO INTERPRET DATA 14
- IV. COURT-CENTERED VS. PEOPLE-CENTERED DATA COLLECTION PRACTICES 15
- V. THE PROBLEM OF VENDOR CAPTURE AND THE NEED FOR
AN OPEN-SOURCE COURT ECOSYSTEM..... 17
- VI. CREATING A COLLABORATIVE RESEARCH ECOSYSTEM..... 20

INTRODUCTION

As most of you know, I am not a data scientist. I come to the question of justice and data from the perspective of increasing access to civil justice. Data, and specifically court data, are important to understand how well the justice system serves the people caught up in it and to assess whether interventions effectively improve access to the court system. As we all recognize, the court-data landscape is, to use a technical phrase, a “hot mess.” Important initiatives such as the Systematic Content Analysis of Litigation Events (SCALES) are beginning to bring coherence and clarity to federal court data. Today, I want to add to the conversation by talking about the promise of state and local court data, the challenges to getting quality data, and the actions we might take together to grow the field of research.

In 2018, the Princeton Eviction Lab, led by sociologist Matthew Desmond, released the first big study of eviction rates in the United States.¹ Announced under a banner headline in the *New York Times* that included a large heat map of the United States and graphs listing the cities with the highest filing rates, the study described an eviction crisis that spanned the country.² Two years earlier, Desmond’s book *Evicted: Poverty and Profit in the American City* had brought national attention to evictions in Milwaukee by giving them a human face—a single mother, typically Black, living in poverty.³ The 2018 study, based on eighty-three million court records, gave evictions both a national and local geography.⁴ North Charleston topped the list of cities with the highest eviction rates⁵ and soon became known in the South Carolina press as “America’s eviction capital.”⁶

¹ Emily Badger & Quoc Trung Bui, *In 83 Million Eviction Records, a Sweeping and Intimate New Look at Housing in America*, N.Y. TIMES (Apr. 7, 2018), <https://www.nytimes.com/interactive/2018/04/07/upshot/millions-of-eviction-records-a-sweeping-new-look-at-housing-in-america.html> [<https://perma.cc/2RCS-H3UB>].

² *Id.*

³ MATTHEW DESMOND, *EVICTED: POVERTY AND PROFIT IN THE AMERICAN CITY* (2016).

⁴ See Badger & Bui, *supra* note 1.

⁵ *Id.*

⁶ See Thad Moore, *SC’s Only Major Attempt to Fix North Charleston’s Eviction Crisis Isn’t Enough*, POST & COURIER (Dec. 27, 2020), https://www.postandcourier.com/business/real_estate/sc-s-only-major-attempt-to-fix-north-charleston-s-eviction-crisis-isn-t-enough/article_68bb54d4-3bda-11eb-83d9-832a639530db.html [<https://perma.cc/CFU4-PT88>]; Hannah Alani, *Eviction Rates in North Charleston, Columbia Some of the Highest in U.S., Study Says*, POST & COURIER (Dec. 28, 2022), https://www.postandcourier.com/news/eviction-rates-in-north-charleston-columbia-some-of-the-highest-in-u-s-study-says/article_bebe69e2-3cd3-11e8-b412-331e6a815e34.html [<https://perma.cc/SH5Q-CV34>].

Beyond putting eviction on the national radar and spurring conversations about the affordable housing crisis in the United States,⁷ the Eviction Lab's findings prompted waves of new policies and interventions.⁸ Two years later, with the onset of the pandemic, the federal government and states implemented rental assistance programs.⁹ The Centers for Disease Control and Prevention and various states also imposed eviction moratoriums.¹⁰ The increased focus on evictions put wind in the sails of a

⁷ See Jason DeParle, *As Need Rises, Housing Aid Hits Lowest Level in Nearly 25 Years*, N.Y. TIMES (Dec. 19, 2023), <https://www.nytimes.com/2023/12/19/us/politics/housing-aid-rent-costs.html> [<https://perma.cc/P82L-CYTS>] (noting that housing assistance for the poorest tenants had fallen to the lowest level in nearly a quarter-century); see also Justine McDaniel & Adela Suliman, *Homelessness Soars by Record 12 Percent as Covid Support Ends, HUD Says*, WASH. POST (Dec. 16, 2023), <https://www.washingtonpost.com/nation/2023/12/16/homeless-hud-data-housing/> [<https://perma.cc/ZMP6-278R>]; Will Parker, *Wave of Rental Resets to Further Deplete Affordable Housing*, WALL ST. J. (July 10, 2023), <https://www.wsj.com/articles/wave-of-rental-resets-to-further-deplete-affordable-housing-62c8d063> [<https://perma.cc/7D8P-9H2N>]; Andra Ghent, *The Affordable Housing Crisis in 2023: Where Do We Stand, and What Are the Solutions*, KENAN INST. OF PRIV. ENTER. (Jan. 24, 2023), <https://kenaninstitute.unc.edu/commentary/the-affordable-housing-crisis-in-2023-where-do-we-stand-and-what-are-the-solutions/> [<https://perma.cc/H8CV-6R6W>]; Alan Mallach, *The United States Housing Affordability Crisis: No Easy Solutions*, ICMA (Dec. 1, 2023), <https://icma.org/articles/pm-magazine/united-states-housing-affordability-crisis-no-easy-solutions> [<https://perma.cc/6ZL6-CMFZ>].

⁸ See, e.g., Solomon Green, Todd M. Richardson, Jemine A. Byron & Richard Cho, *Rise in Homelessness Averted Amidst Worsening Housing Needs in 2021. What Does This Tell Us About How to End Homelessness in the U.S.?*, HUD USER: PD&R EDGE (Aug. 22, 2023), <https://www.huduser.gov/portal/pdredge/pdr-edge-frm-asst-sec-082223.html> [<https://perma.cc/59VH-8F5U>] (referencing Eviction Lab research finding that COVID-19 policies cut eviction filings by more than half); Aaron Schroyer, *Tracking the Impact of the CDC Eviction Moratorium*, U.S. DEP'T. OF HOUS. & URB. DEV. OFF. OF POL'Y DEV. & RSCH. (Mar. 22, 2021), <https://www.huduser.gov/portal/pdredge/pdr-edge-trending-032221.html> [<https://perma.cc/7SRM-JW8S>] (citing Eviction Lab research assessing the degree to which eviction moratorium reduced evictions); U.S. GOV'T ACCOUNTABILITY OFF., GAO-21-370, COVID-19 HOUSING PROTECTIONS: MORATORIUMS HAVE HELPED LIMIT EVICTIONS, BUT FURTHER OUTREACH IS NEEDED (Mar. 2021), <https://www.gao.gov/assets/gao-21-370.pdf> [<https://perma.cc/3PUG-YR9D>] (listing the Eviction Lab as one researcher interviewed for the GAO's report).

⁹ *Treasury Emergency Rental Assistance (ERA) Dashboard*, NAT'L LOW INCOME HOUS. COAL., <https://nlihc.org/era-dashboard> [<https://perma.cc/4KYZ-LKP5>] (demonstrating that the Emergency Rental Assistance program and the American Rescue Plan Act cover every state in the Union, the District of Columbia, Guam, Puerto Rico, the Virgin Islands, and the Commonwealth of the Northern Mariana Islands).

¹⁰ See Temporary Halt in Residential Evictions to Prevent the Further Spread of COVID-19, 85 Fed. Reg. 55292, 55292-97 (Sept. 4, 2020) (halting evictions through December 31, 2020); Continuing Temporary Suspension and Modification of Laws Relating to the Disaster Emergency, 9 C.R.R.-N.Y. § 8.202.66 (Sept. 29, 2020) (temporarily prohibiting evictions in New York); S.B. 91, 2021 Leg., Reg. Sess. (Cal. 2021) (extending eviction protections and establishing California's Emergency Rental Assistance Program). The Center for Disease Control's (CDC's) eviction moratorium was extended several times before the U.S. Supreme Court struck it down on August 26, 2021. See *Ala. Ass'n of Realtors v. Dep't of Health & Hum. Servs.*, 594 U.S. 758, 766 (2021); Adam Liptak & Glenn Thrush, *Supreme Court Ends Biden's Eviction Moratorium*, N.Y. TIMES (Nov. 7, 2021), <https://www.nytimes.com/2021/08/26/us/eviction-moratorium-ends.html> [<https://perma.cc/A6S7-YMR5>] (reporting on extensions of state's and localities' eviction moratoriums after the U.S. Supreme Court struck down the CDC's eviction moratorium).

national tenants’ right-to-counsel movement, which continues to gain momentum today. As of spring 2024, seventeen cities, four states, and one county recognize a right to counsel in eviction cases.¹¹

The study also prompted local efforts to address housing insecurity. At the time of the study, the Charleston area prided itself on its great restaurants and new Southern hip-millennial sensibility. That image was shattered when the study revealed it was also home to the city with the highest eviction rate in the country.¹² The local bar, legal aid groups, and courts went into action.¹³ A new specialized housing court launched, and more resources were dedicated to eviction defense.¹⁴ Fast-forward five years and the South Carolina Supreme Court, in a highly unusual move, has granted an exception to its unauthorized practice of law prohibition to allow housing navigators—nonlawyers—to provide legal advice to defendants in eviction cases.¹⁵

While we do not know how effective these initiatives have been, publication of the Eviction Lab data transformed discussions about eviction, the right to housing, and the housing affordability crisis in the United States. In the last five years, studies based on court data have led to discussions and reforms in other areas. Research has documented the rapidly increasing number of debt cases in state courts, the expanding use of the court system by large debt collectors, and the dramatic rise of “frequent-filer” plaintiffs responsible for hundreds of thousands of filings.¹⁶ Researchers have also

¹¹ *Current Tally of Tenant Right to Counsel Jurisdictions: 17 Cities, 4 States, 1 County!*, NAT’L COAL. FOR CIV. RIGHT TO COUNS., http://civilrighttocounsel.org/highlighted_work/organizing_around_right_to_counsel [https://perma.cc/27SS-68DM]. In 2023, Delaware took a different approach and enacted a “right to representation,” which is not limited to representation by an attorney. *See* DEL. CODE ANN. tit. 25, § 5601 (2024) (stating a new rule that allows lay advocates to represent tenants); *see also* DEL. SUP. CT. R. 57.1 (permitting Qualified Tenant Advocates to represent tenants).

¹² *See Top Evicting Large Cities in the United States*, EVICTION LAB, <https://evictionlab.org/rankings> [https://perma.cc/Z63F-EY7U].

¹³ *See Housing Court Pilot Resources*, S.C. BAR, <https://www.sccourts.org/courtOrders/displayOrder.cfm?orderNo=2019-05-24-01> [https://perma.cc/RD79-W39H].

¹⁴ *See, e.g.*, Charleston Hous. Ct. Pilot Project, No. 2019-05-24-01 (S.C. May 24, 2019), <https://www.sccourts.org/courtOrders/displayOrder.cfm?orderNo=2019-05-24-01> [https://perma.cc/2TUU-UXZJ] (establishing the Charleston Housing Court Pilot Project); Charleston Hous. Ct. Expansion, No. 2023-01-26-01 (S.C. Jan. 26, 2023), <https://www.sccourts.org/courtOrders/displayOrder.cfm?orderNo=2023-01-26-01> [https://perma.cc/F2JF-QGSS] (describing the 2021 expansion of the Charleston Housing Court Pilot Project).

¹⁵ *See* S.C. NAACP Hous. Advoc. Program, No. 2023-001608, 2024 WL 706422, at *5 (S.C. Feb. 8, 2024).

¹⁶ *See* Daniel Wilf-Townsend, *Assembly-Line Plaintiffs*, 135 HARV. L. REV. 1704, 1706 (2022); *see also How Debt Collectors Are Transforming the Business of State Courts*, PEW (May 6, 2020), <https://www.pewtrusts.org/en/research-and-analysis/reports/2020/05/how-debt-collectors-are-transforming-the-business-of-state-courts> [https://perma.cc/Z9S4-54Z7] (studying specific states); UTAH BAR FOUND., UTAH BAR FOUNDATION REPORT ON DEBT COLLECTION AND UTAH’S COURTS 11

combined court data with other public data sources to document the rise of imprisonment for failure to pay court debt.¹⁷

In the criminal sphere, the organization Measures for Justice collects and analyzes criminal justice data so that communities can understand the criminal justice system's impact on their members' lives.¹⁸ Researchers at the Criminal Justice Administration Records System, a large criminal justice data repository, have linked administrative and court data sets to examine, among other questions, the connections between access to public benefits and criminal justice involvement.¹⁹ My colleague Neel Sukhatme is using Florida court data to re-enfranchise people convicted of felonies in Florida and other states.²⁰ He scrapes data to determine whether these individuals still owe fines and assists them in paying the remaining balance.²¹ His initiative has shown that eliminating outstanding fines and educational efforts can increase voter participation among people convicted of felonies by 25%.²²

As these examples show, court data studies can shed light on important social problems such as housing, debt, and criminal justice. They also have the potential to illuminate how well courts work, how fair they are, how well

(Apr. 2022), <https://www.utahbarfoundation.org/static/media/UBF2022.912d30c10e5681bf5f8c.pdf> [https://perma.cc/ZQ99-YKFT] (describing how companies make up the majority of plaintiffs); PEW CHARITABLE TRS., HOW DEBT COLLECTION WORKS IN PHILADELPHIA'S MUNICIPAL COURTS I (Oct. 2022), <https://www.pewtrusts.org/-/media/assets/2022/10/how-debt-collection-works-in-philadelphias-municipal-court.pdf> [https://perma.cc/GA7Q-Z9HD] (analyzing debt collection cases in Philadelphia from 2013 to 2018); MICH. JUST. FOR ALL COMM'N, ADVANCING JUSTICE FOR ALL IN DEBT COLLECTION LAWSUITS (Oct. 2022), https://www.courts.michigan.gov/4ac33d/siteassets/reports/special-initiatives/justice-for-all/jfa_advancing_justice_for_all_in_debt_collection_lawsuits.pdf [https://perma.cc/PQ96-PAM4].

¹⁷ See Johann D. Gaebler, Phoebe Barghouty, Sarah Vicol, Cheryl Philips & Sharad Goel, *Forgotten but Not Gone: A Multi-State Analysis of Modern-Day Debt Imprisonment*, 18 PLOS ONE, Sept. 13, 2023, at 1, <https://doi.org/10.1371/journal.pone.0290397> [https://perma.cc/H7DX-NHCE] (analyzing a first-of-its-kind dataset that documents imprisonment for court debts).

¹⁸ See *Helping Communities Reshape the Criminal Justice System*, MEASURES FOR JUST., <https://measuresforjustice.org/> [https://perma.cc/74NJ-9WYA].

¹⁹ See generally Manasi Deshpande & Michael Mueller-Smith, *Does Welfare Prevent Crime? The Criminal Justice Outcomes of Youth Removed from SSI*, 137 Q.J. ECON. 2263 (2022) (linking Supplemental Security Income (SSI) and criminal justice records to show that SSI removal increases number of criminal charges by a statistically significant 20% over the next two decades); Keith Finlay & Michael Mueller-Smith, *Justice-Involved Individuals in the Labor Market Since the Great Recession*, 695 ANNALS AM. ACAD. POL. & SOC. SCI. 107, 107–22 (2021).

²⁰ See Neel U. Sukhatme, Alexander Billy & Gaurav Bagwe, *Felony Financial Disenfranchisement*, 76 VAND. L. REV. 143, 151 (2023); see also FREE OUR VOTE, <https://freeourvote.com/> [https://perma.cc/TFE9-EXDT]; Alexander Billy, J.J. Naddeo & Neel U. Sukhatme, *Felony Disenfranchisement and Voter Turnout: Randomized Trials in Iowa and Washington*, 119 Nw. U. L. Rev. 221 (2024).

²¹ Sukhatme, Billy & Bagwe, *supra* note 20, at 151.

²² *Id.*

they follow the law, and especially how they treat unrepresented people.²³ Consider how judges determine whether a litigant can proceed in forma pauperis. In 2018, Professor Andrew Hammond published an important study comparing the different and inconsistent information that courts were collecting to rule on in forma pauperis petitions.²⁴ His database was a large collection of forms he hand-coded.²⁵ A few years later, using the federal court data available via SCALES, researchers could now begin to document how judges ruled on in forma pauperis petitions.²⁶ Their work suggested that there was huge variation within districts in judges' rates granting these petitions and raised the serious possibility that judges in the same districts were applying different standards under the same rules.²⁷

SCALES is an extraordinary achievement and a great cause for celebration. Bringing it to life took significant resources and coordination among several talented researchers. It is gratifying to know that, with support from the National Science Foundation (NSF), the efforts to build an open federal court data system will continue and grow. As SCALES shows, rationalizing and making federal court data available for research is essential for advancing civil and criminal justice reform. This is equally true of state and local court data, where the challenges are even greater, and the coordinated effort required to capture, clean, and harmonize the data will be even more Herculean than with SCALES.

In the United States, there are tens of thousands of state, county, municipal, and tribal courts. There are courts of general and limited jurisdiction; courts of equity and law; circuit courts, small claims courts, family courts, housing-conditions courts, and eviction courts. More than eighty million cases are filed in these courts every year.²⁸ In comparison, initial filings in federal civil, criminal, and bankruptcy courts number fewer

²³ See Tanina Rostain & Amy O'Hara, *The Civil Justice Data Gap*, in LEGAL TECH AND THE FUTURE OF CIVIL JUSTICE 368–88 (David Freeman Engstrom ed., 2023); Emily S. Taylor-Poppe, *Courts as Data Guardians for the Public Good*, 73 U. TORONTO L.J. 34, 36 (2023).

²⁴ See Andrew Hammond, *Pleading Poverty in Federal Court*, 128 YALE L.J. 1478, 1482 (2018).

²⁵ *Id.* at 1539.

²⁶ Adam R. Pah, David L. Schwartz, Sarath Sanga, Zachary D. Clopton, Peter DiCola, Rachel Davis Mersey, Charlotte S. Alexander, Kristian J. Hammond & Luis A. Nunes Amaral, *How to Build a More Open Justice System*, 369 SCIENCE 134, 135 (2020). For a discussion of some of the complexities in interpreting this data, see Judith Resnik, Henry Wu, Jenn Dikler, David T. Wong, Romina Lilollari, Claire Stobb, Elizabeth Beling, Avital Fried, Anna Selbrede, Jack Sollows & Julia Udell, *Lawyerless Litigants, Filing Fees, Transaction Costs, and the Federal Courts: Learning from SCALES*, 119 NW. U. L. REV. 111, 134–36 (2024).

²⁷ See Pah et al., *supra* note 26, at 135.

²⁸ See INST. FOR THE ADVANCEMENT OF THE AM. LEG. SYS., JUSTICE WE CAN BELIEVE IN: ANNUAL REPORT 1, 9 (2019) (reporting the number of court cases filed in 2017).

than one million per year.²⁹ State and local courts are where people are sued for failure to pay rent and credit card, student, and medical debt. They are where people are charged with traffic violations, misdemeanors, and felony crimes. State and local courts are where people go to solve marital and custody disputes and to get restraining orders to protect themselves against domestic violence. For most people, state and local courts are the face of the justice system. And what they see is not pretty. Civil court dockets are dominated by housing and debt cases, in which over 90% of defendants are unrepresented and in which over 70% of debt cases end in default.³⁰ Representation is also scarce in family court, and the same is true for courts dealing with traffic violations, misdemeanors, and domestic violence cases.³¹ State and local courts have looked like this for several decades, but access to data to document these problems and their implications for the people enmeshed in state courts remains limited.

I. WHY NOT A DATA COMMONS FOR CIVIL COURT DATA?

In 2019, Amy O'Hara, an economist at Georgetown's Massive Data Institute, and I began to explore the creation of a Civil Justice Data Commons (CJDC). The CJDC would be modeled after Stanford's Center for Population Health Sciences (CPHS), an extraordinary resource for research on population health that Amy helped build.³² CPHS makes hundreds of millions of health records available to researchers in a single place to study population health.³³ If sensitive health data could be collected and harmonized in one secure repository, we reasoned, how hard would it be to launch a similar effort for open court data?

²⁹ See JOHN G. ROBERTS JR., 2023 YEAR-END REPORT ON THE FEDERAL JUDICIARY 10–11 (2023), <https://www.supremecourt.gov/publicinfo/year-end/2023year-endreport.pdf> [<https://perma.cc/8GBQ-YNQF>]; *Federal Judicial Caseload Statistics 2022*, U.S. CTS., <https://www.uscourts.gov/statistics-reports/federal-judicial-caseload-statistics-2022> [<https://perma.cc/NX53-WLH5>].

³⁰ See *How Debt Collectors Are Transforming the Business of State Courts*, *supra* note 16, at 1; see also Casey Chiappetta, *Debt Collection Cases Continued to Dominate Civil Dockets During Pandemic*, PEW (Sept. 18, 2023), <https://www.pewtrusts.org/en/research-and-analysis/articles/2023/09/18/debt-collection-cases-continued-to-dominate-civil-dockets-during-pandemic> [<https://perma.cc/6TH9-7C2T>].

³¹ See Pamela K. Bookman & Colleen F. Shanahan, *A Tale of Two Civil Procedures*, 122 COLUM. L. REV. 1183, 1186 (2022) (describing how at least one party is self-represented in 70% of state civil cases and 30% of federal civil cases); NATALIE ANNE KNOWLTON, LOGAN CORNETT, CORINA D. GERETY & JANET L. DROBINSKE, INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., *CASES WITHOUT COUNSEL: RESEARCH ON EXPERIENCES OF SELF-REPRESENTATION IN U.S. FAMILY COURT 1* (May 2016), https://iaals.du.edu/sites/default/files/documents/publications/cases_without_counsel_research_report.pdf [<https://perma.cc/LJ9Y-F6TM>].

³² See *Stanford Center for Population Health Sciences*, STAN. UNIV., <https://redivis.com/StanfordPHS> [<https://perma.cc/CZ9B-LZEW>].

³³ *Id.*

With the support of an NSF grant,³⁴ we conducted dozens of interviews with stakeholders in four jurisdictions to understand both interest in and obstacles to sharing data. Some of the things we heard were predictable: Many of our interviewees agreed that sharing data was a good thing in principle, but pointed to obstacles such as the need for dedicated resources, disaggregated systems that do not speak to each other, poorly designed data collection processes, and the fear of reputational fallout for courts and judges from making data freely available.³⁵

One interview, conducted early in the pandemic, stood out as an indicator of the obstacles ahead. At the outset of the interview, an Ohio judge told us that she was concerned about the pandemic because it would reduce the number of eviction filings in her court, a principal source of revenue for court operations. She also told us that her court did not routinely collect data. Instead, someone in her court downloaded eviction cases from the court's case management system and sold the data in bulk to data aggregators. When we asked her whether she was interested in making data available to researchers to determine whether systemic patterns in the court's decisions might show bias, she said it was an interesting question but not one that directly concerned her. Her role—and the role of all the judges sitting on her court—was to decide each case on its own merits, one at a time.

After our interviews, we prototyped and launched the Georgetown CJDC.³⁶ The goal was to provide facilitated and frictionless access to harmonized data that researchers could use to study civil justice problems within and across jurisdictions—for example, to compare eviction processes in two comparable cities in the Midwest. In its first iteration, the CJDC would obtain data from courts under data sharing agreements, then clean and harmonize this data before making it available through a secure platform to researchers interested in studying courts. From our interviews, it was clear that we should emphasize a what's-in-it-for-the-court approach to persuade courts to share data. To incentivize courts, we offered to create dashboards that courts could use to understand and track court activity. We also decided

³⁴ *SCC-PG: Just Connect Us: A Community-Oriented Civil Justice Data Commons*, U.S. NAT'L SCI. FOUND., (June 28, 2022) https://www.nsf.gov/awardsearch/showAward?AWD_ID=1952067&HistoricalAwards [<https://perma.cc/YF96-WRQT>] (granting almost \$150,000 to this project as NSF Award # 1952067).

³⁵ *See CIV. JUST. DATA COMMONS, CIVIL COURT DATA AT THE LOCAL LEVEL: INTERVIEWS AND INSIGHTS FROM FOUR LOCATIONS 7* (May 2022), <https://www.law.georgetown.edu/tech-institute/wp-content/uploads/sites/42/2022/06/Civil-Court-Data-at-the-Local-Level-Report.pdf> [<https://perma.cc/83XV-37AR>].

³⁶ *See Civil Justice Data Commons*, GEORGETOWN L. LIBR., <https://www.law.georgetown.edu/tech-institute/initiatives/georgetown-justice-lab/civil-justice-data-commons/> [<https://perma.cc/E88E-3FRS>] (noting that the launch of the CJDC was supported by funding from Pew Charitable Trusts).

to focus on judges committed to sharing data and target courts with rules indicating their data were presumptively open and available. We thought these were low-hanging fruit, but they proved to be outside our reach. Institutional barriers, including financial constraints and organizational disincentives, made it nearly impossible to obtain data from courts. And it turned out that the limited data available were often of poor quality, making them ill-suited for research.

II. INSTITUTIONAL BARRIERS TO OBTAINING DATA

As researchers are increasingly discovering, data sets that should be publicly available under court rules rarely are. Many courts have formal rules stating that court data are public,³⁷ and at least one court has held that access to court data is protected by the First Amendment's public right of access.³⁸ Such rules, however, in no way guarantee that these data are accessible.

Our experience with Connecticut highlights these challenges. In August 2021, we formally requested court data from the Chief Court Administrator, laying out what we needed and why under the applicable court rules.³⁹ We never succeeded, despite multiple attempts to get a response over ten months. (We bought them for a different project investigating medical debt, but under the terms of the data-sharing agreement, we cannot put them in the CJDC.) It turned out that Connecticut courts, like those in Ohio, had a strong interest in selling their court data under a restricted agreement, presumably to fund their operations, rather than provide the data for free.

Maryland, which also has a rule that makes its data broadly available, denied our request on the ground that it would impose a "significant and reasonably calculable burden on their Administrative Office of the Courts." Per the denial letter, fulfilling our request would take 217 hours.⁴⁰ That number is not necessarily inflated. Given the antiquated court data

³⁷ See, e.g., MO. CT. OPERATING R. 2.02(a) (stating that court records are presumed to be open); CAL. RULES OF CT. 2.503(a)(1) (stating that all electronic records must be made reasonably available to the public); MD. CODE ANN., CT. ADMIN. Rule 16-904(a) (West 2024) (stating that judicial records are presumed to be open to the public).

³⁸ S.C. State Conf., of the NAACP v. Kohn, No. 3:22-01007-MGL, 2023 WL 144447, at *7 (D.S.C. Jan. 10, 2023).

³⁹ See *Court Records: Frequently Asked Questions*, STATE OF CONN. JUD. BRANCH, <https://www.jud.ct.gov/faq/courtrec.html> [<https://perma.cc/8SUX-4T7K>] (stating that members of the public have a right to see any document in any court file).

⁴⁰ Letter from Bradley Tanner, Pub. Info. Officer, Md. Judiciary Dep't of Gov't Rels. & Pub. Affs., to Dr. Amy O'Hara, McCourt Sch. of Pub. Pol'y, Georgetown Univ. (May 5, 2022) (on file with *Northwestern University Law Review*); see Kat Albrecht & Kaitlyn Filip, *Public Records Aren't Public: Systemic Barriers to Measuring Court Functioning & Equity*, 113 J. CRIM. L. & CRIMINOLOGY 1, 13 (2023) (demonstrating the inability to obtain access to putatively public court transcripts in Cook County, Illinois).

infrastructure, the estimated time might realistically reflect the hours required to download and transfer the requested data.

Conversations with other courts suggested that time and money are not the only obstacles. Courts expressed concern that their data would be misused—to single out particular judges, for example.⁴¹ They were also concerned that the data would be misinterpreted to produce studies that negatively affect court performance. Additionally, we suspect that court data officers may have been embarrassed at the poor data quality.

III. DIRTY, INCOMPLETE, AND DIFFICULT TO INTERPRET DATA

When we successfully obtained data, they were incomplete, inconsistent, and full of anomalies and outliers.⁴² Many data-capture systems have open fields without standardization functions to input names and addresses, leading to inconsistencies, misspellings, and other errors in recording basic information.⁴³ In many instances, it is difficult to determine if the parties are individuals or corporate entities.⁴⁴ Many eviction dockets do not capture the address of the rental property, and often there are duplicate records.⁴⁵ Many of these problems are a product of manual entry processes.⁴⁶ Around the country, hundreds—if not thousands—of court clerks spend hours every day inputting data into antiquated technology systems.⁴⁷ These outdated processes lead to significant errors in the data.⁴⁸

In addition, there is the difficulty of interpreting and classifying court data. In many states, for example, there is the challenge of identifying consumer debt cases among those that fall under “contract,” understanding different case types that might meet the definition of “consumer debt,” and

⁴¹ See CIV. JUST. DATA COMMONS, *supra* note 35, at 8.

⁴² See Lauren Prunkl, Exploring Eviction Data Collection Methods (2022) (Masters Paper, University of North Carolina at Chapel Hill), https://cdr.lib.unc.edu/concern/masters_papers/2801pr62r [<https://perma.cc/3MZ4-3XFA>]; Dana Goplerud & Craig Pollack, *Prevalence and Impact of Evictions*, HUD OFF. OF POL’Y DEV. & RSCH. (2021), <https://www.huduser.gov/portal/periodicals/em/Summer21/highlight2.html> [<https://perma.cc/9YLT-YJMT>] (describing an analysis of eviction court records showing between 7% and 47% of records were incomplete or inaccurate); see also JAMES CAREY & MARGARET HAUGHNEY, THE CHALLENGES OF IDENTIFYING MEDICAL DEBT IN COURT RECORDS 5–9 (Apr. 2024), <https://georgetown.app.box.com/s/i0kzaqbimv4pv5xjb0d46vwsfb85zutr> [<https://perma.cc/G2ZB-GLL9>]. On the poor quality of federal court data, see Resnik et al., *supra* note 26, at 146 & n.114.

⁴³ See Prunkl, *supra* note 42, at 33–34.

⁴⁴ See AMY O’HARA & STEPHANIE STRAUS, LINKING DEMOGRAPHIC CHARACTERISTICS TO EVICTION RECORDS: A PRELIMINARY ANALYSIS 9 (2024) <https://georgetown.app.box.com/s/qamfrz0oacmheo2jhxhem1zh584fdhkm> [<https://perma.cc/QD4B-CGLA>].

⁴⁵ *Id.* at 10.

⁴⁶ See Prunkl, *supra* note 42, at 17.

⁴⁷ See, e.g., CIV. JUST. DATA COMMONS, *supra* note 35, at 15.

⁴⁸ See Prunkl, *supra* note 42, at 33.

interpreting case-type codes. These problems are not limited to consumer debt. Maryland has about twenty different case types that appear eviction-related.⁴⁹ Apparently, if you fail to pay rent on your mobile home, that is a different eviction type than failure to pay rent on your house; there are also “forcible detainers” and “grantor in possession” case types.⁵⁰ Exactly what the differences are and why they matter is a mystery.⁵¹ The problem of taxonomy becomes even more acute when you look at “disposition” fields. One court’s “disposition” is a ruling on a motion or after a hearing; another court’s “disposition” is a final judgment.⁵² In other courts, it may be something else altogether.⁵³ And it may even mean different things in the same dataset.⁵⁴

IV. COURT-CENTERED VS. PEOPLE-CENTERED DATA COLLECTION PRACTICES

A further challenge is that data are collected based on case units and have insufficient fields to link different cases to the same person. A person living in Evanston, Illinois, might have been involved in an eviction case, a debt case, and a traffic violation, but no unique identifier will connect those three matters. Other than the eviction case, which may have an address, other case types typically have no data besides a party’s name that would allow for linkages among them. This makes it nearly impossible to identify the same person in different court data sets.

The importance of identifying unique people with multiple interactions with the criminal and civil justice systems for understanding the impact of court entanglements cannot be overstated. Identifying people enmeshed in criminal and civil cases would produce knowledge about the collateral

⁴⁹ See THE CIRCUIT CT. FOR HOWARD CNTY., MD. 5TH CIRCUIT, CIVIL DIFFERENTIATED CASE MANAGEMENT PLAN (last updated Sept. 2022), <https://www.mdcourts.gov/sites/default/files/import/circuit/howard/pdfs/dcmcivilhoward2022.pdf> [<https://perma.cc/Z6BJ-9UBP>].

⁵⁰ See, e.g., *id.* (listing codes for several different case types); *Maryland Judiciary Case Search*, MD. JUDICIARY, <https://casesearch.courts.state.md.us/casesearch/inquiry-index.jsp> [<https://perma.cc/E26B-DJNC>] (expand the “Party Search” and then “Advanced Search” tabs, type “G” in the “Last Name” box, then input a filing date range of January 1, 2021 to January 1, 2024) (providing search results with a large number of case types used for filings with just a small subset of parties appearing before Maryland’s state courts); *Tribbett v. David G.*, No. D-036-LT-21-000061 (Md. Dist. Ct. May 4, 2021) (providing an example of a filing listed as case type “FTPR”).

⁵¹ See Prunkl, *supra* note 42, at 11.

⁵² See O’HARA & STRAUS, *supra* note 44, at 12. For a discussion of the poor quality of data, see Adam Porton, Ashley Gromis & Matthew Desmond, *Inaccuracies in Eviction Records: Implications for Renters and Researchers*, 31 HOUS. POL’Y DEBATE 377 (2021).

⁵³ See Prunkl, *supra* note 42, at 35; see also O’HARA & STRAUS, *supra* note 44, at 12.

⁵⁴ See Prunkl, *supra* note 42, at 34.

consequences of criminal justice, eviction, and debt cases.⁵⁵ It would also allow for more holistic interventions to help people and provide data to strengthen the work of new problem-solving courts.⁵⁶ In a world where court data are linked to other administrative datasets, it would be possible to understand the effects of upstream interventions on criminal and civil court involvement.⁵⁷

In a recent article, Professors Charlotte Alexander and Lauren Sudeall documented the extensive technical work required to identify unique people who appeared in the criminal and civil court records of a single county in Georgia between 2016 and 2020.⁵⁸ Over half the people in their study had more than one case, with numbers ranging from 2 to 28. The average number of cases per person was 3.2.⁵⁹ This should not come as a surprise: It stands to reason that people with criminal cases are likely to have civil legal problems too. To track people easily through administrative and court systems and formulate responsive policies—impossible tasks with current court data—it is necessary to understand the relationships between civil and criminal justice problems and the effect of court entanglement on people’s lives.

One approach to solving the problem of access to better quality data is to focus upstream on how data are collected. The Judicial Innovation Fellowship (JIF) program, housed at Georgetown’s Institute for Technology Law and Policy, attempts to do just that.⁶⁰ The program puts technologists and designers in courts to bring them into the twenty-first century.⁶¹ Professor Kat Albrecht, a principal investigator at SCALES⁶² and one of the program’s inaugural fellows, is working to integrate criminal justice, housing assistance, and eviction records in Chattanooga, Tennessee, so that the city can more effectively address the related problems of recidivism and homelessness.⁶³ Making these datasets interoperable and persuading

⁵⁵ See Rostain & O’Hara, *supra* note 23, at 368, 369.

⁵⁶ Charlotte S. Alexander & Lauren Sudeall, *Creating a People-First Court Data Framework*, 58 HARV. C.R.-C.L.L. REV. 731, 742–48 (2023).

⁵⁷ See Rostain & O’Hara, *supra* note 23, at 377.

⁵⁸ See Alexander & Sudeall, *supra* note 56, at 748–59.

⁵⁹ *Id.* at 759.

⁶⁰ See *Fellows & Projects*, GEORGETOWN L. INST. FOR TECH. & POL’Y, <https://www.law.georgetown.edu/tech-institute/initiatives/georgetown-justice-lab/judicial-innovation/projects/> [<https://perma.cc/P6HF-K5KX>] (describing the goals of JIF and introducing its members).

⁶¹ *See id.*

⁶² See *Our Team*, SCALES, <https://scales-okn.org/team-2/> [<https://perma.cc/Q9ZU-87D3>] (listing Professor Kat Albrecht as a team member at SCALES).

⁶³ See *Fellows & Projects*, *supra* note 60; HAMILTON CNTY. & JUD. INNOVATION FELLOWSHIP, STATEMENT OF WORK, https://docs.google.com/document/d/1F-vVUWN5k4iq0o09wgpE7PVUBH6NDZS_reD-UC_VUZ8/edit [<https://perma.cc/88MD-LBFG>].

agencies and courts to coordinate their data collection and sharing practices are two principal challenges of the project.

V. THE PROBLEM OF VENDOR CAPTURE AND THE NEED FOR AN OPEN-SOURCE COURT ECOSYSTEM

Our preparatory research for JIF revealed that vendor capture is one of the major impediments to court technology modernization.⁶⁴ One or two private companies dominate the court technology market.⁶⁵ With annual revenues approaching \$2 billion dollars, Tyler Technologies is the largest technology company exclusively serving the public sector, and it has rolled out its systems in hundreds of county and state courts.⁶⁶ As court informants reported, Tyler and other court vendors encourage new court clients to implement high-cost bespoke systems, which leads to a lack of standardization. One former judge noted that

Tyler’s business model—and that of other court vendors—turns on tailoring their case management systems to a court’s highly individualized requirements. In so doing, vendors exploit each court’s tendency to view its operations as unique and independent of the larger system. As a result of courts’ exaggerated perceptions of institutional differences as a function of local court culture, vendors charge as though each system were uniquely bespoke despite widespread similarity. In this process, the principal benefits of standardization—lower costs and interoperability—that technologies confer are lost.⁶⁷

Consistent with this tailored approach, court vendors impose significant costs on courts with existing contracts to add new features and data fields.

⁶⁴ See JASON TASHEA, GEORGETOWN L. INST. FOR TECH. & POL’Y, THE JUDICIAL INNOVATION FELLOWSHIP: A ROADMAP TO STRENGTHEN STATE, LOCAL, TERRITORIAL, AND TRIBAL COURTS 6 (2023), <https://www.law.georgetown.edu/tech-institute/wp-content/uploads/sites/42/2023/02/Judicial-Innovation-Fellowship-Roadmap-1.pdf> [<https://perma.cc/ZG34-MYNW>]; see also David Freeman Engstrom & R.J. Vogt, *The New Judicial Governance: Courts, Data, and the Future of Civil Justice*, 72 DEPAUL L. REV. 171, 194 (2023).

⁶⁵ As of 2020, over half of U.S. states (twenty-six) used Tyler Technologies’ Odyssey Case Manager system, and fifteen of those states implemented Odyssey statewide. See Jennifer Kepler, *Tyler Technologies Announces Successful Go-Live for Odyssey Case Management Solution in Cook County, Illinois, Criminal Courts*, TYLER TECHS. (Apr. 16, 2020), <https://investors.tylertech.com/news/news-details/2020/Tyler-Technologies-Announces-Successful-Go-Live-for-Odyssey-Case-Management-Solution-in-Cook-County-Illinois-Criminal-Courts-04-16-2020/default.aspx> [<https://perma.cc/K2H7-PQNT>].

⁶⁶ For a discussion of the larger implications of Tyler’s dominance in the court-technology field, see Todd Venook, *Enterprise Justice: Tyler Technologies and the Private Provision of Court Services* (on file with author).

⁶⁷ E-mail from Jennifer Bailey, former J., Thirteenth Circuit (Kanawha Cnty., W. Va.), to author (Mar. 12, 2024, 1:21 PM) (on file with *Northwestern University Law Review*); see also TASHEA, *supra* note 64, at 9.

In addition to being unnecessarily costly, Tyler has a pattern of lengthy delays and significant malfunctions with its court technology rollouts, which it has little incentive to address.⁶⁸ Problems have included wrongful arrests when Tyler systems automatically reopened arrest warrants that had been closed and delayed prison releases for people who had finished serving their sentences.⁶⁹ Despite these publicly documented problems, courts have few alternatives. For example, the North Carolina courts chose Tyler for an \$83 million upgrade, even though court administrators were aware of the company's history of delays, cost overruns, and more serious problems.⁷⁰ After a year-long delay and \$15 million in additional costs, the new system was rolled out in early 2023, only to produce errors and challenges for lawyers and court clerks trying to use the system.⁷¹ Again, the system automatically opened warrants that had been closed and led to the unlawful detention of people for several additional weeks due to system failures.⁷²

One way to loosen the grip of vendors in the court technology space is to develop an open-source case management system. The Open Courts Act of 2023, the newest version of a bill previously sponsored by Oregon

⁶⁸ See, e.g., Heather Hollingsworth & John Hanna, *Kansas Court System Down Nearly 2 Weeks in 'Security Incident' That Has Hallmarks of Ransomware*, AP NEWS (Oct. 25, 2023), <https://apnews.com/article/kansas-courts-cybersecurity-problem-network-offline-ba6c2c1979fd58869cc5d4fb1aed501c> [<https://perma.cc/WW9M-YYMT>].

⁶⁹ See, e.g., Fifth Amended Class Action Complaint, *Turnage v. Oldham*, No. 2:16-CV-2907-SHM/tmp., 2019 WL 1369836 (W.D. Tenn. Feb. 26, 2019) (alleging that system error resulted in delayed prison releases); Russ McQuaid, *Inmates Sue Marion County Sheriff over Delayed Jail Releases*, FOX 59 (Feb. 15, 2015, 5:31 PM), <https://fox59.com/news/inmates-sue-sheriff-for-delayed-jail-releases/> [<https://perma.cc/39ZX-6MKQ>] (reporting suits arising out of delayed releases); Karen Turner, *Court Officials Blame Software for Wrongful Arrests, Other Legal Mishaps*, WASH. POST (Dec. 20, 2016, 3:56 PM), <https://www.washingtonpost.com/news/the-switch/wp/2016/12/20/court-officials-blame-software-for-wrongful-arrests-other-legal-mishapsu20> [<https://perma.cc/4ALK-WTFY>] (reporting comments from court officials and Tyler Technologies representatives in response to the allegations); Cyrus Farivar, *Inadequate Court Software Still Gets People Wrongly Arrested, Lawyers Say*, ARS TECHNICA (Feb. 1, 2017, 4:00 AM), <https://arstechnica.com/tech-policy/2017/02/inadequate-court-software-still-gets-people-wrongly-arrested-lawyers-say/> [<https://perma.cc/NY7W-LGNU>] (reporting public defenders' critiques of the software).

⁷⁰ Russ Bowen, *Company in Charge of \$85 Million NC Court System Overhaul Faced Problems, Delays in Other States*, CBS17.COM (Feb. 7, 2023, 6:22 PM), <https://www.cbs17.com/news/north-carolina-news/company-in-charge-of-85-million-nc-court-system-overhaul-faced-problems-delays-in-other-states/> [<https://perma.cc/X8UM-DPMJ>].

⁷¹ Michael Hewlett & Jeffrey Billman, *A \$100 Million Dollar Mess: The Rollout of North Carolina's Long-Delayed Attempt to Digitize Court Records Has Been a Catastrophe. Can it Be Salvaged?*, ASSEMBLY (Mar. 29, 2023), <https://www.theassemblync.com/politics/courts/a-100-million-mess> [<https://perma.cc/PES9-EZCL>].

⁷² See *id.* As several commentators have noted, court-technology vendors exploit courts' unsophisticated procurement practices, which do not sufficiently take into account the substantial harms for litigants and the public resulting from technical "glitches" in the system. See NAT'L CTR. FOR STATE CTS., *CONTRACTING DIGITAL SERVICES FOR COURTS 2* (May 2022), https://www.ncsc.org/_data/assets/pdf_file/0029/76754/Contracting-Digital-Services.pdf [<https://perma.cc/J5BY-L3C6>].

Democratic Senator Ron Wyden, would do just that.⁷³ The main objective of the Open Courts Act is to authorize and replace PACER with a new case management system that would facilitate access to individual and bulk court records while incorporating current design practices and open data standards.⁷⁴ Importantly, the code would be available to state and local courts that want to rebuild their systems.⁷⁵ The availability of a well-designed case management system that can be white-labeled and fine-tuned to the needs of particular courts would significantly lower financial barriers to implementation.

But even with an open-source system to build on, bringing state and local court technology infrastructure into the twenty-first century will likely cost hundreds of millions of dollars. The only way such resources might become available is if improving state and local court infrastructure becomes a federal priority. The federal government already has longstanding commitments to improving state and local criminal justice processes,⁷⁶ increasing housing stability,⁷⁷ and protecting consumers from overreaching debt collection practices.⁷⁸ Federal funding to improve state and local court technology infrastructure would further these interests,⁷⁹ along with federal interests in providing resources to victims of crime⁸⁰ and eliminating

⁷³ A draft of this bill is on file with the author and *Northwestern University Law Review*. See Open Courts Act of 2021, S. 2614, 117th Cong. (2021).

⁷⁴ *Id.*; see Nate Raymond, *Making PACER Court Records System Free Wouldn't Add to Deficit, CBO Says*, REUTERS (Dec. 8, 2022), <https://www.reuters.com/legal/government/making-pacer-court-records-system-free-wouldnt-add-deficit-cbo-says-2022-12-08/> [<https://perma.cc/XA76-PLLP>]. The costs of the system would be imposed on frequent users of the system (not including the government). Access for the press and researchers would be free.

⁷⁵ S. 2614, 117th Cong. (2021).

⁷⁶ See Omnibus Crime Control and Safe Streets Act, Pub. L. 90-35, 82 Stat. 197–98 (1968). The Bureau of Justice Assistance at the Department of Justice devotes hundreds of millions of dollars to state and local criminal justice reform efforts. See *About the Bureau of Justice Assistance*, U.S. DEP'T OF JUST., <https://bja.ojp.gov/> [<https://perma.cc/T6M5-MQWH>].

⁷⁷ See Housing and Urban Development Act of 1965, Pub. L. 89-117, 79 Stat. 451; see also Coronavirus Economic Stabilization Act, 15 U.S.C. §§ 9001–141 (2020); American Rescue Plan Act, Pub. L. 117-2, 135 Stat. 4 (2021); Eviction Prevention Act of 2023, H.R. 6696, 118th Cong., 1st Sess. (2023); *Eviction Protection Grant Program*, HUD USER, <https://www.huduser.gov/portal/eviction-protection-grant.html> [<https://perma.cc/9XY4-NZCM>].

⁷⁸ See *Understand How the CFPB's Debt Collection Rule Impacts You*, CFPB (Nov. 30, 2021), <https://www.consumerfinance.gov/about-us/blog/understand-how-cfpb-debt-collection-rule-impacts-you/> [<https://perma.cc/T27R-M9UD>]; see also Fair Debt Collections Practices Act, 15 U.S.C. § 1692 (1977) (authorizing the CFPB's debt collection rule).

⁷⁹ On the federal interests reflected in state and local courts and the need for significant attention to and funding for state courts, see Judith Resnik, “Open” Courts and “Remedy by Due Course of Law”: *The Capital and the Investments in Courts, State and Federal*, 99 N.Y.U. L. REV. (forthcoming 2024).

⁸⁰ See Victims of Crime Act, 34 U.S.C. §§ 20101–20145 (1988).

violence against women.⁸¹ The Bureau of Justice Statistics has begun taking small steps in this direction by funding a new state and local court data collection survey.⁸²

VI. CREATING A COLLABORATIVE RESEARCH ECOSYSTEM

I also believe there are opportunities for collective action among researchers that can accelerate the process of making court data available and bring new researchers into our field. Researchers who want to study eviction, debt, criminal, or other court actions face significant barriers to obtaining high-quality data. A researcher who wants to investigate some aspect of the justice system must first identify a court with usable data, usually by looking at many individual case-search court websites and talking to other researchers in the field. Next, the researcher must either build a scraper, enter a data-sharing agreement with the court, or buy the data from the court (and in the last two instances, the researcher is likely disallowed from sharing the data with other researchers). Then, the researcher must devote time to cleaning the data, and if other researchers have also scraped or obtained the same dataset, the researcher must validate her dataset against theirs. Along the way, the researcher must continuously engage with court personnel and other researchers on how to interpret various fields in the data. All these steps pose an impediment to researching courts.

As researchers, we must push for a much more open court-data ecosystem. We are a small but mighty band; we will become a much bigger and much mightier band if we work to eliminate barriers to accessing data. Such efforts include supporting court-data repositories like SCALES and CJDC so researchers can easily access court datasets to see what questions they might answer. Researchers can contribute to the data ecosystem by ensuring that data-sharing agreements contain provisions for sharing data

⁸¹ See Violence Against Women Act, 34 U.S.C. § 12471 (1994). One approach could be for Congress, in enacting or reauthorizing legislation, to condition federal grants to courts on adherence to common open-data-collection standards. Imposing common data standards and funding improved data collection would be justified as part of the infrastructure required for data collection and reporting under the grant.

⁸² See *Legal Aid Interagency Roundtable*, U.S. DEP'T OF JUST., <https://www.justice.gov/atj/legal-aid-interagency-roundtable> [<https://perma.cc/KAZ5-AMHP>] (demonstrating the federal goal to improve coordination among federal programs and increase availability of meaningful access to justice for individuals and families). Working with the American Bar Foundation's Access to Justice Initiative, the Bureau of Justice Statistics is funding the development and testing of a new large-scale civil legal needs study. See *The ABF Access to Justice Research Initiative and NORC Receive \$2 Million Award from Bureau of Justice Statistics for Access to Justice Design and Testing Program*, AM. BAR FOUND. (Jan. 16, 2024), <https://www.americanbarfoundation.org/abfs-access-to-justice-research-initiative-and-norc-receive-2-million-award-from-bureau-of-justice-statistics-for-access-to-justice-design-and-testing-program/> [<https://perma.cc/A2TM-9JLE>].

with repositories, as Pew Charitable Trusts does in research grants, or, if data are scraped, making it a routine practice to pass them along to repositories for further use.⁸³ We need these repositories to grow the field of researchers studying the justice system. With accessible repositories, we can integrate court data into the data science curriculum, offer courses on court data and the justice system, and create a pipeline of data science graduate students and law students with technical skills and genuine interests in studying civil and criminal justice.

In addition, widely available data will promote broader discussions about variations in data, differences in results, and the interpretations of data fields, leading to easier data harmonization. These processes and discussions must be documented and commonly shared so researchers are not continually duplicating each other's efforts to obtain clean data. Improving access will encourage researchers to investigate increasingly sophisticated questions about the formal justice system.

Making court data widely available will also allow for the replication of results—a cornerstone of empirical research. Replicability is critical to ensuring the reliability and validity of scientific findings. As the ongoing crisis in social psychology makes clear,⁸⁴ replicability helps with identification of errors and biases in original studies and contributes to the robustness of scientific findings.⁸⁵

The last, most fundamental (and obvious) reason to work towards more widely available data is that state and local courts are central public institutions. Courts are where the democratic rubber meets the road—where laws passed by duly elected officials are implemented. In principle, courts should be accountable to the individuals before them and the public. In practice, they are black boxes. We need to nurture the growth of an open court-data ecosystem to evaluate whether courts fairly apply laws as intended by the legislatures and determine whether courts are living up to their promise of being just and equitable institutions.

⁸³ See *How to Standardize Court Data for Greater Transparency and Ongoing Improvement*, PEW (Sept. 27, 2023), <https://www.pewtrusts.org/en/research-and-analysis/fact-sheets/2023/09/how-to-standardize-court-data-for-greater-transparency-and-ongoing-improvement> [https://perma.cc/MZ8X-3FUJ].

⁸⁴ See, e.g., Gideon Lewis-Kraus, *They Studied Dishonesty. Was Their Work a Lie?*, NEW YORKER (Sept. 30, 2023), <https://www.newyorker.com/magazine/2023/10/09/they-studied-dishonesty-was-their-work-a-lie> [https://perma.cc/7H7M-RZYK].

⁸⁵ See OPEN SCI. COLLABORATION, *Estimating the Reproducibility of Psychological Science*, 349 SCIENCE 943, 943 (2015), <https://www.science.org/doi/10.1126/science.aac4716> [https://perma.cc/XP63-MVSU]; see also John P.A. Ioannidis, *Why Most Public Research Findings Are False*, 2 PLOS MED. 696, 700 (2005).

