LOW-INCOME LITIGANTS IN THE SANDBOX: COURT RECORD DATA AND THE LEGAL TECHNOLOGY A2J MARKET

Claire Johnson Raba[*]

Abstract:

With the goal of spurring innovation and development of new legal technology in the direct-to-consumer legal services marketplace, states are considering regulatory changes to the practice of law within legal regulatory sandboxes. In these spaces, creative digital innovators are bringing to market new products to address the gap in legal services delivery that leaves more than 80 percent of low- and moderate-income litigants without legal representation, but there are risks to consumers in rapid adoption of new technology by state courts. In states without adequate access to court record data, sandbox exploration committees, legal technology developers, and researchers operate without a complete understanding of how low-income self-represented litigants fare in state court. This Article addresses the need for quantitative data on the experiences and outcomes of low-income litigants in state courts and calls for reform in access to court record data to ensure that the most vulnerable litigants are helped and not harmed by the adoption of new technology.

This Article presents the novel findings of an empirical study analyzing ten years of court record data for debt collection lawsuits in California and makes the normative recommendation that court record data is a critical starting point for the development of rigorous empirical metrics for evaluating legal technology innovation tools. Based on analysis of California debt collection state civil court record data, the findings in this Article form the basis of an argument for better access to state court record data nationwide and inform the model legislation attached hereto.

---

[*] Assistant Professor of Law, University of Illinois Chicago School of Law. This Article benefitted from early feedback by discussants at the 2020 Law and Society Conference (CRN 43) and from being workshopped at the 2021 AALS Conference Works-in-Progress session. Many thanks for invaluable guidance from Michele Gilman and the Law and Technology group at the 2021 NYU Clinical Law Review Writers' Workshop. Workshopping an article during a pandemic is challenging and much gratitude is also due to the academic conference organizers who ensured that these opportunities remained available remotely. ST. JOHN'S LAW REV., Forthcoming 2023.

Table of Contents

INTRODUCTION

A. Courts Fail to Adequately Serve Consumer Debtors.

Katrina was a community college student with two children, trying to juggle work, childcare, and school.[1] During class in the spring of 2018, her phone buzzed incessantly. She looked down to see a message from her roommate saying a process server had shown up at the house to deliver a summons and complaint, naming Katrina in a lawsuit filed in county court by a debt collection company she had never heard of. Katrina turned to the internet for help and found herself overwhelmed with advertisements that began to pop up in her social media feeds trying to get her to enroll in debt settlement companies, or offering help filing bankruptcy, with or without a lawyer. Katrina didn't know which of these tools to trust, and the court self-help website was overwhelming and full of confusing information that was hard to read on her mobile phone. Katrina is one of the estimated 71 million people in the United States with debt in collections[2] and was one of almost a quarter of a million Californians sued for debt in 2018, almost all of whom would have to navigate a state civil court system as unrepresented litigants against professional debt collection lawyers.[3]

Consumer debt collection cases comprise an increasing percentage of the dockets of most state civil courts in the United States.[4] In California, over the last ten years, debt collection cases totaled 20% of all cases filed, and 34% of the limited civil[5] docket, second only to family law matters.[6] It is estimated that of the 71 million consumers who have debt in collections, 15% were sued

---

[1] Katrina's story is a composite of low-income clients the author worked with while practicing as a legal services attorney.

[2] Urban Institute, *71 Million U.S. Adults Have Debts in Collection* (2018), https://www.urban.org/urban-wire/71-million-us-adults-havedebt-collections; Erica Rickard, *Many U.S. Families Faced Civil Legal Issues in 2018*, https://www.pewtrusts.org/en/research-and-analysis/articles/2019/11/19/many-us-families-faced-civil-legal-issues-in-2018.

[3] Author data, Part I, *infra*, for California debt collection case data (2018 limited civil debt collection filing totals).

[4] Pew Charitable Trusts, *How Debt Collectors Are Transforming the Business of State Courts* (hereinafter "Pew Debt Collection") at p. 6,8 https://www.pewtrusts.org/-/media/assets/2020/05/debt_collectors_to_consumers.pdf; Part I, *infra*.

[5] California limited civil jurisdiction covers cases seeking $25,000 or less and includes all eviction matters.

[6] Author data, Part I, *infra*.

in the last year.[7] That means, according to the research of the Aspen Institute, an estimated 12 million people were sued across the United States to collect a consumer debt in the last year (most commonly credit card, medical debt, auto deficiency, or other consumer unsecured debt).[8] The exact number of people sued on consumer debt cases in state courts each year is not known, because these data points are lost in a myriad of state court case management systems. Researchers and advocates know the exact number of businesses and consumers litigating in federal court, through the unified federal court management system PACER, and a rising number of data analytics companies, from Bloomberg, to Gavelytics, to Lex Machina and Ravel Law, promise law firms and corporations ever-detailed information about judicial behavior and case trends.[9] Also available through the federal PACER system is docket-level information about consumer bankruptcy filings, leading to empirical analysis of legislative changes to the bankruptcy code,[10] but in the area of civil justice, as administered by state courts, there is a "severe data deficit."[11] Recently, states have moved to obtain better criminal case record data[12] in recognition of the necessity for empirical data as a predicate for crafting criminal justice policy, but for many types of civil cases, access-to-

---

[7] Aspen Institute, *Lifting the Weight: Solving the Consumer Debt Crisis for Families, Communities, and Future Generations*, at 44 (2018), http://www.aspenepic.org/wp-content/uploads/2018/12/LiftingtheWeight_SolutionsFramework.pdf [hereinafter *Lifting the Weight*].

[8] *Id.*

[9] UNC Law Blog: *Get to Know your Judge, for a Fee: Judicial Analytics Platforms Promise Insight into Judges' Tendencies*; https://ncjolt.org/blogs/get-know-judge-fee-judicial-analytics-platforms-promise-insight-judges-tendencies/; Colorado Bar Association Colorado Lawyer Magazine: *Here Comes the Judge Judicial Analytics* (Aug./Sept. 2019), https://ncjolt.org/blogs/get-know-judge-fee-judicial-analytics-platforms-promise-insight-judges-tendencies/.

[10] Lois. R. Lupica, *The Consumer Bankruptcy Fee Study: Final Report*, 20 AM. BANKR. INST. L. REV. 17, 107 (2012); Beckett Cantley & Geoffrey Dietrich, *Hindsight: The 2005 Bankruptcy Abuse Prevention &Amp; Consumer Protection Act's Unintended Effects on the Poor* (2020), https://papers.ssrn.com/abstract=3707562; Dalié Jiménez, *The Distribution of Assets in Consumer Chapter 7 Bankruptcy Case*s, 83 AM. BANK. LAW J. 795 (2009).

[11] Rebecca L. Sandefur, *Paying down the Civil Justice Data Deficit: Leveraging Existing National Data Collection*, 68 S. C. L. REV. 295–310 (2016).

[12] California AB 1331, passed in 2019, requires criminal court record data to be compiled and reported to the state by criminal justice agencies and clarifies that courts can share data with research organizations in order to increase access and accuracy of research data in criminal cases. AB 1331 amends California Penal Code sections 13150, 13152, and 13202. Measures for Justice is also working in states across the United States to improve criminal justice record data. https://www.measuresforjustice.org/portal.

justice scholars and other vital stakeholders do not know what is happening in many state courts, particularly in states with disaggregated case management systems.[13]

Knowing how legal issues are resolved, and the consequences of legal issues for the people who experience them, has been identified as critical to "knowing who has access to justice and how healthy is the rule of law," yet states do not collect basic information about debt collection cases in a form that can be analyzed and studied.[14] In most states, researchers, policymakers, and technology innovators have no idea how many people were sued to collect a consumer debt, where those lawsuits are happening, how they resolved, or what post-judgment collection activity occurred, and researchers have to go to extraordinary lengths to try to obtain this information.[15] Without robust data, "the result is a lot of policymaking in the dark."[16] This is particularly true of cases that affect poor people and self-represented litigants, such as eviction, debt collection, and small claims matters.[17]

B. States Restrict Access to State Court Record Data.

Pew Charitable Trusts reporting on debt collection case in 12 states found that debt claims filed doubled in the twenty years between 1993 and

---

[13] The data that does exist is often collected and intended for use in administrative court functions, rather than with a focus on legal needs and improving access to justice. Hugh McDonald, *Assessing Access to Justice: How Much "Legal" Do People Need and How Can We Know?*, 11 U.C. IRVINE L. REV. 693, 736 (2021).

[14] Rebecca L. Sandefur, *Paying down the Civil Justice Data Deficit: Leveraging Existing National Data Collection*, 68 S. C. L. REV. at 300 (2016); see also McDonald, *supra* note 13, at 728 ("public legal assistance and judicial case management would, ideally, be based on what works best for different users," but a major access-to-justice challenge is posed by a lack of information and data.)

[15] See, e.g., the methodology used in California by the Center for Responsible Lending, which visited individual courts in person to obtain a representative sample of court record data and extrapolated to estimate the number of filings. Center for Responsible Lending, Court System Overload: *The State of Debt Collection in California after the Fair Debt Buyer Protection Act* (2019), https://www.responsiblelending.org/research-publication/court-system-overload-state-debt-collection-california-after-fair-debt-buyer.

[16] San Francisco District Attorney George Gascón and Amy Bach, *Open Forum: Better Criminal Justice Requires Better Information*, S.F. CHRON., Sept. 19 2019. https://www.sfchronicle.com/opinion/openforum/article/Open-Forum-Better-criminal-justice-requires-14448193.php.

[17] In order to create the first database of eviction cases in the United States, the Princeton Eviction Lab has used an exhaustive methodology to first request court data through public records request, supplemented by data scraping, and then by purchasing data from third party vendors to collect data by county. No such database exists for debt collection cases. https://evictionlab.org/methods/#what-data.

2013 and that in 2013, twenty-four percent of all cases filed were to collect a debt, exceeding the number of other civil, family law, and juvenile case combined.[18] However, Pew found that only 7 states track debt collection matters at the state level with sufficient granularity statewide in order to meaningfully track trends over time, and only Texas and Colorado track debt cases as a category in publicly available reports.[19] Only two states report disposition data, including default judgment rates.[20]

An initial data point for evaluating the scope of the access-to-justice problem in each state, and in each county, is the number of cases in each civil case type where litigants are unrepresented by an attorney. The Legal Services Corporation (LSC), the non-profit corporation that the federal government partners with to facilitate the funding of civil legal aid throughout the United States, is also seeking data to identify legal needs and gaps in services. LSC is engaged in data scraping (obtaining data through the front end access portals) in states that allow such automated acquisition of case record data, but the LSC Office of Data Governance and Analysis reports that only one county, Tarrant County, Texas, affirmatively and intentionally tracked self-represented litigants, and this county has stopped doing so.[21]

In states where more granular docket-level data is accessible through portal scraping or through other means, data analysts normalize the data across cases in which the attorney field is blank through the case to resolution to create a field called "self-represented litigant."[22] This type of analysis for a set of 12,499 private student loan debt collection cases in California found that between ninety and one hundred percent of defendants in debt collection matters were unrepresented, depending on the county.[23] This may be due to

---

[18] Pew Charitable Trusts, *How Debt Collectors Are Transforming the Business of State Courts* (2020), at 6, 8 https://www.pewtrusts.org/-/media/assets/2020/05/debt_collectors_to_consumers.pdf.

[19] Pew Charitable Trusts, *supra* note 18, at 19. A failure to standardize categories can lead to a failure not only to track data in the first place, but also may result in errors in the data. For example, medical data led to an undercount in opioid deaths due to a lack of standardization of categories. Meryl Kornfield, *Researchers find that more people died from opioid deaths than reported - because of lack of standardization of categories*, WASH. POST, February 28, 2020; Christine E. Cerniglia, *The Civil Self-Representation Crisis: The Need for More Data and Less Complacency*, 27 GEO. J. POVERTY LAW & POL'Y 355, 370 (2020).

[20] Pew Charitable Trusts, *supra* note 18, at 19.

[21] Email correspondence with Carlos Manjarrez, Chief Data Officer, LSC Office of Data Governance and Analysis (February 7, 2020).

[22] *Id.*

[23] Claire Johnson Raba, *Co-Opting California Courts: How Private Creditors Have Turned*

the disproportionate cost of retaining an attorney compared to the cost of the debt, it makes little economic sense to hire a private attorney to defend many debt collection cases, and in California as in many states, legal aid and self-help assistance is scarce for consumer debt lawsuits.[24] When data is analyzed, it supports the proposition that the overwhelming result in debt collection lawsuits is a default judgment entered against the consumer defendant.[25]

The research conducted for this paper demonstrates that like many states, California local county courts do track much of this data, but do not report it or aggregate it in a form that is usable for researchers, legal technology app developers, and state access-to-justice commissions. This Article argues that if we are to address the flood of default judgments against unrepresented consumers, it is imperative that state judicial bodies and legislatures require local courts make available docket-level data in matters like consumer debt collection cases.

The lack of data about self-represented litigants in state courts has accurately been described as a crisis.[26] In California, as in many other states, courts place a high administrative cost on mass access to civil court record docket-level data, and this data is not organized in a normalized way that allows for quantitative evaluation of interventions for self-represented litigants. This lack of access to court record data places corporate entities that can purchase civil court record data at an advantage over non-profits, researchers, and state regulatory bodies themselves, and boxes out independent researchers from evaluating the impact of market-based legal technology tools.[27] Because this data is not aggregated and available to the

---

the Judiciary into a Predatory Student Debt Collection Machine at 19 (2021), https://protectborrowers.org/co-opting-california-courts-how-private-creditors-have-turned-the-judiciary-into-a-predatory-student-debt-collection-machine/.

[24] LawHelp California shows only five county court self-help programs that assist with debt collection matters, out of 58 California counties. https://www.lawhelpca.org/legal-directory?field_subtopic_reference_target_id=307&field_topic_target_id=3.

[25] Emily S. Taylor Poppe, *Why Consumer Defendants Lump It*, 14 NW. J. L. & SOC. POL'Y. 149, 152 (2019); Johnson Raba, *supra* note 23 at 21.

[26] Cerniglia, *supra* note 19 at 370 (noting that only three states report reliable data on self-represented litigants to the National Center for State Courts).

[27] California's disaggregated system can be contrasted with the unified court case management system in Utah. The Utah regulatory sandbox project has established a partnership with University of Denver's Institute for the Advancement of the Legal System (IAALS). IAALS and other researchers in Utah are able to effectively engage in empirical research on Utah's access-to-justice reforms in part due to access to the full set of civil court record data from Utah's unified court case management system. Through its Unlocking Legal Regulation Project, IAALS is a third-party researcher evaluating Utah's regulatory sandbox.

public, it may also be that upon implementation of California's regulatory sandbox, legal aid and non-profit entities may be excluded or discouraged from participation in legal services regulatory reform efforts because of an inability to determine the scope of the problem to be solved.

C. Lack of Data Imperils Unrepresented Litigants.

A stated goal of the movement to explore regulatory sandboxes and spur innovative technology solutions is to close the justice gap.[28] In some states, like Utah, which has partnered with the University of Denver's Institute for the Advancement of the American Legal System for evaluation of the Utah regulatory sandbox, these regulatory changes are accompanied by empirical research.[29] However, the movement toward rapid adoption of market-based innovations glosses over the current state of state court record data, particularly in states where there is not a unified court system. For example, The California ATILS commission, which led to the creation of California's Closing the Justice Gap Working Group, sets the following lofty goal: "In connection with the goal of increasing access to justice, this statement represents ATILS' strong interest in a deliberate effort to identify and evaluate metrics that can assess the actual impact of the implementation of the ATILS regulatory reform options on consumer access to legal services, including but not limited to, the justice gap."[30] This Article proposes that state courts and legislatures need to take an affirmative role in the collection and aggregation of civil court record data in order to facilitate an effective implementation of access to justice technologies in a manner that produces statistically significant positive outcomes for self-represented litigant users as well as for the courts and with an assurance that they will not cause harm.

A body of research on self-represented litigants (SRLs) as plaintiffs in consumer law matters that utilizes federal court record data is centered on enforcement of rights as plaintiffs coming to the court to seek redress.[31] Very

---

IAALS, *Unlocking Legal Regulation* https://iaals.du.edu/projects/unlocking-legal-regulation.

[28] *Id. See also* California Closing the Justice Gap Working Group, https://www.calbar.ca.gov/About-Us/Who-We-Are/Committees/Closing-the-Justice-Gap-Working-Group

[29] *Id.*

[30] ATILS Memo for Public Comment (2019), http://www.calbar.ca.gov/Portals/0/documents/public/Comment/2019/List-of-Tentative-Recommendations-Memo-For-Public-Comment.pdf.

[31] Mark D. Gough & Emily S. Taylor Poppe, *(Un)Changing Rates of Pro Se Litigation in*

little research has examined the psychology of the role of the SRL as a defendant in a civil case.[32] Consumer debt defendants are different, even from other defendants in state court, including those facing an eviction or family law matter, where the litigant realizes that there is a legal issue to be adjudicated, and the stakes – loss of housing or custody of a child – are high.[33] Consumer debtors are not, like criminal defendants or recipients of traffic tickets, hailed into court to be adjudged for their wrongful acts, but they subject to a similar moral evaluation.[34] Consumers are often made to feel that they owe the money and they must pay, a moral and ethical trap that creditor attorneys have exploited to their advantage, bringing along the court system as their accessories.[35] When evaluating legal technology products, and particularly online dispute resolution,[36] researchers need empirical data on what happens in to consumer debt defendants before judgment is entered, and in post-judgment collection, in order to ensure that the rights of consumers are not being lost in the implementation of legal technology processes that run the risk of allowing creditor plaintiffs to use technology to more efficiently automate the process of obtaining default judgments against consumers.

Federal Court, 45 LAW & SOCIAL INQUIRY 567–589 (2020) (using federal court docket data to engage in empirical study of pro se litigants in federal court): Anna E Carpenter et al., *Studying the "New" Civil Judges*, 2018 WIS. L. REV. 249 at 268 (a gap in knowledge about state courts, in part due to a federal court bias in scholarship and research)

[32] Taylor Poppe, *supra* note 25, at 152.

[33] Matthew Desmond et al., *Evicting Children*, 92 SOCIAL FORCES 303–327 (2013) (using court record data to analyze case outcomes in eviction matters); Ellen Degnan et al., *Trapped in Marriage* (2018), https://papers.ssrn.com/abstract=3277900 (a randomized control study examining the effect of pro bono counsel for self-represented litigants in family law cases).

[34] Consumer defendants are more likely than other civil litigants to not respond to service of legal papers, in what Emily Taylor Poppe calls "lumping it," resulting in an entry of default judgment. Taylor Poppe, *supra* note 25, at 152; Rebecca L. Sandefur, *What We Know and Need to Know about the Legal Needs of the Public*, 67 S.C. L. REV. 443–460 at 448 (2015) (recognizing that many people do not "think of [justice] problems as legal . . . the legal nature of any given civil justice problem is socially constructed").

[35] David J. Griener et al., *Self-Help, Reimagined*, 92 IND. L.J. 1119 (2017).

[36] Online dispute resolution (ODR) and other forms of technology-assisted alternative dispute resolution methods require a critical lens in debt collection matters due to the imbalance of power between the parties. Settlement does not look the same from the perspective of the creditor as the debtor, and there is a risk of losing that lens when ODR is viewed as an efficiency solution for courts. See Owen Fiss, *Against Settlement*, 93 YALE L. J. 1073, 1075 (1984) Settlement should be treated "as a highly problematic technique for streamlining dockets . . . and although dockets are trimmed, justice may not be done."

There are multiple places in the process of a debt collection lawsuit in which legal technology tools can help or automate processes in order to obtain more buy-in and participation from self-help centers and legal aid organizations,[37] but those seeking to help solve this problem need data from the courts. State legislation, such as that proposed in Appendix A, requiring courts to provide comprehensive, granular, and sufficient data, will ensure that tools that are developed to help self-represented consumer plaintiffs obtain outcomes that in the long term are measurably good for the consumer, not merely an improvement on judicial efficiency.

In Part I, I report on the results of attempts to obtain debt collection civil record data through public records act requests, data portal scraping, and the existing California judicial branch reporting system and discuss findings in the exploratory analysis of this normalized set of data.

In Part II, I discuss the importance of court record data access for empirical study of the courts and legal technology innovations in context with the access-to-justice movement and the legal empiricism of New Legal Realism scholarship.

In Part III, I discuss norms for evaluating legal technology tools for positive, neutral, and negative outcomes and discuss the granularity of aggregated data necessary to make empirical evaluations based on these norms.

In Part IV, I discuss the lack of data available for evaluating changes to the civil justice system in state courts and the lack of aggregated state court data in the United States and in California, including some innovative and diverse efforts currently underway to attempt to access, aggregate, taxonomize, and normalize this data. I then propose model legislation to promote transparency and access to civil court records.

I. CALIFORNIA DEBT COLLECTION CASE RECORD DATA

A. A Blizzard of Administrative Records Requests.

California's respect for access to public records is well-defined in state law. The California Public Records Act, codified in 1968, was modeled

[37] In 2010, 4 million people visited the California court self-help website. Bonnie Hough, *Self-Represented Litigants in Family Law: The Response of California's Courts,* CAL. L. REV. CIRCUIT, (Feb. 10, 2010).

after the federal Freedom of Information Act of 1967.[38] The Ninth Circuit has recently held that data belonging to private companies but publicly available on the internet is also subject to a right of public access, in holding that public LinkedIn data was not protected from data scraping.[39] Docket-level court data is both public record data and information that is published on the internet, similar to the LinkedIn public profiles; accordingly, the recent decision should make it easier for private companies engaged in data scraping through the front-end portals of the 58 counties in California. However, data scraping through a portal is limited to the data produced on the public facing docket, and cost-prohibitive access fees for documents actually filed in a court case creates a barrier to access to information. In a reflection of how a market-driven approach to need can skew the manner in which data is produced, docket scraping by private companies often leads to cleaned and normalized data available for purchase that is optimized for analytics for a particular customer base and is not made widely available to the public for research. [40] Courts collect significantly more information than they choose to list on the docket, including disposition data, but this information is not uniformly reported in any field on the public docket.[41]

To engage in statewide civil court record data research specific to debt collection cases, I chose to ask the 58 California county courts for a variety of administrative record data, in part to determine what information each court is tracking. California permits the public to request copies of administrative records through California Rule of Court 10.500 which became effective January 1, 2010. [42] The rule allows public access to administrative information about the operation of the courts, and is differentiated from access to records filed in court cases. Individual case

---

[38] Scott A. Baxter, *Information Privacy and the California Public Records Act Public Entities Officers and Employees*, 30 MCGEORGE L. REV. 778–801, 785 (1998); Cal. Gov. Code § 6250 et seq.

[39] h*iQ Labs, Inc. v. LinkedIn Corp*., 938 F.3d 985, (9th Cir. 2019); Marissa Boulanger, *Scraping the Bottom of the Barrel: Why It Is No Surprise That Data Scrapers Can Have Access to Public Profiles on LinkedIn*, 21 S.M.U. SCI. & TECH L. REV. 77 (2018).

[40] As examples, RavelLaw, Gavelytics, UniCourt, and other court data analytics companies have profit models that rely on subscribers who are primarily law firms interested in outcomes of business litigation, driving a focus on party, judicial, and law firm analytics, and less analysis of granular disposition data and post-judgment docket activity.

[41] Part I, *infra*, data on file with author.

[42] Authorized by California Government Code Section 68106.2(g), which requires adoption of rules of court that provide public access to non-adjudicative court records.

records are also available to the public but courts are allowed to charge a fee for electronic access to individual documents filed in cases, and this cost varies by county with some counties charging $7.50 per page. I set out to determine how the 58 county courts in California would interpret a request for aggregate information about the filing of a particular type of court case and the case outcomes as administrative records, and to obtain these records. The intent of the Judicial Council in promulgating this rule "clarifies and expands the public's right to access to judicial administrative records and must be broadly construed to further the public's right of access."[43]

Methodology: I sent letters requesting the following aggregated information from each of the 58 California counties for limited civil debt collection cases, which are cases filed to collect less than $25,000 and designated as a California Rule of Court 3.470[44] limited civil debt collection matter on the mandatory civil case cover sheet:

- The number of debt collection cases filed each year from 2010 through 2019.
- The number of unrepresented or self-represented parties in debt collection cases each year from 2010 through 2019.
- Disposition data for debt collection cases resolved each year from 2010 through 2019 broken down by :
    - Prevailing party;
    - Outcome (dismissal, with or without prejudice or coded as settled and dismissed), default judgment, judgment after trial, pre-trial dispositive motions (motions for judgment on the pleadings and for summary judgment), and stipulated entry of judgment;
- The number of motions to compel filed in limited civil debt collection cases each year;
- The number of self-represented litigants in limited civil debt collection cases by year;
- The number of debt collection case in which litigants were assisted by the court's self-help center, a legal aid program, or a law school clinic;

[43] CRC 10.500(a)(2)
[44] Cal. Rule of Court 3.470, https://www.courts.ca.gov/cms/rules/index.cfm ?title=three&linkid=rule3_740

- The number of cases each year in which a legal technology tool was used by a litigant.

Research Questions:

1) The first research question asked what data courts would return. The responses created baseline of data for the counties that returned results, either in lists of cases or in aggregate totals to determine what percentage of dockets were debt collection case in each county. The value in these findings show how accessible this data is, and the level of detail available for researchers.

2) The second research question asked what data courts are currently tracking. This relied on asking the court for categories of information with the hypothesis that courts track a wide variety of data points and that the results would contain a spectrum of responses, from fully responsive data sets to responses stating that the court was in possession of no responsive information.

3) The third research question asked how many debt collection cases were filed annually per county and the disposition of those cases, if known.

Upon receipt of responses to administrative records requests, I combined aggregate and raw data responses with information reported to the Judicial Council of California through the Judicial Branch Statistical Information System,[45] and with data scraped from the front-end portals of 16 of the largest county courts in California, using some records from a commercial data vendor.

I aggregated the court record data received and engaged in data normalization and cleaning of the individual and aggregate data using the R programming language. Where I received full data sets for some counties and aggregate data for others for the same data fields, I created a combined data set of aggregate data to compile state-wide data. I used Tableau for data visualization of the findings.

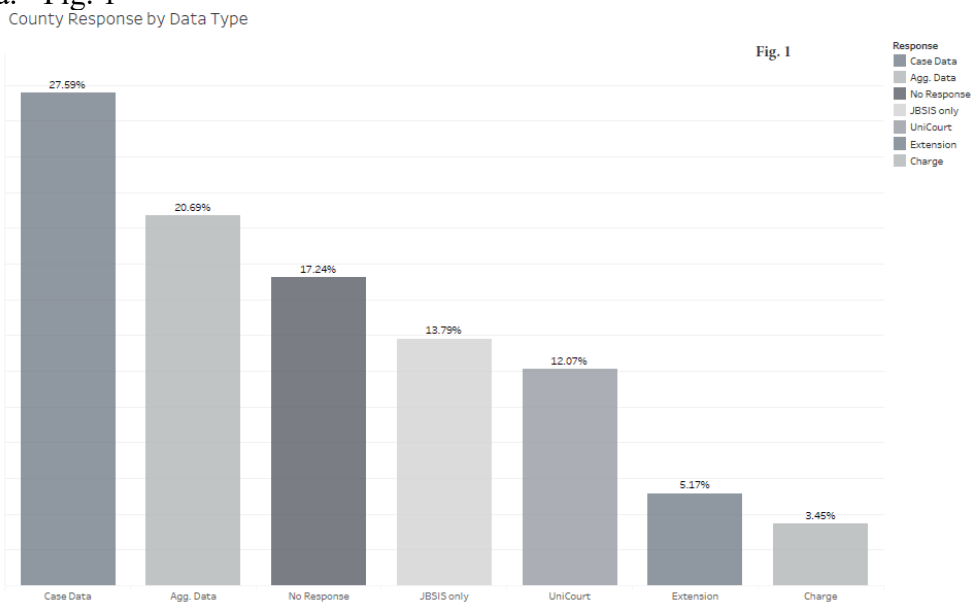
[45] The report used in the data analysis in this Article was the 2019 Court Statistics Report Statewide Caseload Trends 2008-09 through 2017-18, https://www.courts.ca.gov/documents/2019-Court-Statistics-Report.pdf

1. Aggregated and Normalized Debt Collection Data.

Of the 58 counties in California, 48 provided a response as required by California law. 12 counties responded with aggregate data and 15 responded with in-depth court management system output, in the form of .csv and Excel tables with fields responding to the requests for information. Four county courts asked for an extension, and by the time the Covid-19 response shut down the courts, had not yet responded. Eight courts directed me to the Judicial Branch Statistical Information System or otherwise responded that they do not track limited civil debt collection case information. Seven courts that were otherwise nonresponsive had data available through UniCourt, a commercial data scraping service, including fields that allowed for the determination of the number of limited civil debt collection filings. A summary of data type responses is shown in Figure 1.

Unsurprisingly, the data returned contained no aggregate information on current use of legal technology tools. In response to a request for data on self-help center involvement or legal aid assistance with debt collection cases, only two county courts track the involvement of the court's own self-help center anywhere in their administrative court records provided. From the record response data, it does not appear in any county that self-help center tracking involvement of court staff in debt cases is integrated with court case management systems.

a. Fig. 1



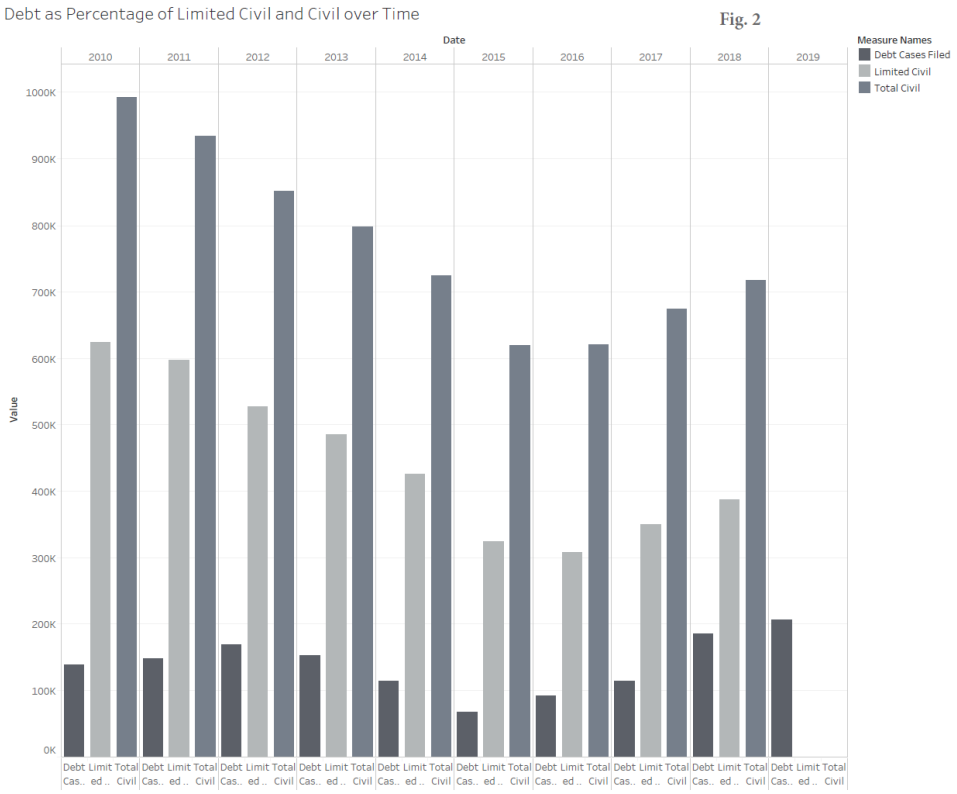
In Figure 1, Case Data refers to a full export of comma delimited file of case file docket data, and Aggregate Data refers to a compiled list of total values for annual filings. Although both types of data were responsive to the Administrative Records Request, it was evident that for a thorough and effective analysis of the debt collection case records that the full Case Data was significantly easier to merge across counties for a full picture of debt collection cases in the state. Through a compilation of the Case Data and Aggregate Data returned data sets, I was able to review detailed filing data information for 34 of the 58 counties in California, including a large number of full sets of data of the smaller counties, which is information not available to commercial analytics firms through online portal scraping. Smaller counties may have fewer filings and be of less interest to commercial analytics firms, but the comparison of data from urban, suburban, and rural communities is a key factor to designing legal technology tools that address the needs of diverse populations.

Across almost all counties, limited civil debt collection cases comprise between 10 and 25 percent of the total cases filed, and frequently were close to half of the limited civil filings. As the remainder of limited civil filings are generally unlawful detainer (eviction) matters, there are interesting data points to be investigated by county during times of economic downturn and recovery. Figure 2 shows a decrease in the number of civil filings total

in California over the study period, with 2015 and 2016 showing the fewest filings and a steady increase again through 2018, the last year that full data was available during the study period.[46] Figure 2 shows that the central value measured, limited civil cases, show a slower increase after the dramatic decrease in filings in 2014-2015, indicating an area that needs further study, and an area of inquiry that would benefit from a comparison with data on non-debt collection cases filed in limited civil matters, the majority of which are eviction cases. Figure 2 shows that debt collection cases constitute an increasing percentage of civil filings, and of limited civil filings, over the most recent years of the study period, increasing from 113,697 cases filed in 2017, which was 17 percent of civil filings and 33 percent of limited civil filings, to 184,589, which was 26 percent of total civil and 48 percent of limited civil matters.

[46] Data analysis used the Judicial Council of California 2019 Court Statistics Report Statewide Caseload Trends report JBSIS 2019, *supra* note 45.

b. Fig. 2



Responsive data was received from 58 percent of counties, including from the 16 most populous counties in the state of California. The dataset contains records for debt collection cases in 87 percent of the court cases filed in California (total case filings in counties reporting data/total case count for the state) over ten years. Over the full data set studied over ten years of filings, debt collection cases comprised 20 percent of total case filings and 24 percent of limited civil debt collection filings.

c. Fig. 3

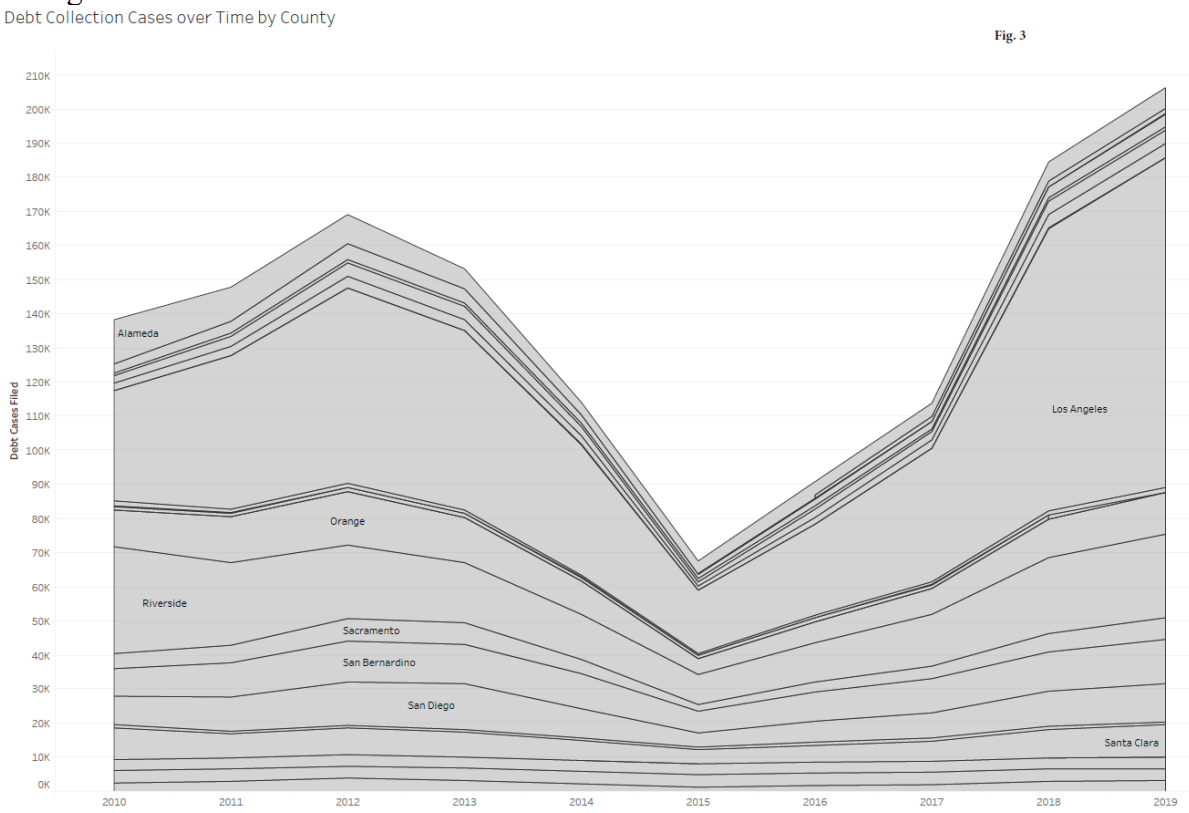


Figure 3 shows the distribution of debt collection filings by year, split by county. Filings throughout the state ranged from a low of only 67,473 cases filed statewide in 2015, to a high of 206,208 in 2019. In Los Angeles alone, the full study period yielded a totaled 488,771 cases, with annual distributions showing the same drop in 2015 and 2016 as across the state, but then skyrocketing to 82,688 filings in 2018 and 96,508 in 2019. Los Angeles County is an area for additional study, as it was one of the counties that provided no data in response to the request for administrative records and responded that it only collected data as necessary to report to the Judicial Council for the JBSIS report that does not itemize by a debt collection case category.[47] Debt collection filings rose during the financial crisis and then dropped precipitously in 2014 and 2015 when two regulatory events occurred: the Consumer Financial Protection Bureau brought enforcement

[47] Los Angeles data was extracted from aggregate data on annual filing totals provided by third-party data vendor UniCourt.

actions against the two largest debt buyers in the United States[48] and California passed the Fair Debt Buying Practices Act.[49] However, in almost every county in which data was reported, the number of debt collection cases filed in 2018 and 2019 are as high or higher than in the worst days of the recession. These findings provide an area for future study on the impact of regulatory and legislative approaches to consumer protection against unfair and deceptive acts and practices by debt collectors. These data points present an opportunity for study on the demographic and geographic distribution of debt collection cases to map why, where, and how the economic recovery did not trickle down to the part of the of the population of California who is sued on defaulted debt. Additional study, and a more robust data set of court record data, is needed to make additional findings about geographic and demographic differences for California populations and consumer credit and debt.

The most valuable data point that is not being measured with accuracy by courts the number of cases in which a debt collection defendant is self-represented. Some courts, such as Orange County, did return data on the percentage of self-represented litigants in debt collection cases, affirming that in one of the largest counties in California, fewer than five percent of debtors are represented by an attorney. This data may be extracted from the full court record data by extrapolating a value from empty attorney fields, but the courts do not generally track this information as a metric, and those that responded with aggregate data asserted that they did not have this information. Figure 4 shows the breakdown of representation of parties in debt collection by year in Orange County.

---

[48] *CFPB v Encore Capital Group Inc, Midland Funding LLC, Midland Credit Management Inc. And Asset Acceptance Capital Corp*. File No. 2015-CFPB-0022 (2015), https://files.consumerfinance.gov/f/201509_cfpb_consent-order-encore-capital-group.pdf.

[49] Cal. Civ. Code § 1788.61 et seq. (imposes evidentiary requirements for a default judgment in consumer debt cases brought by third party debt buyers)

d. Fig.4

Fig. 4

Orange County Representation by Counsel by Year

	Representation		
Year of Filing Date	plaintiff and defendant pro per	plaintiff and defendant rep	plaintiff rep only
2010	0.31%	2.67%	97.03%
2011	0.40%	3.09%	96.52%
2012	0.62%	4.03%	95.35%
2013	0.38%	3.71%	95.92%
2014	0.37%	3.73%	95.90%
2015	0.42%	4.79%	94.78%
2016	0.46%	4.04%	95.50%
2017	0.41%	3.74%	95.85%
2018	0.53%	3.11%	96.36%
2019	0.23%	3.46%	96.31%

Orange County also tracks disposition data with a degree of specificity and particularity that is unusual for California courts. Upon an initial review of a combined set of data from the courts that returned full data sets, analysis showed an unusually high number of cases with a disposition value of "not yet classified by court" or "pending." A case lookup for a random sampling of these case numbers showed that the cases had actually been disposed and that many courts are not updating the case disposition field at the resolution of a case. Orange County data, supports findings by researchers in other states and nationally that find an overwhelming number of debt collection cases result in a default judgment. Figure 5 shows the outcomes of cases in Orange County, and shows about 45 percent of cases result in a default judgment, entered by either a court clerk or a judge, and only one percent of cases result in a court finding, an adjudication on the merits.[50]

[50] Total number of filings in Orange County over the study period equal 194,592 cases. Totals in Figure 5 do not add up to 100 percent because values with fewer than 500 records were excluded, as were pending cases.

e. Fig. 5

Orange County Disposition Distribution 2010-2019 Fig. 5

Request for Dismissal	70,863
Default judgment by clerk	62,133
Default judgment by court	23,881
Court Dismissal	3,534
Stipulated judgment	2,868
Court finding	2,283
Summary judgment	993
Stipulated settlement per..	634

The model legislation and court administrative rules proposed in Appendix A seek to address these issues by including requirements for courts to collect data on these values. By requiring adequate collection and reporting of data, legislatures ensure that researchers, technology developers, regulatory sandbox committees, and administrative offices of state courts can looks meaningfully at the experiences and case outcomes for the vast numbers of self-represented litigants in the state court system.

2. A Gap in Outcome and Disposition Data.

The research conducted herein shows that state courts, at the county level, are capable of providing valuable raw output data containing fields that can be normalized. The work done in front-end portal scraping, by private companies, by the LSC, and by Pew, also shows this. However, courts need guidance to understand which fields they should be tracking because some of the data is simply not present in the docket information, in part due to a state court administrative office reporting procedure that emphasizes active case management and efficacy of processing dockets through the courthouse,

rather than identifying how the court processes are and are not working for litigants, particularly those unrepresented by counsel.[51]

A gap in the data exists for normalized disposition data. While some courts provided information about disposition of debt collection cases, the largest counties referred me to the information collected and published by the Judicial Council's Judicial Branch Statistical Information System, which does not report any case data at a level granular enough to analyze debt collection lawsuits. Moreover, the JBSIS data tends to focus on time to disposition, rather than on substantive outcomes. As discussed *infra,* access to justice for self-represented litigants requires a deep understanding of the outcomes of cases to measure impact of interventions through legal technology and innovative provision of legal services.

Legislative changes, like those proposed in Part IV *infra* and Appendix A, can take the form of rewriting the contracts that local courts have with third-party vendors. Courts can pivot and begin tracking data which will provide a meaningful baseline against which to measure access to justice technology solutions. Although regulatory and legislative changes are more likely to be well received if they are forward looking and do not tackle the issues of the legacy data, such guidance can require reporting out of legacy data in the data set. Given the work already demonstrated in data normalization across large and diverse data sets, and the data normalization done in the data set in this research, it is not unreasonable to expect to obtain a usable set of millions of records, if the courts are required to so report.

The results of the administrative records act requests, data scraping, and analysis of JBSIS data demonstrates the issues with a disaggregated state court system, but the number of county courts that responded with usable data also shows that this is an area of data collection that is ripe for reform. The data produced for this Article, along with the data in the Pew Debt Collection study, demonstrates the need throughout the United States for legislative and regulatory change to obtain usable and meaningful data for case in which most litigants are self-represented. Further discussion of proposed legislation is in Part IV of this Article and the model legislation attached in Appendix A.

---

[51] The National Center for State Courts has created a Data Governance Policy Guide with some recommendations for data standards. National Center for State Courts, DATA GOVERNANCE POLICY GUIDE (2019) at 12-13, https://www.courtstatistics.org/__data/assets/pdf_file/0014/23900/data-governance-final.pdf

II. NEW LEGAL REALISM AND DATA ACCESS

A. Legal Empiricism and Quantitative Data.

Lack of access to state civil court record data creates a barrier to comprehensive empirical research on how the vast number of unrepresented litigants in state court experience the legal system, leaving legal empiricists without a baseline of data against which to evaluate interventions.[52] Research on access to justice as a body of empirical study was recognized in 2013 by Sandefur and Albiston as being at a "theoretical crossroads" amid the emergence of a growing number of researchers engaging in the study of the effectiveness of current civil legal services delivery.[53] Over the last eight years, the study of access to justice has expanded beyond the concept of access to an attorney to include the study of new and innovative ideas to address the thousands of individuals whose needs go unmet by traditional legal aid offices.[54] Beyond civil legal services delivery and the self-reflection of the Legal Services Corporation's Justice Gap Study,[55] an effective empirical study of how individuals interact with the courts must also study the systems and structures, the mechanisms by which civil legal services are delivered.[56] Over the last eight years, scholars have answered this call, incorporating into studies of courts and how individuals experience the civil legal system empirical research design and data-driven normative recommendations.[57]

---

[52] Cernigla, supra note 19 at 370; Rebecca L. Sandefur, *Paying down the Civil Justice Data Deficit: Leveraging Existing National Data Collection*, 68 S.C. L. REV. 295–310 (2016);

[53] Catherine R. Albiston & Rebecca L. Sandefur, *Expanding the Empirical Study of Access to Justice Colloquium*, 2013 WIS. L. REV. 101–120, 107 (2013).

[54] Rebecca L. Sandefur, *Access to What*?, 148 DAEDALUS 49–55 (2019); Rebecca L. Sandefur, *What We Know and Need to Know about the Legal Needs of the Public White Papers*, 67 S. C. L. REV. 443–460 (2015); Anna E Carpenter et al., *Studying the "New" Civil Judges*, WIS. L. REV. 39; Erin York Cornwell, Emily S. Taylor Poppe & Megan Doherty Bea, *Networking in the Shadow of the Law: Informal Access to Legal Expertise through Personal Network Ties*, 51 LAW & SOC. REV. 635–668 (2017).

[55] The methodology behind the LSC annual Justice Gap Study involves asking civil legal aid offices to self-report the number of cases turned away. Sandefur and Albiston's research instead asks people about their civil legal needs, whether or not they approach a legal aid office for assistance.

[56] Sandefur and Albiston, *supra* note 53 at 107, encompassing a broad scope of research in access to justice, "considering not only individuals, but also institutions, not only resources, but also social meaning, not only how civil legal services are provided, but how demand for those services is shaped, to name just a few issues."

[57] See, e.g., Dalie Jimenez et al., *Improving the Lives of Individuals in Financial Distress Using a Randomized Control Trial: A Research and Clinical Approach*, 20 GEO. J. ON

This data-driven approach to studying "law in action" was brought under the framework of "New Legal Realism" in 2005 by researchers who take an interdisciplinary and empirical approach to legal scholarship.[58] Legal scholars within the "big tent" of New Legal Realism have expanded beyond the early New Legal Realist empirical study of judicial behavior and courts to incorporate qualitative and quantitative research design in analysis of how law and society intersect, giving "concerted thought to the problem of translation between the legal academy and empirical work on the law."[59] New Legal Realism as applied to the study of inequality in the law often takes a multidimensional approach, making normative recommendations that not only addressing policy, but also the political implications of proposed legislative or regulatory change.[60]

The emergence of legal technology and states' adoption of new regulatory frameworks for how legal services are provided are squarely within the recognition by the original legal realists that "legal officials do not always enforce rules as written, and that social behavior regularly diverges from legal dictates."[61] Modern legal empiricism as applied to the study of

---

POVERTY L. & POL'Y 449–478 (2012) (utilizing a randomized control trial studying two interventions for consumer debt defendants); Tonya L. Brito, *The Child Support Debt Bubble*, 9 UC IRVINE L. REV. 953–988 (2018) (using qualitative data from ethnography and interviews to make normative recommendations regarding challenging child support arrears); Sudeall, et al., *Paradox: Inside the Black Box of Eviction Court* (March 2, 2021). Forthcoming VANDERBILT L. REV., 2021, https://ssrn.com/abstract=3796279 (an approach to studying the divergent legal processes and outcomes in eviction court using quantitative analysis of court record data and qualitative data from interviews).

[58] Howard Erlanger et al., *Is It Time for a New Legal Realism New Legal Realism Symposium: Is It Time for a New Legal Realism: Foreword*, 2005 WIS. L. REV. 335–364 (2005) at 341, discussing the interdisciplinary paradigm in new legal realism.

[59] Elizabeth Mertz, *New Legal Realism: Law and Social Science in the New Millennium, in The New Legal Realism*: TRANSLATING LAW-AND-SOCIETY FOR TODAY'S LEGAL PRACTICE, ed. Mertz, Stewart Macaulay, and Thomas W. Mitchell (2016), at 5 (comparing early Miles and Sunstein work on judicial behavior to the modern "big tent" new legal realism); see also Cass Sunstein, et al. ARE JUDGES POLITICAL? AN EMPIRICAL ANALYSIS OF THE FEDERAL JUDICIARY (2006).

[60] Thomas Mitchell, *New Legal Realism and Inequality*, in THE NEW LEGAL REALISM: TRANSLATING LAW-AND-SOCIETY FOR TODAY'S LEGAL PRACTICE, ed. Mertz, Stewart Macaulay, and Thomas W. Mitchell (2016) at p. 216. Mitchell's work on land partition is an example of applied empirical legal research, culminating in model legislation that has been adopted in 20 states halting the facially neutral, but *de facto* discriminatory, partition of real property following the intestate death of Black landowners. Uniform Law Commission on the Partition of Heirs Property Act, https://www.uniformlaws.org/committees/community-home?CommunityKey=50724584-e808-4255-bc5d-8ea4e588371d.

[61] Brian Z. Tamanaha, *Legal Realism in Context*, in THE NEW LEGAL REALISM:

how individuals engage with the courts, including the measure of efficacy and outcomes of interventions in delivery of legal services, applies interdisciplinary studies in law and society and adopts the original legal realist Karl Llewelyn's second approach to realism - applying quantitative and qualitative approaches to the study of law.[62]

This is not a new idea. Studying the "law in action" shows up in the theories of the legal realists[63] who were thinking about courts and justice prior to and during the New Deal,[64] and is applied in both access to procedural justice and access to outcomes approach of modern New Legal Realists.[65] A focus on the relationship between social power structures and the legal system was restated by critical legal theorists, [66] and has been reinforced by critical race theorists and feminist legal theorists who bring a racial justice and feminist lens to decisions of judges and the operation of courts and the justice system in people's lives. The ideas proposed here, arguing for the value of empirical study in order to make normative recommendations to improve socio-legal systems, arise within the framework of New Legal Realism, in the empirical study of judges and judging and the embrace of

---

TRANSLATING LAW-AND-SOCIETY FOR TODAY'S LEGAL PRACTICE, ed. Mertz, Stewart Macaulay, and Thomas W. Mitchell (2016), at 165.

[62] Tamanaha, *supra* note 61 at 165, describing the two "main strains" of original legal realism as "a theoretical perspective on law as a social institution" and the second being the empirical study of law, overlapping with research design in the social sciences. The latter is applied here. See also Mark C. Suchman & Elizabeth Mertz, *Toward a New Legal Empiricism: Empirical Legal Studies and New Legal Realism*, 6 ANNUAL REVIEW OF LAW AND SOCIAL SCIENCE 555–579 (2010), at p. 559 discussing rigorous study design by qualitative and quantitative legal empiricists.

[63] Karl Llewellyn, trained as an anthropologist, stated that Legal Realism in practice is a method to "see [the law] fresh," "see it as it works" and adds the function questions "what is it for?" and "how has it been working?" concluding that realism is "a technology. That is why it is eternal . . . The fresh inquiry into results is always the needed checkup." Mertz, *supra* n. 59 at 30.

[64] Arguably, realist legal theory and an approach to how courts and the law impact people in practice ("law in action") led to redistributionist policy changes as legal realists left academia to join the New Deal as policymakers and government lawyers. Marcus J. Curtis, *Realism Revisited: Reaffirming the Centrality of the New Deal in Realist Jurisprudence*, 27 YALE J.L. & HUMAN. (2015).

[65] Rebecca Sandefur, *Access to Justice*, RESEARCH HANDBOOK ON MODERN LEGAL REALISM (2020) at 329.

[66] Challenging the "vision of law as a neutrally benevolent technique" and identifying "both 'law' and 'the economy' as belief systems that people have externalized and allowed to rule their lives." Gordon, Robert W. *New Developments in Legal Theory* in THE POLITICS OF LAW: A PROGRESSIVE CRITIQUE, edited by David Kairys, Pantheon Books, 1982 at p. 283 and 290.

ethnographical and qualitative studies of courts, legal services, the legal profession, and litigants.[67]

Scholars from law schools as well as from non-legal disciplines, such as sociology and econometrics, bring rigorous quantitative and qualitative methodologies to the theoretical framework of New Legal Realism and the study of unrepresented and self-represented litigants, making normative recommendations for closing the justice gap and improving substantive access to justice.[68] As Mitchell observes, however, "legal scholars, including [New Legal Realism] and law-and-society scholars, who are interested in having their scholarship shape the development of the law, must understand that law and policy do not develop in a vacuum and that any ideas they have for legal reform will not be self-executing."[69]

B. Access to Justice and Access to Data.

Recent, innovative, and insightful empirical studies of state court reform and the provision of legal services for self-represented litigants include studies qualitative and quantitative evaluation of new and innovative changes. The work of Stacy Butler and the Innovation for Justice Lab at the University of Arizona utilizing empirical study design to evaluate Arizona's pilot program for licensed legal advocates (LLAs) to provide legal services to domestic violence survivors and that of Anna Carpenter at the University of Utah also engaging in empirical study on the access to justice interventions for self-represented litigants are the forefront of empirical study of the provision of legal services that are not provided by a lawyer. [70] During the

---

[67] In a push for more qualitative and quantitative empirical research in law, Stewart Macaulay posits, "If we put law in context, we should begin to produce a more accurate picture of the actual operation of the legal system. We could abandon ideological claims based on idealistic images of law. However, what are the consequences of looking at law from the ground up? Would an unretouched picture of any nation's legal system damage its claims to legitimacy?" Macaulay, Stewart *A New Legal Realism: Elegant Models and the Messy Law in Action* in THE NEW LEGAL REALISM at p. 44.

[68] See Anna E. Carpenter et al., *Studying the New Civil Judges*, 2018 WIS. L. REV. 249–286 (2018) (a call for empirical analysis of the judicial role in state civil courts in pro se cases); Emily Taylor Poppe, *Homeowner Representation in the Foreclosure Crisis,* 13 JOURNAL OF EMPIRICAL LEGAL STUDIES, 809 (2016) (a quantitative analysis of demographic and geographic variables and correlations in disposition outcomes for judicial foreclosure cases);

[69] Mitchell, *supra* note 59 at 221.

[70] Stacy Butler, *Designing a New Tier of Civil Legal Professional for Survivors of Domestic Violence* (2019), https://law.arizona.edu/sites/default/files/Final%20Report%20with%20Appendices%20May%2024%202019.pdf; Colleen F. Shanahan, et al., *Can a Little Representation Be a Dangerous Thing Symposium: Advancing Equal Access to Justice:*

Covid-19 pandemic, empirical study of evictions has contributed significantly to the national and state conversations about housing and public health.[71]

The groundbreaking work of sociologist Rebecca Sandefur has contributed significantly to an interdisciplinary understanding of how people interact with the legal system.[72] Sandefur's research, and that of other scholars seeking to understand civil legal aid within the broader ecosystem of poverty and inequality, finds that most legal problems are addressed outside of the legal system.[73] Along with scholars in the legal design and design justice disciplines,[74] many access-to-justice scholars have proposed looking outside the traditional provision of legal services in order to reach these people that are caught in the "justice gap" that traditional legal services

---

Barriers, Dilemmas, and Prospects, 67 HASTINGS L.J. 1367–1388 (2015);

[71] Emily Benfer, et al., *Health Justice Strategies to Combat the Pandemic: Eliminating Discrimination, Poverty, and Health Disparities During and After COVID-19*, 19 YALE J. HEALTH POL'Y L. & ETHICS 122; Emily Benfer, *COVID-19 Eviction Crisis: An Estimated 30-40 Million People in America Are at Risk* (2020), https://pesquisa.bvsalud.org/global-literature-on-novel-coronavirus-2019-ncov/resource/pt/grc-739731; Anne Kat Alexander, *Residential Eviction and Public Housing: Covid-19 and Beyond*, 18 IND. HEALTH L. REV. 243, Stephen Metraux and Alex Guterbock, *Eviction and Legal Representation in Delaware-An Overview* (2020), https://udspace.udel.edu/handle/19716/26352; Sudeall, et al., *supra* note 57 (application and use of the Princeton Eviction Lab data set for empirical study of eviction during Covid-19 and normative recommendations for helping low-income renters avoid eviction using a public health and provision of public services as a social good lens).

[72] Rebecca L. Sandefur, *What We Know and Need to Know about the Legal Needs of the Public*, 67 S. C. L. REV. 443–460 (2015) (using qualitative research and comparative research to understand why many legal issues never make it to the court system); Sandefur, *Access to What?*, *supra* note 54; Sandefur, *Paying down the Civil Justice Data Deficit: Leveraging Existing National Data Collection*, 68 S. C. L. REV. 295 (2016) (identifying gaps in the data and knowledge about the provision of civil legal services and our "spotty knowledge about what kinds of help seem to be useful in specific contexts but no systematic understanding of who needs help with what kinds of problems, what kinds of help they need, and how to reach them to give it").

[73] Sandefur, *Access to What?, supra* note 54 at 51 (2019) (findings from Sandefur's empirical research show that most people do not cite the cost of an attorney as the reason for not seeking legal help, but that people often do not "consider law a solution for their justice problems."); Tanina Rostain, *Techno-Optimism and Access to the Legal System*, 148 DAEDULUS 93 (2019).

[74] Margaret Hagan, *A Human-Centered Design Approach to Access to Justice: Generating New Prototypes and Hypotheses for Interventions to Make Courts User-Friendly*, 6 IND. J.L. & SOC. EQUAL. 199 (2018); Margaret Hagan, *Participatory Design for Innovation in Access to Justice*, 148 DAEDALUS 120 (2019); Sasha Costanza-Chock, DESIGN JUSTICE: COMMUNITY-LED PRACTICES TO BUILD THE WORLDS WE NEED at 165 (MIT Press 2020) (discussing hack-a-thons in the legal services community that include community voices in the design of legal technology tools).

models do not reach.[75] Finally, the role of law school centers and institutes that study civil access to justice cannot be overlooked in valuable contributions to empirical study of the courts, and as models for evaluation of proposed and pilot reforms in the provision of legal services. The Civil Justice Reform and Legal Profession Reform projects of the Institute for the Advancement of the American Legal System takes an empirical approach with measurable results to informing normative recommendations for improving civil justice,[76] and are in partnership with the Utah courts to evaluate the Utah regulatory sandbox and engage in recommendations and evaluations for new systems of regulation of regulation for legal services providers.[77]

The research conducted in this project on the lack of court record data in California contributes to the scholarship and work of these and other empirical legal scholars working to reform courts and legal services to better serve self-represented litigants by identifying a gap in access to baseline state court record data against which to measure interventions and proposes normative recommendations to solve this problem through regulatory and legislative solutions.

III: ESTABLISHING NORMS FOR THE EVALUATION OF LEGAL TECHNOLOGY TOOLS.

A. A. Metrics for Coding Quantitative Court Data.

Along with recommendations for better data collection, this project proposes a set of norms for how to use quantitative data to evaluate legal technology tools that requires researchers take into account the way the law works in practice for self-represented litigants in state court, and specifically how consumer defendants experience the court system.[78] The National

---

[75] See Sandefur et al, *supra* note 53; Butler et al., *supra* note 70, Carpenter et al., *supra* note 31.

[76] IAALS, *Transforming Our Civil Justice System for the 21st Century: The Road to Civil Justice Reform*, https://iaals.du.edu/publications/transforming-our-civil-justice-system-21st-century-road-civil-justice-reform (2020); IAALS, *Preventing Whack-a-Mole Management of Consumer Debt Cases* (2020), https://iaals.du.edu/publications/preventing-whack-mole-management-consumer-debt-cases.

[77] IAALS *Project: Unlocking Legal Regulation*, *supra* note 27.

[78] "[C]ivil legal disputes are handled in more than 15,000 courts, resulting in a patchwork of jurisdictions among state, county, municipal authorities. An estimated 46 million people are appearing [annually] in the courts, handling cases involving divorce, custody, child support, guardianship, housing, and consumer. These courts consistently report through sampling that

Center for State Courts Court Statistics Project has created a Data Governance Policy Guide that makes recommendations for the collection and storage of court record data, which is an excellent step in the right direction for data normalization, data storage, and the sharing of data with outside entities.[79] Once that data is collected, unlike in qualitative research where users will subjectively report positive and negative experiences with court processes, determinations by researchers must be made as to how to code case outcomes and other events in court record data so that interventions, such as legal technology tools, may be measured against a baseline of case results.

Australian scholar Tania Sourdin notes that while there is a "dearth of research evaluating justice apps," frameworks for evaluating health apps can be adapted and used to evaluate substantive outcomes of justice apps.[80] Traditionally, "access to justice" has been a call to action to provide more attorneys, either through legal aid or pro bono services, [81] but a recent emerging movement to reform the court system to better serve unrepresented litigants looks to use technology to empower litigants to achieve optimal outcomes without a lawyer, leading to a vibrant access-to-justice technology movement not only in legal technology startup spaces, but also in non-profit organizations and in legal aid.[82] This approach looks at access to outcomes,

---

75% or more of these cases have at least one self-represented litigant." Self-Represented Litigants Network *SRLN Brief: How Many SRLs*? (2019), https://www.srln.org/node/548/srln-brief-how-many-srls-srln-2015.

[79] National Center for State Courts, *supra* note 51, (discussing data collection standards at 12-13 and access for internal and external data users at 17-19).

[80] Tania Sourdin, et al., DIGITAL TECHNOLOGY AND JUSTICE: THE USE OF JUSTICE APPS, 2020 ROUTLEDGE, LONDON UK (2021), at 27 (Sourdin proposes that researchers use the following criteria to evaluate justice apps: ease of use, effectiveness, privacy and security considerations, and interoperability.)

[81] "[T]raditional solutions to reducing the access to justice gap, such as increased funding for civil legal aid, more pro bono work, or court assistance programs have had some success, but are not likely to resolve the gap, which is only increasing in severity." Conference of Chief Justices, Resolution 2, *Urging Consideration of Regulatory Innovations Regarding the Delivery of Legal Services* (Feb. 5, 2020); Sandefur, *Access to Justice*, *supra* note 65 at 329.

[82] See, e.g., the Access to Justice Tech Fellows Program that equips "the next generation of justice leaders to better ensure equitable access to justice for all." https://www.atjtechfellows.org/. Many of these A2J tech projects are connected with law school clinics, such as the Suffolk Legal Innovation Lab, the Vanderbilt Program on Law and Innovation, the University of Denver Law and Innovation Lab, and through other partnerships between legal aid programs and law school technology programs, such as the UC Irvine Consumer Law Clinic partnership with Inland Counties Legal Services to create digital services deliver in the San Bernardino County Consumer Rights Legal Clinic.

as opposed to access to process.[83] Although more common in the criminal justice space, community organizers and economic justice advocates in civil justice activist spaces are amplifying the voices of those affected by the legal system in helping to redefine what is meant by access to justice.[84] However, when it comes to studying the efficacy and impact of new and innovative solutions, while legal technology innovations seek to "disrupt" the provision of legal services and the operation of courts in the interest of access to justice, there is still a strong bias toward efficiency and an assumption of neutrality in access to justice.[85] There is also the problem of what Tanina Rostain calls "techno-optimism" – an aspirational belief among legal technologists that that technology can "address the system's failings and the conditions of poverty more generally."[86]

The norms proposed here are premised on the non-neutral value statement that an optimal outcome for a self-represented consumer litigant in a debt collection lawsuit is one that does not contribute to the consumer's financial instability.[87] That courts should do justice in a way that is meaningful and serves the whole person – and that measurable outcomes for the litigant should inform how courts provide services – is an approach to law

---

[83] Rebecca Sandefur, *Access to Justice*, *supra* note 65 at 329 (research on procedural access to justice often looks at access to attorneys or legal services, while access to outcomes studies might examine "whether law achieves specific policy goals" or "allow people to assess the just-ness of their own experiences," or, as here, may adopt an explicit normative definition of a just outcome.

[84] Decolonizing Justice: Advancing Community-Grown Justice Solutions, a 2020 symposium, brought together thinkers and activists exploring how to radically change the approach of legal aid and the courts in making the justice system more just for traditionally marginalized communities and Black, indigenous, and BIPOC populations.

[85] See, e.g. Deno G. Himonas and Tyler J. Hubbard, *Democratizing the Rule of Law*, 16 STAN. J. C.R. & C.L. 261 (2020), discussing utilizing measurements of efficient disposition of cases as a metric for increased access to justice.

[86] Rostain, *supra* note 73 at 94 (2019). Rostain proposes that legal technology solutions alone overlook the "cultural, material, and educational hurdles" that low-income populations face when confronted with a legal problem and that "[e]ven when people recognize that their problem is legal, they face significant impediments to finding, understanding, and using online legal tools effectively," recommending that legal technology tools be "placed in the hands of people in positions of trust … so they can function as intermediaries between disadvantaged people and the legal system."

[87] *Aspen Institute Convening on the Rights of Debtors*, February 21, 2020, Washington D.C.; Catherine R. Albiston and Rebecca L. Sandefur, *Expanding the Empirical Study of Access to Justice Colloquium*, 2013 WIS. L. REV. 101–120 at 111 (2013) (measuring outcomes requires a definition of "effectiveness").

of letting practice guide theory, and vice versa.[88] The idea that justice is more than neutrally deciding cases on doctrinal law and appellate precedent is grounded in the idea the way people interact with courts matters, and that when courts are studied, researchers must study this interaction in order to improve justice.[89]

B. Mapping Debt Collection Case Outcomes.

Debt collection cases are not substantively value neutral.[90] They are a built-in tool in the business model of lenders who build rates of default into interest rates, and who prey on consumers by offering credit on unaffordable terms with high interest rates, penalty late fees, and penalty interest upon default.[91] Consumer installment loans and payday loans often have annual percentage rates of over 100%, and credit cards marketed to low-income borrowers frequently have interest rates of 24.99%, jacked to 29.99% under the terms of the contract if a payment is missed or late. Credit card lawsuits, and those brought by third-part debt buyers and assignees, seek interest and fees far beyond any amount by which the borrower may have been enriched, and that justice would ask she disgorge as unjust enrichment.[92]

In California, most debt collection cases are not brought as breach of contract, as the plaintiff often lacks the original contract or knowledge of the terms to plead the contract with specificity. Instead, consumer debt collection cases are brought as common counts actions, suing on the final balance on an account, with little inquiry into whether the consumer obtained a benefit of

---

[88] Mertz, *supra* note 59 at 10.

[89] The idea that "law lags behind social change," "moneyed interests are often served by law," and "a full appreciation of law must go outside legal doctrine to observe its actual consequences" are a "a complex of attitudes" named by the Legal Realists. Tamanaha, *supra* note 59; see also Christoper Bradley, *The Consumer Protection Ecosystem: Law, Norms, and Technology*, 97 DENVER L. REV. 35 (2020) for a discussion of consumer protection law and technology from a realist perspective.

[90] Cass Sunstein, FREE MARKETS AND SOCIAL JUSTICE, (Oxford University Press USA, 1999) at 3 "Free markets can produce economic inefficiency and (worse) a great deal of injustice. … In fact, free markets depend on a range of coercive legal interventions." Sunstein refers to property law here, but the same can be said of consumer contracts.

[91] Installment lender CashCall's loans are over 100% APR, with some as high as 343%. From 2011-2013, CashCall originated at least $269 million in loans in 17 states. "Four in every 10 CashCall borrowers defaulted on their loans." https://slate.com/news-and-politics/2019/10/cashcall-reddam-cfpb-loans-lawsuit.html. See also *De La Torre v. CashCall*, 5 Ca. 5th 966 (2018), holding that an interest rate may be so high as to violate California's unconscionability law.

[92] Dalié Jiménez, *Dirty Debts Sold Dirt Cheap*, 52 HARV. J. LEGIS. 41 (2015).

the bargain, or might be entitled to offset for unconscionable interest or other terms of contract drafted by the original creditor, in a contract of adhesion. So, as a starting point, in debt collection cases, unlike commercial breach of contract cases, judges don't operate to make the parties whole by calculating damages for breach based on an expected benefit baked into the terms of an agreement reached through the negotiation of the contracting parties. Instead, even more than in landlord-tenant cases, the court operates to enforce unilaterally against a borrower, who has already shown that she is unable to meet her financial obligations.[93]

Beyond the substantive problems with using the courts as a debt collector to collect from victims of predatory lending schemes, there are a significant number of debt collection lawsuits in which a consumer may raise a meritorious defense but finds herself precluded from prevailing because of procedural barriers.[94] However, the majority of debt collection lawsuits are ones in which a consumer did originally owe a debt to someone. Creditor attorneys have capitalized on the psychological effects of debt and shame[95] and due to internalized blame about debt and personal responsibility,[96] the consumer may approach the litigation seeking to make amends. A component of successful legal technology and self-help in this area is centered on consumer empowerment.[97] The findings from empirical research and randomized controlled trials on the provision of self-help materials demonstrate the same.[98] When we shift the frame of the consumer from a

---

[93] Johnson Raba, *supra* note 23 at 25 (student loan borrowers sued on defaulted private loans in California had executions of judgment for wage garnishment and bank levies returned unpaid 92.5 percent of the time, showing that borrowers who default simply do not have the money to pay their debts.)

[94] Peter A. Holland, *The One Hundred Billion Dollar Problem in Small Claims Court: Robo-Signing and Lack of Proof in Debt Buyer Cases*, 6 J. BUS. & TECH. L. 259, 259 (2011); Peter A. Holland, *Defending Junk-Debt-Buyer Lawsuits*, 46 CLEARINGHOUSE REV. 12, 14 (2012)

[95] Prohibitions on unfair debt collection activity such as contacting consumers at work show that one of the legislative goals of the Fair Debt Collection Practices Act recognized embarrassment as a component of debt collection abuse. 15 U.S.C. § 1692c; *see also* 15 U.S.C. § 1692e, prohibiting misrepresentations intended to disgrace the consumer.

[96] Taylor Poppe, *supra* note 25, at 152.

[97] The Consumer Financial Protection Bureau's Your Money Your Goals program is focused on consumer empowerment, https://www.consumerfinance.gov/consumer-tools/educator-tools/your-money-your-goals/; *see* also the San Francisco Financial Justice Project, https://sfgov.org/financialjustice/.

[98] Greiner, Daniel James and Jiménez, Dalié and Lupica, Lois R., Self-Help, Reimagined (February 15, 2016). IND. L. JOURNAL, Vol. 92, No. 3, 2017; IAALS, *The Future of Legal Services Speaker Series,* Rohan Pavuluri, co-founder/CEO Upsolve (2020),

person who owes to a person who can fight the lawsuit, we take the first step to achieve a better for that SRL. Litigant engagement is a quantitative measurement of empowerment.[99] A confusing and overwhelming court process and a sense that the consumer is doomed to lose results in default judgments.[100] An empowerment approach that demystifies the procedural process should result in more engagement by litigants and fewer default judgments. Accordingly, coding qualitative data to identify consumer involvement in debt collection lawsuits at any point of the process should be coded as a value against which to measure intervention, demonstrating engagement with the court system and locating points of intervention and improvement.

The second piece of setting norms for data and coding values in quantitative court record data is evaluation of outcome data.[101] Determining an optimal outcome for a self-represented litigant is tiered and branching. Tiered outcomes explore the impact that a default judgment or other objectively negative outcome may have on a consumer, based on the consumer's credit score, income, source of income, and housing stability. Additional quantitative data beyond court records may be necessary to identify and describe optimal good outcomes based on the consumer's ability to pay to the creditor's claim. Identification of additional information, such as patterns in credit data among consumer debtors, may be necessary to map the distribution of outcome data and identify points of intervention that are targeted to consumer needs.[102] Outcome norms are branching in that other factors inform best outcomes for the consumer based on factors arising from both the consumer debtor, i.e., personal financial circumstances, other consumer debt, the long-term ability to abide by any possible settlement

---

https://iaals.du.edu/events/future-legal-services-rohan-pavuluri-andrea-s-jarmon-and-andrew-arruda.

[99] *But see* Catherine R. Albiston and Rebecca L. Sandefur, *Expanding the Empirical Study of Access to Justice Colloquium*, 2013 WIS. L. REV. 101–120, 112 (2013) (sociolegal research shows that "legal rights affect individuals' self-conception and identity in complex and counterintuitive ways" and that not all people find empowerment from engaging in legal processes). Albiston and Sandefur also note at 113 that creditors may take advantage of social stigma around debt to persuade consumers to enter into a repayment plan on a defaulted debt. Katherine Porter, *Bankrupt Profits: The Credit Industry's Business Model for Postbankruptcy Lending*, 93 IOWA L. REV. 1369, 1401-02 & fn. 164-65 (2008).

[100] Taylor Poppe, *supra* note 25, at 163.

[101] Sandefur, *Access to Justice*, *supra* note 44 at 329.

[102] Ongoing research by the author under a project funded by the Pew Charitable Trusts explores this further. Data on file with author.

agreement, eligibility for bankruptcy, and factors arising from the creditor plaintiff, i.e., whether the plaintiff is an original creditor or third party debt collector, the amount of the debt, and whether the creditor has a "consumer rights policy" of considering hardship for dismissals or settlements. Many of the second set of factors are knowable from the quantitative court record data and an examination of consumer creditors' practices.

C. Baseline Quantitative Data and Research Design.

Quantitative court record data can contribute to the effective design of research studies on legal technology tools that seek to improve outcomes for self-represented litigants, but in the absence of available data, there is a concern that the free market, and for-profit advertising driven business models, may inadvertently may harm certain demographics of consumers in a greater and more significant way than others who have less access to resources to understand the limited scope of tools that engage in creatively skating on the boundary of the unauthorized practice of law. For example, most researchers and legal technology app developers would agree that a randomized controlled trial in which one self-represented litigant consumer receives inaccurate and misleading information and another receives accurate information is harmful and unethical. It may be posed that it is also unethical to provide assistance only with an answer (first responsive pleading) to a debt collection lawsuit for all users but charge money for attorney review or other services. Some legal technology apps run by for-profit companies and currently on the market do just that.[103]

Quantitative court record data that shows when and where cases resolve through dispositive motions can identify need for technology tools that provide comprehensive know-your-rights information. Without such data, regulatory sandboxes run the risk of greenlighting tools that may assist with an answer but fail to provide any follow-up for procedural next steps and how the consumer litigant might respond when served with discovery. The consequences of such partial assistance may result in a consumer's failure to respond to discovery and a motion to deem admissions admitted,

[103] An example of such products include Solosuit, an entrant into the Utah regulatory sandbox that is built on a "freemium" model, in which only an answer is provided for free and the user must pay to have an attorney review the document or provide any additional assistance. https://www.solosuit.com/ Such a model does not translate well to California where the civil litigation process is complex and discovery and law-and-motion practice are common tools of the debt collector plaintiff.

which may lead to a motion for judgment on the pleadings. Regulatory sandbox committees are looking closely at standards to ensure that legal technology apps that provide inaccurate information, but there are additional evaluation metrics that are difficult to name without sufficient quantitative data about how cases with self-represented travel through the court system and how these cases resolve. While these regulatory questions seek to provide guidance about using technology to augment, or even replace, the practice of law, these are at heart data-based questions that can be answered in part by examining civil court record data and determining where the self-represented consumer litigant fails out of the system and needs support, assistance, and a trusted resource to complete the lawsuit to a just outcome.

What do we mean when we say a just outcome? A number of self-help centers in California report that they do not assist with consumer debt collection lawsuits because they are worried about doing more harm than good.[104] They express concern that if they help a consumer file an answer, the litigant may end up incurring attorney's fees and costs and may end up owing more than if they defaulted on the lawsuit. The self-help centers also report that they do not know how to handle assisting with discovery responses without crossing the barrier between legal information and legal advice, which self-help centers cannot provide. These are valid concerns and legal technology application developers should consider these among other potential harms caused by a "one-touch" solution or other unbundled service that is not adequately connected to trusted and vetted resources.

Access to civil court record data intersects with human-centered design and demographic research by identifying demographic, cultural, and geographic differences necessary in designing legal technology tools that can be used effectively by self-represented litigants in all counties in large states, or across multiple states. While some of this can be addressed through qualitative research, without a baseline of court record data and outcomes, and a coding of that data to measure intervention, researchers are missing a huge set of information necessary to inform evaluation of new technologies. Quantitative data from court case information in California demonstrates the

[104] The California State Bar Closing the Justice Gap Working Group at its June 18, 2021 meeting considered risk-based criteria for entrants to the regulatory sandbox with metric of considering whether a consumer (user) of a justice app has a better outcome than if they remained self-represented (no intervention). https://board.calbar.ca.gov/Agenda.aspx?id=16253&t=0&s=false.

necessity for state-level reporting of normalized data fields in order that researchers and legal technology app developers can identify the needs, identify the points of failure in the existing support systems for self-represented litigants, and revise and update the tools and empirically analyze the effects of court systems on litigants.

D. Civil Court Record Data and Digital Courts.

The move to digital courts is described by Richard Susskind as an inevitable outgrowth of client needs and technological improvements.[105] Susskind says that clients seek the "outcomes that professionals bring," which are "practical results (a job done) and emotional effects (an appropriate feeling, perhaps of reassurance and confidence."[106] Susskind's conclusion is that when technology can produce these outcomes "cheaper, better, quicker, or more convenient[ly] than the current offering, we can expect the market to switch to the alternatives."[107] Which brings us to the question of how we measure these outcomes or if we are letting the market drive the definition of outcomes. [108]

Legal technology solutions can mean any or all of the pieces that help a self-represented litigant, or an attorney, automate parts of the procedural pieces of the legal system. The role of a legal technology solution can also be to engage in outreach to litigants to educate them about their rights. Legal technology solutions developed under regulatory frameworks which restrict the unauthorized practice of law may succeed by "tying a local lawyer to its sale of forms," as Rocket Lawyer has done, or by designating legal forms as legal information, not legal advice, as LegalZoom has done in response to a number of lawsuits.[109] Mass marketed legal forms are available for free on many court self-help websites, and back in in 2010, 4 million people visited the California self-help website.[110] However, these forms are confusing and

---

[105] Richard Susskind, ONLINE COURTS AND THE FUTURE OF JUSTICE, (2019) at 48.
[106] Susskind, as cited in Thomas Kuhn, THE STRUCTURE OF SCIENTIFIC REVOLUTIONS (1996) at 151
[107] Susskind*, supra* note 105 at 48.
[108] Clayton M. Christensen et al., *What is Disruptive Innovation ?*, HARVARD BUSINESS REVIEW MAGAZINE (2015), https://hbr.org/2015/12/what-is-disruptive-innovation; Ian Bremmer, *State Capitalism Comes of Age: The End of the Free Market*, 88 FOREIGN AFF. 40 (2009); Katherine Porter, BROKE – HOW AMERICA BANKRUPTS THE MIDDLE CLASS (2012).
[109] Barton, Benjamin and Stephanos Bibas, REBOOTING JUSTICE (2017) at 136-37.
[110] Barton, *supra* note 109 at 120; *Colloquium: The Legal Profession's Monopoly on the Practice of Law: Protecting the Profession or the Public? Rethinking Unauthorized-Practice*

difficult to navigate for many consumers. Some self-help centers, such as the Sacramento Law Library, provide PDF guides to help consumers to complete forms and draft motions.[111]

Some pilot projects for online courts aimed at consumer debtors have not seen widespread adoption, in part because of a lack of data to evaluate efficacy and outcomes. In New York, attempts to institute an online dispute resolution (ODR) program for credit card debt collection cases was halted out of concern for lack of evaluation metrics and ability to compare the ODR program to existing self-help services run by LawHelpNY to help SRLs engaging in rigorous study of the regulatory sandbox, the ODR program has instituted a mandatory online dispute resolution process with a trained ODR facilitator in small claims matters. [112] Preliminary findings showed that cases move faster with ODR – previously, cases took an average of 144 days to disposition, and with ODR, they take 84 days.[113] This begs the question of whether case management efficiency is consistent with seeking more just outcomes for self-represented users of the state court civil legal systems. Justice Himonas of the Utah Supreme Court concluded that the use of ODR by consumer debtors expedited access to justice because of the shorter time and cites to the fact that default rates fell by 4%, informed defaults decreased to 3% of defaults, and settlement increased by 4%.[114] These numbers start to show how empirical data case be used to show outcomes, but particularly with ODR, unless such research is combined with qualitative research to ensure that the participants are obtaining a just outcome, such as an affordable payment plan, the numbers alone present insufficient data for metrics of outcomes. Indeed, further empirical analysis of the Utah ODR program coupled with interviews with ODR participants shows that over a longer study period, the default rate for consumer cases increased significantly following the imposition of mandatory ODR for consumer debt cases, with defaults localized among a small set of institutional debt collection plaintiffs.[115]

Enforcement, 82 FORDHAM L. REV. 2587, 2591-91 (2014).

[111] Sacramento Law Library forms, available at https://saclaw.org/self-help/civil-self-help-center/.

[112] Himonas, *supra* note 74 at 58.

[113] *Id* at 58.

[114] *Id*.

[115] Todd Feathers, *Payday Lenders Are Big Winners in Utah's Chatroom Justice Program*, THE MARKUP, Mar. 16, 2022, reporting findings that support the hypothesis that "any system

The California State Bar recognizes that the "California Justice Gap Survey revealed that a significant portion of the justice gap in California is caused by a lack of knowledge about the legal system."[116] Through a survey conducted with Californians who experienced a legal problem, the State Bar found that many people do not "recognize the legal aspects" of issues and try to "deal with problems on their own rather than seek legal help for their otherwise actionable civil legal issue."[117] So do we choose adoption of a legal tool as an effective measurement of success? Does mere use increase access to justice? Barton in Rebooting Justice notes that there is a "relative dearth of reported harm from LegalZoom, which has existed since 2001," and draws from this the conclusion that if "we are going to ban or curtail vastly less expensive online legal services for consumer protection, we need actual evidence of harm. As of yet, there is none."[118]

How do legal technology developers, particularly those which in the legal aid and public space, demonstrate efficacy beyond mere adoption or user engagements? This data is hiding in civil court record data and civil court case outcomes. Tools such as A2J Author and HotDocs and Tyler Guide and File are notoriously non-user friendly. The next phase of legal technology tools operate in the same mobile-friendly space, user-centered design, and with the same flexible and intuitive one-click operations that consumers have come to expect from their online experiences from shopping to registration.

An analysis of civil court record data at a granular level can provide significant initial guidance as how to define harmful, helpful, neutral outcomes. These same analyses may very well demonstrate that we need to simplify the court processes. The Oregon Judicial Department, by decree of the Administrative Office of the Courts, has relaxed the rules of evidence in family law matters where both sides are self-represented, demonstrating that collaboration between courts and legal technology applications can help to simplify overly complex civil litigation processes.[119] Oregon, like Utah, uses

---

that makes it easier and faster to litigate debt collection lawsuits is likely to disadvantage low-income individuals."

[116] State Bar of California, *Justice Gap Study Exec. Summ*. (2019) at 10, https://www.calbar.ca.gov/Portals/0/documents/accessJustice/Justice-Gap-Study-Executive-Summary.pdf.

[117] *Id.*

[118] Barton, *supra* note 109 at 134.

[119] William J. Howe III and Jeffrey E. Hall, Oregons' Information Domestic Relations Trial: A New Tool to Efficiently and Fairly Manage Family Court Trials 55 FAM. COURT REV. 70 (2017), https://www.ncsc.org/__data/assets/pdf_file/0031/19696/oregons-

a unitary system across the entire state, implemented on the Odyssey system owned by Tyler Technologies for case management, financial management, and electronic documents in one system, making it easier to evaluate and understand civil court record data.[120]

One concern in letting the market and market share take the place of carefully defined norms and setting standard for outcomes is shown by the story of tax filing in the United States. Filing taxes is unnecessary burdensome and the market for the generation of tax returns is dominated by a few industries – TurboTax owned by Intuit, on the forms generation side, and non-accountant para-professional tax preparers on the services side, such as Jackson Hewitt and H&R Block. Intuit effectively shut down a plan to allow taxpayers to automatically pay their taxes through the ReadyReturn program run by the IRS, by lobbying "heavily in opposition, including $13 million in federally lobbying and a million dollars to oppose a candidate for the California comptroller who supported ReadyReturn."[121] This is a cautionary tale for courts regarding the entrance of market-based solutions that propose to solve the justice gap.

When evaluating which court systems benefit consumers and which stand in the way of progress, "[a]n evidence-based approach can also justify when burdens are needed and providing demonstrable value."[122] Access to justice combined with the disruptive nature of technological advances means that sometimes court processes will also need to change to ensure that legal technology is not merely replicating unjust systems.

E. Applications for a Regulatory Sandbox.

A regulatory sandbox creates a new set of regulations for approved technology companies to provide legal services without creating an attorney-

---

informal-domestic-relations-trial-a-new-tool-to-efficiently-and-fairly-manage-family-court-trials.pdf.

[120] The Oregon Judicial Department (OJD) uses the Odyssey Case Management System, a Tyler Technologies product. This system was installed, as part of the Oregon eCourt project, across all Oregon circuit courts and the Oregon Tax Court between June 2012 and June 2016. The system provides case management, financial management, and electronic document management within a single system. https://www.courts.oregon.gov/Pages/faq.aspx.

[121] Herd, Pamela and Donald P. Moynihan, ADMINISTRATIVE BURDEN: POLICYMAKING BY OTHER MEANS, at 10; Liz Day, *How the Maker of TurboTax Fought Free, Simple Tax Filing.* PROPUBLICA (March 26, 2013), https://www.propublica.org/article/how-the-maker-of-turbotax-fought-free-simple-tax-filing.

[122] *Id.* at 254.

client relationship and avoids concerns under existing regulations regarding the unauthorized practice of law. Utah pioneered the idea of a regulatory sandbox for legal services in 2019. [123] In Utah, the Supreme Court authorized a sandbox "overseen by a regulator in which legal entities will experiment with consumer-focused innovation."[124] As Justice Deno Himonas notes, "form reform" and a deep dive into making forms available to the public, including through the expansion of Utah's Online Court Assistance Program (OCAP) is an excellent start but is insufficient to solve the substantive problem of access to justice.[125] Utah, it must be noted, as a unitary court system and a centralized case management system, which creates a fertile ground for empirical study by the Regulatory Reform Task Force of the both the pre-reform Utah court system, and the effects of changes implemented through the sandbox. Thus far, the California working group has not considered evaluation of sandbox entrants using civil court record data, electing instead to focus on qualitative evaluation such as focus groups and surveys.[126]

The California Closing the Justice Gap Working Group was formed in 2020 to explore the formation of a regulatory sandbox and follows the genesis of a working group to explore licensing paraprofessionals to engage in work that under current licensing regulations would be considered the unauthorized practice of law.[127] The California Task Force on Access

---

[123] Other states that have looked into creative and non-traditional solutions for addressing the access-to-justice crisis include Arizona (AZ Code of Jud. Admin. § 7-201(A) Limited License Legal Practitioner), Illinois (Future of Legal Services Report 2016), Oregon (Oregon Futures Task Report on Paraprofessionals), Virginia (Virginia Future of Law Practice), and Washington (Limited Practice Rule for Limited License Legal Technicians). Utah is the first to launch a legal regulatory sandbox. See Himonas, *supra* note 74. North Carolina is the most recent state to pass a law authorizing the creation of a regulatory sandbox, in House Bill 624. HB 624 is awaiting the Governor's signature as of the current draft of this paper. https://www.billtrack50.com/BillDetail/1366450. Some consumer protection concerns have been raised regarding the North Carolina model. See, e.g., Lee Reiners, *North Carolina's Proposed Regulatory Sandbox Needs Work*, THE FINREG BLOG GLOBAL MARKETS CENTER DUKE UNIVERSITY SCHOOL OF LAW (May 28, 2019), https://sites.law.duke.edu/thefinregblog/2019/05/28/north-carolinas-proposed-regulatory-sandbox-needs-work/.

[124] Himonas, *supra* note 85 at. 59.

[125] Himonas, *supra* note 85 at 51-52.

[126] Agenda for first meeting of the California Closing the Justice Gap Working Group, http://board.calbar.ca.gov/Agenda.aspx?id=16015&t=0&s=false.

[127] Closing the Justice Gap Working Group Appointment of Members Agenda Item 702 (September 24, 2020), http://board.calbar.ca.gov/docs/agendaItem/Public/agendaitem1000026576.pdf.

Through Innovation of Legal Services recommended on January 31, 2020, in advance of its February 4, 2020 meeting, to establish a regulatory sandbox that permits the formation of entities permitted to provide legal services that would previously have been prohibited under the existing California Rules of Professional Responsibility. California authorized the formation of the Closing the Justice Gap Working Group on May 14, 2020 and subsequently approved a working group charter and put out a call for nominations for working group members on September 24, 2020.[128] The call for nominees asked for participants from a wide variety of disciplines, showing promise for a diverse set of voices from the access-to-justice advocacy community, including consumer and defense attorneys, an ethics expert, a technology expert, a regulator, a trial court judge, a consumer of legal services, an economist, a national legal services organization, and a member of the Utah Regulatory Reform Task Force, among others.[129]

The working group as formed has a number of diverse voices and is considering the structure and governance of a proposed sandbox regulatory authority.[130] The California working group considered model applications for entrants into the regulatory sandbox at the August 18, 2021 meeting.[131] Of concern, the model application does not inquire into applicants' business model, does not ask how they intend to close the justice gap, and does not require any data on unmet need and how the proposed entrant intends to engage in qualitative or quantitative evaluation of the efficacy and outcomes of their product.

On December 7, 2021, the chairs of the California Senate and Judiciary Committees expressed concern with the State Bar's focus on a regulatory sandbox and the proposed regulatory changes to the rules governing the unauthorized practice of law, suggesting that the State Bar focus instead on its "core mission of policing attorney misconduct" and offering traditional pro bono, legal aid, and court-sponsored self-help services.[132] In response, the California Closing the Justice Gap placed on hold

---

[128] *Id*.
[129] *Id.*
[130] Closing the Justice Gap Working Group Agenda (Aug. 18, 2020), https://board.calbar.ca.gov/Agenda.aspx?id=16332&tid=0&show=100030305#10037995.
[131] *Id.*
[132] California State Legislature, letter from Assemblymember Mark Stone and Senator Tom Umberg, Dec. 7, 2021, https://s3.documentcloud.org/documents/21151650/state-bars-ctjg-

its regulatory sandbox committee meetings, "to allow time for further conversations and determine the next best steps."[133] This pause in the regulatory sandbox is an opportunity for the State Bar and the Judicial Council of California to engage in effective reform of collection of civil court record data in order to provide a data-driven response to lawmakers regarding identified areas for intervention for technology providers in the access-to-justice space.[134]

IV. THE LANDSCAPE OF ACCESS TO CIVIL COURT RECORD DATA.

A. The Federal Courts' PACER System.

Researchers, academics, and developers of legal technology startups are making great strides in attempting to collect court record data, but state court civil record data has challenges that make it hard to follow behind other kinds of court record data that has moved into the realm of big data. Technology is not new to the courts. In the United States, the federal court system has been using a centralized e-filing system (Electronic Court Filing, or ECF) for represented parties, began in 1996 and by 2007, 98 percent of federal court filings had moved to electronic filing.[135] State civil court systems have been significantly slower to adopt e-filing, and adoption is scattered in a non-unitary court system like California's.[136] The use of third party vendors for e-filing and case management creates a further lack of aggregated data.

Federal court record data is kept on a single centralized database system, PACER, and bankruptcy, federal criminal, and federal civil cases have been scraped and the data recombined by both public and private entities. A class action lawsuit was filed by NCLC, Alliance for Justice, and National Veterans Legal Services Program,[137] arguing that excessive costs

---

concerns-12-7-21.pdf

[133] Law.com, *State Bar Pauses 'Regulatory Sandbox' Work After Criticism by Lawmakers*, THE RECORDER, (Dec. 2021), https://www.law.com/therecorder/2021/12/13/state-bar-pauses-regulatory-sandbox-work-after-criticism-by-lawmakers/

[134] Claire Johnson Raba, *Court Record Data Access and Non-Profit Innovators Can Help Close the Justice Gap*, UNICOURT BLOG (Jan. 24, 2022), https://unicourt.com/blog/justice-gap/

[135] National Center for State Courts Trends 2007 report http://cdm16501.contentdm.oclc.org/utils/getfile/collection/tech/id/570/filename/571.pdf.

[136] Data on e-filing systems in California's 58 counties on file with author.

[137] *Nat'l Veterans Legal Servs. Program v. United States*, 235 F. Supp. 3d 32, 2017 U.S. Dist. LEXIS 9447, 96 Fed. R. Serv. 3d (*Callaghan*) 928, 2017 WL 354084

violates federal law committing PACER fees solely to the operation of the PACER database. The PACER system fees are authorized by the E-Government Act of 2002, which allows fees "only to the extent necessary" to operate the database.[138] Amicus briefs filed by journalists argue that the fees inhibit journalists' ability to access data[139] and former judges, including Richard Posner argued for a removal of all access fees.[140] The bipartisan Open Courts Act, H.R. 8235,[141] introduced by Representatives Hank Johnson (D-Ga.) and Doug Collins (R-Ga.), was passed by the House on December 8, 2020. House Judiciary Committee Chairman Jerry Nadler said in his opening statement, "It is a disservice that in today's digital age, the public's access to public records in public proceedings is so resource-intensive and burdensome. This fragmented and costly approach does not reflect the modern standards of access the public deserves."[142]

The Open Courts Act requires the creation of a new "system for all [federal] public court records" and will require the establishment of data standards and industry standard design standards, including user-centered design, provide search functions, and "make public court records automatically accessible to the public upon receipt of such records."[143] State court record data in many states is significantly more fragmented, necessitating a not only a migration of data but a normalization across case management systems within states in order to provide public access, as discussed in Part VI, and in the proposed model legislation. The cost of migrating the existing unified PACER system will be funded by increased filing fees in federal court, as necessary to fund the new system, and the cost of ensuring ongoing access to the system will be funded by an amount equal

---

[138] Matt Ford, *The Courts Are Making a Killing on Public Records*, THE NEW REPUBLIC, (Jan. 31, 2019), https://newrepublic.com/article/153003/courts-making-killing-public-records-pacer-fees

[139] *Nat'l Veterans Legal Servs. Program v. United States* Case No. 1901081(L) and 1901083 in the U.S. Court of Appeals for the Federal Circuit, Brief of *Amici Curiae* The Reporters Committee for Freedom of the Press and 27 Media Organizations in Support of Plaintiff-Appellants, (Jan. 23, 2019), https://www.rcfp.org/wp-content/uploads/2019/01/2019-01-24-NVLSP-v-US.pdf.

[140] Cite to Posner and judicial amicus brief.

[141] H.R. 8235 (2020), https://www.congress.gov/bill/116th-congress/house-bill/8235.

[142] Jacqueline Thomsen, *House Judiciary Advances Bipartisan Bill to Make PACER Free to Public,* THE NATIONAL LAW JOURNAL, (Sept. 15, 2020), https://www.law.com/nationallawjournal/2020/09/15/house-judiciary-advances-bipartisan-bill-to-make-pacer-free-to-public.

[143] H.R. 8235 §§ 1, 4.

to federal agencies' 2018 annual fees paid into PACER, with no cost to the public.[144]

B. California's Disaggregated and Outsourced System.

State court data collection systems vary widely in their systems, constructions, and access to data for the public. Court systems like Utah, that are run on a unified case management system, provide clean and normalized data access for researchers, but many other states have disaggregated systems in which multiple counties or jurisdictions have separate contracts with vendors for case management services. Data collection problems which arise from disaggregated systems in state court civil cases are exemplified by California's 58-county system. California's legislature spent two billion dollars on a multi-year failed attempt at implementing a centralized case management system.[145] The California Centralized Case Management System (CCMS) was authorized by California Government Code Section 68106.2(g) and was intended to integrate all the legacy systems in use by the county courts but was abandoned by the legislature after sinking millions of dollars into a system that never materialized.[146]

In California, the two houses of the legislature have a Democratic supermajority, but many issues and advocates compete for parts of the budget allocation, ensuring that any bill that includes budget appropriation will receive particularly careful scrutiny and require that it serves a necessary purpose. Proposed legislation for transparency and affirmative data sharing must consider the history of civil court record data in the 58 California counties and the "two-billion-dollar boondoggle" of the California Court Case Management System (CCMS), on which the Legislature pulled the plug in 2012.[147]

In a forward-thinking move toward data access through direct democracy, California voters passed Proposition 220 in 1998, beginning the

---

[144] HR. 8235 §§ 2(f)(3), 3(c).

[145] Maria Dinzeo, *Two Billion-Dollar Boondoggle Criticized by California Judges*, COURTHOUSE NEWS SERVICE (Apr. 14, 2010), https://www.courthousenews.com/two-billion-dollar-boondogglecriticized-by-california-judges/.

[146] Charles Horan, *California's CCMS boondoggle is but a symptom*, SFGATE (Feb. 24, 2011), https://www.sfgate.com/opinion/openforum/article/California-s-CCMS-boondoggle-is-but-a-symptom-2474386.php.

[147] Dinzeo*, supra* note 145; Horan, *supra* note 146.

process of unifying California's trial and municipal courts.[148] AB233, passed in 1997, and effective January 1, 1998, authorized funds for a system of state funding for trial courts and followed the Trial Court Realignment and Efficiency Act of 1991. In its preamble, AB 233 states, "[i]t is increasingly clear that the counties of California are no longer able to provide unlimited funding increases to the judiciary" and that the stated intent of the Legislature is to "[p]rovide that the State of California shall assume full responsibility for any growth in costs" beyond a cap for county contributions to be held at the county court operations cost in the 1994-95 fiscal year.[149] Among other changes, AB 233 amended section 77009 of the Government Code to allocate a Trial Court Operations fund for centralized court services within each county.[150] Accordingly, funding for all 58 county courts is already a centralized process and it would be appropriate for the state legislature to require data collection and reporting from the courts.

In 2002, the California Legislature authorized the Administrative Office of the Courts to contract with a vendor to create a California Court Case Management System in an attempt to modernize and standardize California's trial courts. Originally bid at $260.2 million to serve six counties, by 2010 it carried a cost of $1.9 billion and contractor Deloitte was unable to deliver a functional system.[151] In 2010, the county of Sacramento withdrew from the project, stating that it required clerks to manually input more information and that the system "asked for the same information over and over."[152] The judges, in exiting the CCMS project, raised concerns that the Deloitte project included many functions that the court did not need but did not meet the needs of the existing court case management.[153]

---

[148] California State Auditor, *Administrative Office of the Courts: The Statewide Case Management Project Faces Significant Challenges Due to Poor Project Management* (2011), http://www.auditor.ca.gov/pdfs/reports/2010-102.pdf.

[149] California AB 233 (1997), http://www.leginfo.ca.gov/pub/97-98/bill/asm/ab_0201-0250/ab_233_bill_19971010_chaptered.html.

[150] *Id.*

[151] Dinzeo*, supra* note 145; Horan, *supra* note 145; California State Auditor, *supra* note 148 at 6.

[152] Dinzeo*, supra* note 145.

[153] A judge's words, regarding the CCMS capability to find out how judges in other counties ruled on small claims matters were, "The [Administrative Office of the Courts] might care, but I certainly don't," said Maino, likening CCMS to a Ferrari with the capability of going 220 miles an hour, stating, "I can't legally go that fast. So why are we doing this?" Dinzeo*, supra* note 145. Modern judicial analytics software programs now regularly report how various judges rule in types of cases.

Attempts were made by the Legislature to address the concerns raised by the auditor about the CCMS in the 2012 legislative session and salvage it, but the Legislature determined that the CCMS was a failure, creating a barrier to future contracts for a centralized case management system.[154] In 2020, a bill was introduced seeking to require reporting of aggregated data to the Judicial Council of California, but this unfunded mandate was rejected by the JCC as unworkable and prohibitively expensive.[155]

Solutions for requiring the necessary level of data collection can come from court administrative agencies or from state legislatures. California's Judicial Branch Statistical Information System, discussed *supra*, is potentially one place to start, but the current set of aggregate data collected does not provide a level of detail necessary to run analysis on this data. The model legislation in Appendix A proposes the creation of an office of data analysis within the court administrative agency, to be properly funded, and to require that all county courts provide docket level reporting information that include all relevant fields. Legacy data must be normalized and this data must be made available to researchers, legal aid organizations, academics, and those who seek to develop technology tools to further access to justice.

C. National Efforts to Reform Court Data Collection.

In recognition of the need to get a more accurate baseline of data on what occurs in state courts, efforts are underway to collect state court record data nationally. The National Center for State Courts Court Statistics Project Data Governance Policy Guide provides standards for court administrative offices that are working to make their data more uniform.[156] A cohort of court docket data researchers are working to develop a taxonomy for common fields necessary for data analysis that would dramatically improve the ability of legal technology innovators to identify areas for automation as appropriate

---

[154] AB 2119 (Nielson 2012), which died at committee desk, sought to establish an office of the California Technology Agency and required any expenditure of the AOC over $500 million to be approved by this new office, http://www.leginfo.ca.gov/pub/11-12/bill/asm/ab_2101-2150/ab_2119_bill_20120223_introduced.html.

[155] AB 2271 (Gabriel 2020) sought to collect data on unlawful detainer (eviction cases). The Judicial Council of California objected to this bill's unfunded mandate requiring the Judicial Council to aggregate and host eviction case data for researchers and tenants' rights advocates. https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201920200AB2271.

[156] National Center for State Courts, *supra* note 51.

points for intervention and to allow the testing of outcomes for tools.[157] The Legal Services Corporation is actively working to on a project to scrape court data through the front-end portals of many states and counties. However, this effort is hampered by access issues. The LSC is only able to access data in a limited number of counties in some states because of two primary barriers: terms of service which prohibit the automated retrieval of public record data; and electronic captchas which prohibit such data scraping.[158] Academics, researchers, and non-profit or court-sanctioned legal technology tools should have access to the information that so many parties are going to tremendous lengths to obtain. Data scraping creates additional problems in that the data retrieved requires significant normalization because the fields are not uniform.

Among the most egregious issues, as discussed in the case study in Part I, is the lack of a field to indicate that a party is self-represented. Normalization of data across the entire time that the case is open in the case management system is required to determine with any accuracy when parties are self-represented, and if they are receiving limited scope representation or self-help assistance. In many counties in the data studied for California, only those who self-identify as "in pro per" by so stating on their pleadings are coded as such in most of the court case management systems. The LSC reports that in its research, there is not a single county in the United States from which it has scraped data that affirmatively reports whether a user is self-represented. A county in Texas used to report this data up through the mid-2000s when it stopped reporting this information.[159]

This reflects that the case management systems used by courts are designed to be used by lawyers and for case management of judicial dockets without consideration of datapoints that would be easy to collect and report at filing and disposition that would lead to baseline data information which would significantly increase the ability to use empirical court record data against which to test legal technology tools. Court record data access is recognized as a need by national researchers.[160] The Pew report on state courts and debt collection cases demonstrates the value of such data, but in

---

[157] Pew Charitable Trusts Civil Court Modernization Project, Court Docket Data Researchers Convening, data on file with author.
[158] Manjarrez, *supra* note 21.
[159] *Id.*
[160] Pew Debt Collection, *supra* note 4 at 18.

conjunction with the findings by LSC, the states missing from the Pew report show the severity of limitations due to disaggregated systems such as the California 58-county system.[161] Reliance only on data from unified court systems where the information is easy to access excludes from research datasets huge amounts of data and failing to recognize geographic and population differences that might require development of legal technology tools tailored to the needs of diverse populations.

D. The Limits of Justice Gap Studies.

In an absence of quantitative state court record data, one of the primary current tools states use for identifying legal needs for identifying the most pressing access-to-justice needs is a Justice Gap study, modeled after the Legal Services Corporation's legal needs study.[162] The State Bar of California's 2019 Justice Gap Study [163] is a report from data gathered through research conducted by survey of legal needs users, utilizing primarily multiple-choice questions, and of legal aid offices, engaging in a four-week tracking of unserved people. Such studies have been several limits, as identified by the OECD and the Open Society Foundation's guidance on legal needs surveys.[164] This can be compared to research utilized in other countries, in which the legal needs survey is the first step, followed by a refined checklist to "proactively ascertain the legal needs of individuals before the legal need is addressed."[165] This 'second generation' approach, used in community legal clinics in Canada, "could also be used at a systemic level

---

[161] *Id.*

[162] LEGAL SERVICES CORPORATION, *The Justice Gap: Measuring the Unmet Civil Legal Needs of Low-Income Americans*, (Jun. 2017), https://lsc-live.app.box.com/s/6x4wbh5d2gqxwy0v094os1x2k6a39q74

[163] The State Bar of California, *The California Justice Gap: Measuring the Unmet Civil Legal Needs of Californians* (2019) at 43, http://www.calbar.ca.gov/Portals/0/documents/accessJustice/California-Justice-Gap-Report.pdf.

[164] Organisation for Economic Co-operation and Development (OECD) and Open Society Foundation, Legal Needs Surveys and Access to Justice at 30-31, https://www.oecd.org/gov/legal-needs-surveys-and-access-to-justice-g2g9a36c-en.htm; McDonald, *supra* note 13 at 730.

[165] Organisation for Economic Co-operation and Development (OECD) and Open Society Justice Initiative *Understanding Effective Access to Justice* (2016) at 9, https://www.oecd.org/gov/Understanding-effective-access-justice-workshop-paper-final.pdf

by governments and legal services providers" to identify and anticipate legal needs based on patterns and demographics of communities.[166]

Both at the national level and at the state level, the primary study model for evaluating the need for "closing the justice gap" are survey-driven studies conducted with consumers of legal services and IOLTA-funded legal aid offices in order to draw a conclusion on unmet legal needs.[167] This type of study, first conducted by the Legal Services Corporation in 2005, and most recently conducted in 2017, has consistently shown that more than 80 percent of the civil legal problems reported by low-income Americans received inadequate or no legal help.[168] While the national and California Justice Gap studies survey data conducted through the intake census show that certain people are unserved by traditional LSC-funded legal aid programs, in seeking to make the leap to using a Justice Gap study to recommend solutions, there is an initial problem of establishing a common definition and measurable framework for terms like "underfund" and "need" in an exploring the challenging in calculating the benefits of providing access to legal services, placing at issue the very premise the research necessary to identify a "justice gap" is commiserate with identifying the solutions to address unmet need.[169]

Based on its legal needs study, the California State Bar recommendations include "[i]dentify technology and nontechnology based approaches to create more affordable legal service for those who will not qualify for legal aid, but who cannot pay the current market rate for attorney services" and "fund projects addressing the most common types of problems faced by Californians: health, finance, employment, and income maintenance."[170] In addition, the State Bar notes that California "recognizes the need for legal innovation and regulatory reform" to "stimulate the creation of new legal service models to reduce the justice gap" and references the ATILS commission, including the delivery of legal services through technology, AI, and online legal service delivery models. The State Bar recognized its limitations in this report in recommending that California

---

[166] OECD, *supra* note 165 at 9.
[167] Legal Services Corporation, *The Justice Gap: Measuring the Unmet Civil Legal Needs of Low-income Americans* (2017) at 10, https://www.lsc.gov/sites/default/files/images/2017/06/TheJusticeGap-FullReport.pdf.
[168] *Id* at 8.
[169] Prescott, J. J. "The Challenges of Calculating the Benefits of Providing Access to Legal Services." Fordham Urb. L. J. 37, no. 1 (2010): 303-46.
[170] State Bar Justice Gap Executive Summary, *supra* note 116 at 26.

"[c]ollect more robust data on self-represented litigants so that approaches to addressing the needs of this population can be informed by current and comprehensive data"[171]

The use of a study model designed to increase funding for traditional legal aid programs lacks support as a framework against which new and innovative solutions can be tested because by design it measures the existing service structure. Reliance on Justice Gap studies reinforces the need for additional data for empirical analysis of these issues and a framework against which to evaluate the outcomes for low-income litigants in their interactions with state court systems. Justice Gap studies are valuable initial tools, but they cannot be utilized to make normative recommendations. Traditional Justice Gap study design evaluates "inputs and outputs, as distinct from outcomes."[172] There is a movement to increase the study of outcomes for legal aid clients,[173] but the current model of the Justice Gap study does not include data to make those evaluations. While the Justice Gap study shows a need to fund legal aid at higher rates in order to do more legal aid, if state judicial administrative bodies are thinking about new and innovative solutions, those solutions needs to be targeted at the needs of the litigant. The Justice Gap studies do not measure that information, nor do they provide a baseline framework against which to evaluate pilot projects or technology solutions.

Utilizing survey data that reports the status of existing systems to make recommendations to change those systems skips the important step of studying the impact of proposed changes. As an example, the Justice Gap study draws a conclusion that problems related to housing and immigration are the most common types of problems people seek help from legal aid with, but this is drawn from the fact that these are the most common type of cases closed by legal aid offices.[174] While this reflects community needs, it also

---

[171] State Bar Justice Gap Executive Summary, *supra* note 116 at 26.

[172] National Center for Access to Justice, *Tracking Outcomes: A Guide for Civil Legal Aid Providers and Funders* (2018) at 12, https://ncforaj.org/wp-content/uploads/2018/06/NCAJ-Outcomes-Guide-complete-for-6-20-18.pdf

[173] *See,* e.g., Tanina Rostain's *Just Connect Us* Data Commons Project (2019) https://www.nsf.gov/awardsearch/showAward?AWD_ID=1952067&HistoricalAwards=false.

[174] *Id* at 44. ("Figure 14 shows the cases closed by State Bar-funded legal aid organizations in 2018, according to their annual case summary reports.29 Housing problems, including rental housing and homeownership/foreclosure issues, comprised 21 percent of the cases closed by these legal aid organizations in 2018. Immigration comprised an additional 13

primarily reflects the priorities of particular legal aid offices and often, sources of funding that are tied particular practice areas.

Utilizing a Justice Gap study to make normative recommendations for solutions overlooks that this survey data is only the first step in devising solutions to unmet legal needs. An example of this is that the two leading issues identified by respondents are not actually legal needs. Respondents report as legal needs a perception that they were billed incorrectly for medical services and that health insurance wouldn't cover needed medical procedures.[175] These are problems with the health insurance system, not the legal system, and filing a lawsuit would not resolve these types of issues. The report used only surveys of legal aid offices and of families living in poverty but engaged in no analysis of civil court record data and no evaluation of outcomes of cases for litigants engaged with the court system.[176] Although there is no data showing that litigation would resolve the self-identified issues, the State Bar recommendations include "[i]dentify technology and nontechnology based approaches to create more affordable legal service for those who will not qualify for legal aid, but who cannot pay the current market rate for attorney services" and "fund projects addressing the most common types of problems faced by Californians: health, finance, employment, and income maintenance."[177] The State Bar recognized its limitations in this report in recommending that the "[c]ollect[ion of] more robust data on self-represented litigants so that approaches to addressing the needs of this population can be informed by current and comprehensive data."[178]

The solutions to the justice gap, nationally and in California, should be broad and comprehensive, and solutions need to be undertaken holistically, not piecemeal. The American Academy of Arts & Sciences project Making Justice Accessible: Designing Legal Services for the 21st

---

percent.30 Health problems, including medical insurance denials and disputes as well as other health and long-term care issues, comprised 10 percent of closed cases, as did problems categorized as "miscellaneous.")

[175] State Bar Justice Gap, *supra* note 116 at Fig. 4 at 24; *see also* the self-identified "legal issue" that is online scam, which is not a judiciable civil legal problem that can be remedied through the courts, as recovery from such scammers is rarely possible.

[176] State Bar Justice Gap Study Technical Report, 2019, State Bar of California, https://www.calbar.ca.gov/Portals/0/documents/accessJustice/California-Justice-Gap-Survey-Technical-Report.pdf

[177] State Bar Justice Gap Executive Summary, *supra* note 116 at 26.

[178] *Id.*

Century engaged in extensive study to provide a set of "clear, national recommendations for closing the justice gap," culminating in the comprehensive report Civil Justice for All.[179] The study methodology utilized by Prof. Minow and her team demonstrates a way to use Justice Gap studies to engage in further study and makes a set of recommendations, including recognizing the needs for civil court record data to evaluate the efficacy of new delivery models for providing help with civil justice needs.[180]

E. Recommendations for Legislative Change.

It is in the furtherance of the public good for state legislatures and court administrative bodies to take the lead in reform of data access in order that civil court record data information be reported in a way that permits the development of legal technology tools that can be tested, evaluated, and used by self-represented litigants in conjunction with self-help, legal aid, and law schools in order to advance access to justice.[181]

The greatest barrier to passing an authorizing statute requiring increased reporting of civil court record data is the cost to the local county courts, which requires budget appropriation by the state legislature. However, in Assembly Bill 1331, California has already recognized the importance of data collection in policy decisions in relation to criminal case record data, which makes California an ideal model for reform of civil record data collection from civil cases.[182] The proposed legislation in Appendix A ensures that this data would be collected in a way that benefits researchers and would make anonymized court data available to researchers, non-profits, law school clinics, and policy advocates. Existing JBSIS data is already available on the court administrative office website in downloadable tables; the proposed legislation and rulemaking would extend this public record access to a set of records that would contain meaningful and usable data.

The proposed model legislation in Appendix A is premised on the idea that the state should bear the burden to collect the raw data from the each of county courts and normalize the data. Given the proof of concept of this in

---

[179] American Academy of Arts & Sciences, *Civil Justice for All* (2020) at 6, https://www.amacad.org/publication/civil-justice-for-all.

[180] *Id* at 25, recommending a national organization to collect data on the state of civil justice.

[181] American Academy of Arts & Sciences, *supra* note 179 at 9. 35.

[182] Measures for Justice, *AB 1331 Passage in California a Win for Increased Data Transparency in the Justice System* (Sept. 11, 2019), https://www.measuresforjustice.org/news/2019-09-12-ab-1331-passage-in-california.

collecting the data on debt collection cases, this is not an insurmountable, nor even an incredibly challenging, task from a data normalization perspective. The proposed model legislation recommends the establishment of public database to be hosted at a university research facility or non-profit data science research entity, such as the California Policy Lab. In addition, third-party for-profit data aggregators and vendors are also engaged in collecting state court data. The model legislation does not place restrictions requiring that the data host and aggregator be a non-profit, opening the option to utilize any third-party vendor that complies with open data standards to aggregate, normalize, anonymize, store, and host state civil court record data.

CONCLUSION

As states consider innovations to close the justice gap and examine alternatives to the traditional provision of civil legal aid, these actions must consider the impact on vulnerable self-represented litigants in state court. Ensuring consumer protection isn't possible without sufficient data to understand the landscape of the problem and empirically evaluate pilot projects. Building effective legal technology tools that result in just outcomes for self-represented litigants requires a baseline of data and access to ongoing data in a way that provides a method for testing tools in the field over longitudinal data points. As state courts look to adopt user-friendly and mobile-friendly technology tools, the developers of those tools should not be designing in a vacuum and should not need to engage in costly and burdensome market research to devise tools that are effective in addressing access-to-justice issues, effectively excluding non-profit and legal aid technology developers from the regulatory sandbox. As courts move to adopt legal technology that behaves in ways that digital users expect from modern and mobile-friendly interfaces, it is critical that those developing the legal technology do so in ways that is data-driven, trustworthy, responsive, led by community-centered design principles, and builds on relationships with governmental, community-based, and non-profit partners.

APPENDIX A

Model Legislation for State Court Data Collection

Statement of Intent:

The state of [State] recognizes the need for robust data collection of state court record data in order to effectively implement access-to-justice initiatives. The [State] recognizes that there is a "justice gap" in the ability of low-income litigants to access traditional legal aid services. While [State] seeks to encourage creative and innovative solutions to increase access to justice for low- and moderate-income litigants, [State] has an interest in ensuring that access-to-justice initiatives do not cause harm to self-represented litigants and that services as implemented ensure and improve equal access to services for all residents of [State].

[State] further recognizes that access to public record data is a constitutionally and statutorily protected right of the public and acknowledges the need for a centralized and regularly updated normalized dataset of docket-level state court record information. This bill proposes a method of data collection, normalization, and anonymization in order to allow access-to-justice technology innovators, researchers, and advocates access to anonymized public record information in order to improve the operations of the courts of [State].

Fiscal Impact:

Due to the disaggregated and non-centralized case management systems operated by the counties in [State], the fiscal impact of creating a centralized database of civil court record information will require ongoing budget allocation in order to maintain an up-to-date set of normalized records. First-year start-up costs will be a one-time cost for a vendor or university to create a database system and normalize fields . Funds shall be allocated through an increase to the [State] Trial Court Operations Fund.

SEC. 1. Section is added to the Government Code, to read:

(a) A centralized database shall be established for the purpose of providing public access to civil court record data.
    (1) The Legislature shall solicit bids from vendors for the creation of a database to be made available to the public civil court record data.
    (2) Civil court record docket data, as described in Sec 2, shall be collected, maintained, and processed to be made available to the public in the following manner:
        a. Fields shall be normalized;

b. Cases in which litigants are self represented shall be identified and a field shall be created for the purpose of identifying such cases;
c. Party name fields shall be anonymized if they are associated with cases for self-represented litigants;

(b) The centralized database shall be updated monthly to contain current docket information on case filings and dispositions.
(c) The database system vendor shall make available to the public at no cost anonymized civil court docket data, including but not limited to, field identified in Sec. 2 of this section, in a downloadable format, such as Excel, .csv, or other format utilized by standard data analysis software.

SEC. 2. Section is added to the Government Code, to read:

(a) For each case filed or dispositioned in each county court, the superior court of the county in which the case was filed shall provide docket-level data to the database vendor established by the Legislature in Sec. 1. County courts shall provide reporting as follows:
(1) County courts shall cause to be added to their case management system the following fields:
a. A field to indicate whether a litigant received assistance from the court's self-help office;
b. Disposition data fields that include as a minimum the following: type of judgment, including whether judgment was default, after dispositive motion, after bench or jury trial, or pursuant to agreement; whether dismissals are with or without prejudice; if disposed of by motion, the type of motion;
c. A field to track the document type of any post-judgment filings;
d. A field to indicate whether an interpreter was requested and the language for interpretation.
(2) County courts shall provide a monthly dataset in Excel, .csv., or other format readable with standard data analytics tools, an output of docket fields, containing case management system output for each case filed or dispositioned during the month prior, including at a minimum, as applicable:
a. Case number, party names, attorney names and addresses, jurisdiction, case type, judicial officer assignment, filing date, detailed disposition data including all fields enumerated in Sec. 2(a)(1)(b), disposition date, and the dates and document types of post-judgment filings.
b. Whether a case received assistance from the court's small claims self-help office;

c. Whether an interpreter was requested and the language for interpretation.

(3) Data for the prior month shall be generated and caused to be sent to the database vendor established in Sec. 1 by the third day of each month.

SEC. 3. If the Commission on State Mandates determines that this act contains costs mandated by the state, reimbursement to local agencies for those costs shall be made pursuant to Part xxxx of the Government Code.