DePaul Law Review

Volume 72
Issue 2 *Spring 2023*

Article 4

The New Judicial Governance: Courts, Data, and the Future of Civil Justice

David Freeman Engstrom

RJ Vogt

THE NEW JUDICIAL GOVERNANCE: COURTS, DATA, AND THE FUTURE OF CIVIL JUSTICE

*David Freeman Engstrom**
*R.J. Vogt***

DEPAUL L. REV. (CLIFFORD SYMPOSIUM)

INTRODUCTION

The pandemic has been a potent disruptor of the American legal system. Yet, as with so many other aspects of American life, COVID-19 was most powerful as an *accelerant* of trends already in motion. Nowhere has this been more evident in law than in the civil justice system's uptake of new legal technologies. Consider five ways this is so:

* LSVF Professor in Law, Co-Director, Deborah L. Rhode Center on the Legal Profession, Stanford Law School; tel.: 650-721-5859; email: dfengstrom@law.stanford.edu.
** J.D. Candidate, Deborah L. Rhode Center Civil Justice Fellow, Stanford Law School.

- After years of futurist proclamations,[1] court-based online dispute resolution (ODR) platforms have entered the mainstream. Pre-pandemic, ODR was already in use in at least sixty-six active sites in twelve states.[2] Early empiricism on the pandemic's effect suggests that number has nearly doubled,[3] and will likely grow from here.[4]
- Only a trickle pre-pandemic, virtual court proceedings—not ODR, but online versions of formal court proceedings—are now a flood, to the tune of millions of hours in many states.[5] Switch-

1. *See, e.g.*, RICHARD SUSSKIND, ONLINE COURTS AND THE FUTURE OF JUSTICE 194-95 (2019) (predicting ODR's pervasive presence as one version of technology displacing attorneys where services can be routinized and commodified); David A. Larson, *Technology Mediated Dispute Resolution (TMDR): A New Paradigm for ADR*, 21 OHIO ST. J. ON DISP. RESOL. 629, 630–31 (2006) (imagining the way technology and dispute resolution would intersect in the 2020s); Arno R. Lodder and John Zeleznikow, *Developing an Online Dispute Resolution Environment: Dialogue Tools and Negotiation Support Systems in a Three-Step Model*, 10 HARV. NEGOT. L. REV. 287, 300 (2006) ("There is no reason why offline disputes could not be resolved online."); ETHAN KATSH & ORNA RABINOVICH-EINY, DIGITAL JUSTICE: TECHNOLOGY AND THE INTERNET OF DISPUTES 151 (2017).

2. *See* ABA Center for Innovation, *Online Dispute Resolution in the United States: Data Visualizations 2* (Sept. 2020), https://www.americanbar.org/content/dam/aba/administrative/center-for-innovation/odrvisualizationreport.pdf (recording 66 "active sites" of court-linked ODR by November 2019); Amy J. Schmitz, *Expanding Access to Remedies through E-Court Initiatives*, 67 BUFF. L. REV. 89, 92–93, 119 (2019) (discussing how courts in Michigan, Ohio, and New York are innovating and developing online dispute resolution pilot projects to resolve certain types of disputes; further noting that 50-60 courts are "looking to launch new projects").

3. In a survey of state chief justices and court administrators, eight states reported that their state's courts had begun using ODR during the pandemic, often in local pilots. *See Court Structure and Technology Survey*, on file with author. By April 2020, a single ODR platform provider, Matterhorn, reported that it was operating in 70 jurisdictions. *See* Avital Mentovich, J.J. Prescott, & Orna Rabinovich-Einy, *Are Litigation Outcome Disparities Inevitable? Courts, Technology, and the Future of Impartiality*, 71 ALA. L. REV. 893, 930 n.211 (2020).

4. For more details, *see note* 62, *infra* and accompanying text. ODR growth is likely because key institutional actors support its expansion, many even before the pandemic hit. *See, e.g.*, ABA Commission on the Future of Legal Services, *Report on the Future of Legal Services in the United States* 6 (2016) (recommending that "[c]ourt-annexed online dispute resolution systems should be piloted and, as appropriate, expanded"), https://www.americanbar.org/content/dam/aba/images/abanews/2016FLSReport_FNL_WEB.pdf; Conference of Chief Justices, *Call to Action: Achieving Civil Justice for All* 37 (2016) (calling for "creating on-demand court assistance services"), https://iaals.du.edu/sites/default/files/documents/publications/cji-report.pdf; National Center State Courts, *What Is ODR?* (last visited Jan. 17, 2023) (offering a wide range of supportive materials on ODR), https://www.ncsc.org/odr/guidance-and-tools/. For a fascinating case study from the pandemic, *see* David Allen Larson, *Designing a State Court Small Claims ODR System: Hitting a Moving Target in New York During a Pandemic*, 22 CARDOZO J. OF CONFLICT RESOLUTION, 569, 569–70 (2021).

5. The statistics from Michigan are striking but representative of many states. By mid-February 2021, Michigan trial courts had conducted more than 2.3 million hours of online hearings. Michigan residents had used the state's Virtual Courtroom Directory to find a hearing on YouTube more than 200,000 times. Collectively, trial court YouTube channels had nearly 75,000 subscribers. *See* Email from Chief Justice Bridget McCormack, Michigan Supreme Court, to David Freeman Engstrom (2021), on file with author.

ing costs have been paid in full, meaning remote access will continue even as the pandemic recedes.[6]

- Many courts that had stubbornly resisted have now adopted e-filing for civil pleadings and papers.[7] The old days of paper files stashed in far-flung courthouses, where they were pricey to maintain, time-consuming to retrieve, and nearly impossible to search,[8] are giving way to centralized repositories of machine-readable documents and metadata.
- Lawyers, forced out of their offices and outside their comfort zones, have quickened their embrace of legal tech tools that augment and even supplant what they do. Powered by yet another leap in the power of natural language processing[9]—the branch of machine learning that performs text analytics—machines are rapidly improving their capacity to perform core legal cognitions, from legal analytics to predicting the outcomes of cases. The race is on among law firms and tech companies to capture the value these new tools offer.[10]

---

6. A Reuters survey of more than 238 judges found that 93% conducted remote hearings in 2020 and 89% were still doing so in 2021. Gina Jurva, *The Impacts of the Covid-19 Pandemic on State & Local Courts Study 2021: A Look at Remote Hearings, Legal Technology, Case Backlogs, and Access to Justice*, THOMSON REUTERS 2–3 (2021), https://legal.thomsonreuters.com/content/dam/ewp-m/documents/legal/en/pdf/white-papers/covid-court-report_final.pdf.

7. *See* notes 240–44 *infra* and accompanying text. For a useful overview, including historical trends in adoption of e-filing, *see* Michael Thompson et al., *How Courts Embraced Technology, Met the Pandemic Challenge, and Revolutionized Their Operations*, The PEW CHARITABLE TRUSTS 1 (Dec. 2021), https://www.pewtrusts.org/-/media/assets/2021/12/how-courts-embraced-technology.pdf. Examples of states that have increased e-filing for at least some case types include Massachusetts, Alaska, and New Hampshire. *See Housing Court Standing Order 1-20: Implementation of mandatory electronic filing for attorneys in summary process and small claims cases in the Housing Court Department*, MASS.GOV, https://www.mass.gov/housing-court-rules/housing-court-standing-order-1-20-implementation-of-mandatory-electronic-filing; *see E-Filing Information*, ALASKA COURT SYSTEM, https://courts.alaska.gov/efiling/index.htm#current-courts; *see also Supplemental Rules of the Circuit Court of New Hampshire for Electronic Filing*, NEW HAMPSHIRE JUDICIAL BRANCH, https://www.courts.state.nh.us/rules/dmcr/dmcr-sup-efile.htm.

8. *See, e.g.*, Jenni Bergal, *Courts Plunge into the Digital Age*, THE PEW CHARITABLE TRUSTS: STATELINE (Dec. 8, 2014), https://www.pewtrusts.org/en/research-and-analysis/blogs/stateline/2014/12/8/courts-plunge-into-the-digital-age (noting that the status of courthouses' digital use "has been changing dramatically in many courthouses across the country. States are moving to systems in which documents are submitted electronically, file rooms are disappearing and the judicial system is going paperless").

9. One notable advance is OpenAI's GPT-3 model, and its ChatGPT offspring, which any member of the public can use to perform drafting tasks that are nearly indistinguishable from human-authored ones. *See generally* OpenAI, *ChatGPT: Optimizing Language Models for Dialogue*, OPENAI, https://openai.com/blog/chatgpt/ (last visited Jan. 17, 2023). For a celebratory but critical analysis of GPT-3's capacities, *see* Will Douglas Heaven, *OpenAI's New Language Generator GPT-3 Is Shockingly Good—And Completely Mindless*, MIT TECH. REV. (July 20, 2020), https://www.technologyreview.com/2020/07/20/1005454/openai-machine-learning-language-generator-gpt-3-nlp/.

10. For a comprehensive overview of the legal tech industry and the academic literature that has begun to grow up around it, *see generally* David Freeman Engstrom & Jonah B. Gelbach, *Legal Tech, Civil Procedure, and the Future of Adversarialism*, 169 U. PA. L. REV., 1001 (Mar.

- Finally, the pandemic deepened what was already a growing recognition that American courts at all levels are in the grips of a pro se and access crisis.[11] Rising economic insecurity and tech-powered institutional plaintiffs have steadily transformed our courts, particularly state and local ones, into debt collection and eviction mills whose main function is generating judgments for debt collectors and landlords against millions of debtors and tenants who overwhelmingly lack lawyers.[12] Accompanying this transformation is the realization that the standard proposals to

---

2021). For indications of an uptick in adoption of lawyer-driven legal tech, *see* Kate Beioley, *The Battle To Win at Legal Tech*, FINANCIAL TIMES (May 26, 2021) (tracking venture capital flows), https://www.ft.com/content/66853b7c-b62a-461e-9d85-bb805e8dff97; *see* A.J. Shankar, *The Pandemic Might Be the Tech Disruptor the Legal Industry Needs*, FORBES (Feb. 8, 2021) (summarizing recent survey research), https://www.forbes.com/sites/forbestechcouncil/2021/02/08/the-pandemic-might-be-the-tech-disruptor-the-legal-industry-needs/; *see* Nicole Black, *Top 5 Legal Technology News Stories of 2021*, ABA JOURNAL (Dec. 21, 2021), https://www.abajournal.com/columns/article/the-top-5-legal-technology-news-of-2021 (noting "the unprecedented consolidation of legal technology companies" through acquisitions). Another key development is tech-forward law firms' formation of subsidiaries that are explicitly designed to monetize the firm's bespoke service expertise in product form. A good example is Wilson Sonsini's SixFifty, *see* https://www.sixfifty.com/.

11. For a small slice of the rapid growth of thinking and concern around access to justice issues, *see generally* Brittany K.T. Kauffman & Brooke H. Meyer, *Transforming Our Civil Justice System for the 21st Century: The Road to Civil Justice Reform*, IAALS (Apr. 2020), https://www.ncsc.org/__data/assets/pdf_file/0019/36424/IAALS-113-Transforming-Civil-Justice_FINAL.pdf; *see generally Access to Justice*, DAEDALUS (Winter 2019), https://www.amacad.org/daedalus/access-to-justice; *see generally* Legal Services Corporation, The *Justice Gap: Measuring the Unmet Civil Legal Needs of Low-Income Americans* (June 2017), https://www.lsc.gov/sites/default/files/images/TheJusticeGap-FullReport.pdf; *see generally* New York State Permanent Commission on Access to Justice, *Report to the Chief Judge of the State of New York* (Nov. 2018), http://ww2.nycourts.gov/sites/default/files/document/files/2019-10/18_ATJ-Comission_Report.pdf. Some even suggest national/federal reform might be in the offing. *See generally* Daniel Wilf-Townsend, *National Civil Justice Reform: A Proposal for New Federal-State Partnerships*, THE GREAT DEMOCRACY INITIATIVE (Mar. 2020), https://greatdemocracyinitiative.org/wp-content/uploads/2020/03/Civil-Justice_Townsend_Final.pdf. Moreover, the DOJ just revived an office devoted to access to justice issues. *See* U.S. Dept. of Justice, *Office of Access to Justice* (last visited Jan. 17, 2022), https://www.justice.gov/archives/atj. For more views on the baleful state of the market for legal services and related reform imperatives, *see generally* GILLIAN K. HADFIELD, RULES FOR A FLAT WORLD: WHY HUMANS INVENTED LAW AND HOW TO REINVENT IT FOR A COMPLEX GLOBAL ECONOMY (2017); *see generally* Gillian K. Hadfield, *Legal Markets*, 60 J. ECON. LIT. 1264 (Dec. 2022); *see generally* FREDERICK WILMOT-SMITH, EQUAL JUSTICE: FAIR LEGAL SYSTEMS IN AN UNFAIR WORLD (2019).

12. Paula Hannaford-Agor, Scott Graves, & Shelley Spacek Miller, *The Landscape of Civil Litigation in State Courts,* NAT'L CTR. FOR STATE CTS., at iv, vi, 32 (2015) (finding that one side lacks a lawyer in some 75% of filed civil cases), https://www.ncsc.org/__data/assets/pdf_file/0020/13376/civiljusticereport-2015.pdf; *How Debt Collectors Are Transforming the Business of State Courts*, THE PEW CHARITABLE TRUSTS (May 6, 2020), https://www.pewtrusts.org/en/research-and-analysis/reports/2020/05/how-debt-collectors-are-transforming-the-business-of-state-courts ("Research on debt collection lawsuits from 2010 to 2019 has shown that less than 10 percent of [debtor] defendants have counsel, compared with nearly all [debt collector] plaintiffs."). For more on the prevalence of pro se representation at the federal level, *see* Judith Resnik, *A2J/A2K: Access to Justice, Access to Knowledge, and Economic Inequalities in Open Courts and Arbitrations*, 96 N.C. L. REV. 606, 607–08 (2018) (compiling statistics).

> fix what is broken, from civil *Gideon*[13] to simplification,[14] are inadequate or even chimerical. Surveying this sorry landscape as the pandemic recedes, many states are considering a bold solution: deregulating the legal services industry in order to welcome new, non-lawyer legal service providers—including the non-human, software sort—into the system.[15]

Each of these tech trends—ODR, online courts, e-filing, legal tech, and lawyer deregulation—will transform the civil justice system as we enter the post-COVID-19 era. And a flowering of scholarship, including several contributions to this Clifford Symposium (Symposium), explores the possibilities and perils of each trend within its own four corners and on its own terms.[16] New studies are measuring court-linked ODR's effect on case outcomes, particularly default judgment rates, while also questioning assumptions about what kind of cases are suitable for ODR-based adjudication.[17] Research is also asking urgent questions about new virtual fora, including whether online systems ad-

13. Among the many essays describing and critiquing "civil *Gideon*" are Rebecca Aviel, *Why Civil Gideon Won't Fix Family Law*, 122 YALE L.J. 2106 (2013); Jessica K. Steinberg, *Demand Side Reform in the Poor People's Court*, 47 CONN. L. REV. 741, 745 (2015); Tonya L. Brito, David J. Pate Jr., Daanika Gordon, & Amanda Ward, *What We Know and Need To Know About Civil Gideon*, 67 S. CAR. L. REV. 223, 224–25 (2016); Benjamin H. Barton, *Against Civil Gideon (and for Pro Se Court Reform)*, 62 FLA. L. REV. 1227 (2010). For the case that established a right to counsel in certain criminal proceedings, *see* Gideon v. Wainwright, 372 U.S. 335, 343 (1963).

14. Colleen F. Shanahan & Anna E. Carpenter, *Simplified Courts Can't Solve Inequality*, 148 DAEDALUS 128, 130 (2019).

15. *See generally* DAVID FREEMAN ENGSTROM, LUCY RICCA, GRAHAM AMBROSE, & MADDIE WALSH, DEBORAH L. RHODE CTR. ON THE LEGAL PRO., LEGAL INNOVATION AFTER REFORM: EVIDENCE FROM REGULATORY CHANGE, (Sept. 2022), https://law.stanford.edu/wp-content/uploads/2022/09/SLS-CLP-Regulatory-Reform-REPORTExecSum-9.26.pdf; *see generally* Rebecca L. Sandefur, Thomas M. Clarke, & James Teufel, *Seconds to Impact?: Regulatory Reform, New Kinds of Legal Services, and Increased Access to Justice*, 84 LAW & CONTEMP. PROBS. 69 (2021).

16. *See, e.g.*, Paula Hannaford-Agor, *Our New Normal? How COVID-19 Accelerated Pre-Pandemic Trends in State Court Litigation*, 71 DEPAUL L. REV. 279 (2022); Jean R. Sternlight & Jennifer K. Robbennolt, *In-Person or Via Technology?: Drawing on Psychology to Choose and Design Dispute Resolution Processes*, 71 DEPAUL L. REV. 537 (2022); Valerie P. Hans, *Virtual Juries*, 71 DEPAUL L. REV. 301 (2022); Herbert M. Kritzer, *COVID-19 and the Multiple Worlds of Litigation,* 71 DEPAUL L. REV. 393 (2022).

17. *See generally* J.J. Prescott & Alexander Sanchez, *Platform Procedure: Using Technology to Facilitate (Efficient) Civil Settlement*, *in* SELECTION AND DECISION IN JUDICIAL PROCESS AROUND THE WORLD: EMPIRICAL INQUIRIES (Yun-chien Chang, ed.) (Cambridge Univ. Press, 2020); Norman W. Spaulding, *Online Dispute Resolution and the End of Adversarial Justice?*, *in* LEGAL TECH AND THE FUTURE OF CIVIL JUSTICE 251, 257-262 (Engstrom ed., 2023); Alex Sanchez & Paul Embley, *Access Empowers: How ODR Increased Participation and Positive Outcomes in Ohio*, in TRENDS IN STATE COURTS at 14, NCSC (2020), https://www.ncsc.org/__data/assets/pdf_file/0018/42156/Trends_2020_final.pdf.

vance or impair the truth-seeking function of litigation,[18] advantage or disadvantage pro se litigants,[19] or alter the perspectives and even the psychological makeup of the system's various participants.[20] Last, a fast-growing literature examines the use of legal tech tools, including their impact on the litigation playing field[21] and their uneasy fit within the existing regulation of legal services and lawyer discipline.[22]

But sitting beneath these otherwise disparate trends are a deeper set of challenges and opportunities that will shape the civil justice system's future trajectory. Indeed, what binds together the above developments is that each has increased the *amount* of data in the system, data's *centrality* to the system's workings, or both things. A key question for the future, then, is whether the civil justice system—and, in particular, the courts that sit at its center—can harness new data flows in ways that promote the just, equitable, and efficient administration of justice. With pandemic-buffeted courts entering the digital age, a new mode of judicial governance is emerging; choices made today will have big consequences tomorrow.

In the pages that follow, we make and defend two core arguments. First, digitization means datafication—and, with the legal system generating data like never before, its ability to deliver justice will increasingly depend on the health of its data ecosystem. New forms of digitization, from online courts to ODR, are facilitating efforts by courts to collect and analyze data on cases, case outcomes, and the characteristics of litigants.[23] Chief justices of state supreme courts now regularly discuss how structured and usable their data are—or can be-

18. *See* Renee Danser, D. James Greiner, Elizabeth Guo, & Erik Koltun, *Remote Testimonial Fact-Finding*, *in* LEGAL TECH AND THE FUTURE OF CIVIL JUSTICE 93, 94-111 (Engstrom ed., 2023). For a focused analysis in the context of juries, *see generally* Hans, *supra* note 16.

19. Victor Quintanilla, *Digital Inequalities and Access to Justice: Dialing into Zoom Court Unrepresented*, *in* LEGAL TECH AND THE FUTURE OF CIVIL JUSTICE 225, 232-250 (Engstrom ed., 2023); Sternlight & Robbennolt, *supra* note 16.

20. *See, e.g.*, Sternlight & Robbennolt, *supra* note 16.

21. Engstrom & Gelbach, *supra* note 10, at 1038, 1040; David Freeman Engstrom & Nora Freeman Engstrom, *Legal Tech and the Litigation Playing Field*, *in* LEGAL TECH AND THE FUTURE OF CIVIL JUSTICE 133, 134-54 (Engstrom ed., 2023).

22. *See* Benjamin H. Barton, *Regulation, Culture, Markets: The Future of Legal Tech*, *in* LEGAL TECH AND THE FUTURE OF CIVIL JUSTICE 21, 22-43 (Engstrom ed., 2023); Benjamin H. Barton & Deborah L. Rhode, *Access to Justice and Routine Legal Services: New Technologies Meet Bar Regulators*, 70 HASTINGS L.J. 955, 958 (2019).

23. For a useful overview of pandemic-fueled court digitization, *see generally* National Center for State Courts, *Post-Pandemic Planning: Technology Resource Guide* (July 1, 2020), https://www.ncsc.org/__data/assets/pdf_file/0020/42482/Post-Pandemic-Planning.pdf (summarizing state-level initiatives around, among other things, digital divide kiosks, digital recording, electronic filing and signatures, legal assistance portals, live streaming, and remote proceedings) (hereinafter Post Pandemic Planning). *See also* notes 40–53, *infra* and accompanying text.

come.[24] Other important efforts are underway to create national data standards and facilitate interoperability among roughly 14,000 federal, state, and local court jurisdictions, many of them exercising independent control over their records and data flows.[25] The online migration is also creating entirely new data pools of digitally recorded proceedings, which were previously only memorialized, if at all, in written transcripts that obscured much of the system's workings.[26] Even the basic pandemic-fueled shift to e-filing means a larger store of primary legal materials—especially electronic versions of previously paper-only files—that can be more readily mined for insights about the system's outputs. We are emerging from the pandemic with more insight into the core workings of the civil justice system than we have ever had before—a low bar for sure, but still, significant.[27]

---

24. *See, e.g.*, National Center for State Courts, *Data Governance Policy Guide* v (Dec. 2019), https://www.courtstatistics.org/__data/assets/pdf_file/0014/23900/data-governance-final.pdf (offering an overview of data infrastructure issues, including the value of structure). Another example is the JTC's 2015 report on big data and courts, which repeatedly mentions data structure. *See generally* Joint Technology Committee, *Big Data: What State Courts Should Know* (Dec. 2014), https://www.ncsc.org/__data/assets/pdf_file/0028/17938/big-data-1-0-1-23-2015-final.pdf.

25. *See* notes 254–58, *infra* and accompanying text. The leading data standardization effort, as described in more detail below, is the National Open Court Data Standards' (NODS) project of the National Conference of State Courts. *See* https://www.ncsc.org/services-and-experts/areas-of-expertise/court-statistics/national-open-court-data-standards-nods.

26. *See* Justin Weinstein-Tull, *The Structures of Local Courts*, 106 VA. L. REV. 1031, 1044–45 (2020); Jessica E. Steinberg, Anna E. Carpenter, Colleen F. Shanahan, & Alyx Mark, *Judges and the Deregulation of Lawyers* (unpublished manuscript, 2020); Anna E. Carpenter, Jessica K. Steinberg, Colleen F. Shanahan, & Alyx Mark, *Studying the "New" Civil Judges*, 2018 WISC. L. REV. 249, 252–53 (2018); Jessica K. Steinberg, *Adversary Breakdown and Judicial Role Confusion in "Small Case" Civil Justice*, 2016 BYU L. REV. 899, 938–43; Anna E. Carpenter, *Active Judging and Access to Justice*, 93 NOTRE DAME L. REV. 647, 652 (2018); Elizabeth G. Thornburg, *The Managerial judge Goes to trial*, 44 U. RICH. L. REV. 1261, 1291 (2010).

27. *See Making Justice Accessible Initiative, Measuring Civil Justice for All: What Do We Know? What Do We Need to Know? How Can We Know It?*, AMER. ACAD. ARTS AND SCI. 11–12 (2021), https://www.amacad.org/sites/default/files/publication/downloads/2021-Measuring-Civil-Justice-for-All.pdf. For discussions of the opacity that generally dogs the civil justice system, *see* Elizabeth Chambliss, *Evidence-Based Lawyer Regulation*, 97 WASH. U. L. REV. 297, 321 (2019); Nora Freeman Engstrom, *Measuring Common Claims About Class Actions*, JOTWELL (Mar. 16, 2018). For analysis regarding that opacity's impact on access to justice, *see generally* James Greiner, *The New Legal Empiricism & Its Application to Access-to-Justice Inquiries*, 148 DAEDALUS 64 (2019); Erika J. Rickard, *The Agile Court: Evidence-Based Approaches to Improve Access to Justice and the Court User Experience*, 39 W. NEW ENG. L. REV. 227, 249–48 (2017) (advocating more empirical evaluation of the efficacy of innovations); Laura K. Abel, *Evidence-Based Access to Justice*, 13 U. PA. J.L. & SOC. CHANGE 295 (2009) (calling for a more evidence-based approach in civil legal aid programs). D. James Greiner & Andrea Matthews, *Randomized Control Trials in the United States Legal Profession*, 12 ANN. REV. L. & SOC. SCI. 295 (2016) (noting the lack of randomized control trials or other rigorous research designs in the civil justice realm); Deborah L. Rhode, *Access to Justice: An Agenda for Legal Education and Research*, 62 J. LEGAL EDUC. 531, 533 (2013) (noting the lack of empirical research).

Importantly, digitization means more than just *more* data: it also means *higher-value* data that can be put to more effective use.[28] As a result, data quality, and access to quality data, will increasingly shape the civil justice system's mechanics and outcomes. Predictions vary, however, as to whether new technologies will level or slant the litigation playing field. On the one hand, litigants with more or better data and the technical capacity to analyze it will be more able to assess case prospects and value, thus making smarter litigation decisions and capturing more of the settlement surplus at the bargaining table.[29] Legal tech may prove to be one more way that litigation's "haves" come out ahead.[30] On the other hand, data-driven legal tech tools can expand access to counsel by allowing lawyers to do more with less, making services more affordable for those with legal needs that currently go unmet.[31] Moreover, new data matched with potent new analytics will make possible new legal tech applications, from legal help chatbots to document-assembly tools to algorithmically-mediated "triage" systems. These data-fueled innovations could serve self-represented parties who must go it alone while also expanding the reach and resources of legal aid organizations.[32] Potentially more transformative are increasingly sophisticated ODR platforms, including those that arm disputants with an algorithmic, data-based outcome prediction in order to guide them toward a fair settlement without the assistance of counsel.[33] Regardless of which of these visions comes to pass, data will

28. *See* Karen Yeung, *Algorithmic Regulation: A Critical Interrogation*, 12 REGULATION AND GOVERNANCE 505, 509 (2018) ("The excitement surrounding Big Data is rooted in its capacity to identify patterns and correlations that cannot be detected by human cognition, converting massive volumes of data (often in unstructured form) into a particular, highly data-intensive form of knowledge, and thus creating a new mode of knowledge production."); *see also* Pew Charitable Trusts, *How States Use Data to Inform Decisions: A National Review of the Use of Administrative Data to Improve State Decision-Making* 1 (2018) (noting increasing use of "administrative data" that was previously low-value to make "strategic, data-informed decisions about how "to implement and oversee a program"). One way of thinking about this is that ML has allowed public section automation efforts to move into a set of unstructured and semi-structured problems, sometimes called "gray zones," previously thought to be insufficiently tractable to be susceptible of quantification, reduction, or encoding in automated systems. *See* Michael Veale & Irina Bass, *Administration by Algorithm: Public Management Meets Public Sector Machine Learning*, in KAREN YEUNG & MARTIN LODGE, EDS., ALGORITHMIC REGULATION (2019).

29. *See* Engstrom & Gelbach, *supra* note 10, at 1039.

30. *See generally* Engstrom & Engstrom, *supra* note 21; Engstrom & Gelbach, *supra* note 10, at 1039.

31. *See* Engstrom & Gelbach, *supra* note 10, at 1038.

32. *See* notes 65–72, *infra* and accompanying text. For an introduction to "document assembly" and its relationship to access to justice efforts, *see* Claudia Johnson, *Resource: Document Assembly: An Essential Building Block for the Access to Justice Ecosystem* (LHI 2016), SELF-REPRESENTED LITIGATION NETWORK (2016), https://www.srln.org/node/848/document-assembly-essential-building-block-access-justice-ecosystem-news-2016.

33. *See* notes 65–70 *infra* and accompanying text.

steadily move to the heart of the civil justice system, becoming its lifeblood.

The second argument follows directly from the first: Because data will be increasingly central to civil justice, some of our courts' most consequential post-pandemic responsibilities will lie in data governance.[34] Navigating the opportunities presented by a newly digitized system, while avoiding its many perils, will require careful regulation to ensure that the resulting data are distributed and used in fair, equitable, secure, and privacy-protecting ways.[35] For better or worse, the difficult job of crafting and enforcing those all-important regulations will largely fall to the nation's courts. Some of their most important work will no longer center on providing litigants their day in court, whether in-person or online. Instead, it will revolve around data.

To make all of this more concrete, post-pandemic courts will play at least three distinct data governance roles as digitization deepens throughout the civil justice system. First, courts will be *users* of data when they design and oversee new data-based tools, including court-linked legal help chatbots and ODR systems.[36] Second, courts will be *dispensers* of data when they collect the mountains of data generated by the legal system and set the terms on which that data is made available to outside actors who seek to use it.[37] Third, courts will be *regulators* of data—particularly others' *use* of data—as they determine which software providers can, or cannot, provide legal services consistent with existing lawyer regulation and the rules of professional responsibility.[38]

In performing these new governance roles, courts will make innumerable choices about how to reach and communicate decisions regarding the collection, use, and disposal of data. They will quickly

34. We define data governance as rules, regulations, and practices that shape the application of analytics tools to data, including the hardware and other systems in which those processes are housed. I do not, contrary to some, distinguish between data governance functions and IT functions, as the line between the two can be blurry. *See e.g., Data Governance Policy Guide*, *supra* note 24, at 7 (distinguishing between IT functions, including maintenance of hardware, systems software, networks, and facilities, and data governance, including logical data modeling, data categorization, data "access/sharing," and data "quality/integrity").

35. We are not the first to note that courts and, in particular, chief judges will play a central role in reform efforts. *See, e.g.*, Jonathan Lippman, *The Judiciary As the Leader of the Access-to-Justice Revolution*, 89 N.Y.U. L. Rev. 1569, 1569–70 (2014) (discussing judicial leadership on access to justice issues); Richard Zorza, *Access to Justice: The Emerging Consensus and Some Questions and Implications*, 94 Judicature 156, 156–57 (2011) (same); Gerry Singsen, *Observing Change*, 3, 6–7 MSBF.ORG (same), http://www.msbf.org/futuresandatj/Observing-Change-article.pdf.

36. *See* Part I, *infra*.

37. *See* Part II, *infra*.

38. *See* Part III, *infra*.

come to see data as a strategic asset, not just a by-product of case processing or court management.[39] Courts will also face pervasive make-or-buy questions: how much of their own internal technical capacity to build and rely upon and how much to contract out to outside vendors? Additional choices will include deciding how to standardize data formats across jurisdictions, as well as how that data is accessed and by whom. How well courts perform these new and uncharted governance roles—as data users, data dispensers, and data (and data-use) regulators—will, perhaps more than any other single force, shape the future of the American civil justice system in the post-pandemic era.

The remainder of this Article tours the challenges inherent in the courts' new data governance roles and offers some initial, tentative thoughts on how best to meet them. Part I explores courts as data users. Part II examines courts as data curators and dispensers. Part III considers courts as data (and data use) regulators. A brief conclusion identifies themes cutting across the trio of data governance roles and plots some fruitful avenues for further inquiry.

I. Courts as Data Users

Of all the ways courts will be thrust into the role of data governors, perhaps the most direct is that courts will increasingly use data themselves to provide litigants with digital legal services and to automate the process of managing and adjudicating cases. Doing so will put courts squarely in the role of data *users*—and, by extension, digital system designers.

A. Automating Court Services: From Portals and Chatbots to ODR

A range of data-based, court-centered projects are currently underway,[40] but two use cases seem likely to be most impactful over the near- to medium-term. First, a cluster of initiatives creates sorting systems to guide people with civil justice needs into appropriate legal

39. This language is a gloss on the National Center for State Court's excellent *Data Governance Policy Guide*, *supra* note 24, at v, which defines data governance as "[a] framework by which courts reach and communicate organizational decisions around data, ensure that business activities and data management are synchronized, and develop and document long- and short-term strategies around the collection, use, and disposal of data."

40. *See generally* National Center for State Courts, *Post-Pandemic Planning: Technology Resource Guide* (July 1, 2020) (providing an expansive overview of initiatives, with concrete examples by state); Rebecca Love Kourlis & Riyaz Samnani, *Mapping the Future of User Access Through Technology* (2017), https://iaals.du.edu/sites/default/files/documents/publications/court_compass_mapping_the_future.pdf (providing a more succinct overview).

help tracks.[41] The result is an increasingly digitized version of the courthouse self-help centers, help desks, and facilitator offices, often staffed by volunteer and non-profit-funded lawyers, that courts created back in analog times to address access concerns.[42] Many jurisdictions, for instance, are actively developing and deploying legal help chatbots.[43] Arizona's "Gavel" chatbot,[44] New Jersey's Judicial Information Assistant, and Mississippi's Lex[45] are examples.[46] Another is LACourtConnect, which provides answers to a pre-programmed set of questions and also allows users to submit a question to the system's "knowledge base" as a suggestion for future inclusion.[47] For the moment, these chatbots provide low-level information, typically about court processes. With time, they will surely grow more sophisticated.

A related set of tools aspires to something closer to soup-to-nuts litigation assistance: "litigant portals" that offer a range of services to litigants, particularly self-represented ones who must navigate the civil justice system alone.[48] The Michigan courts are currently collaborating with Michigan Legal Help, a non-profit organization, to provide self-help tools across numerous civil justice areas—ranging from family law, to housing law, to immigration law—that route litigants into help channels and provide general legal information and document assembly services.[49] Numerous other states beyond Michigan have created landing pages linking SRLs to legal tech for creating and filing

41. For overviews of this family of initiatives, *see* Schmitz, *supra* note 2 at 92–93 (2019); J.J. Prescott, *Improving Access to Justice in State Courts with Platform Technology*, 70 VAND. L. REV. 1993, 1999 (2017). For descriptions at a somewhat earlier stage in development, *see* James E. Cabral et al., *Using Technology to Enhance Access to Justice*, 26 HARV. J.L. & TECH. 241, 259–60 (2012); Michael J. Wolf, *Collaborative Technology Improves Access to Justice*, N.Y.U. J. LEGIS. & PUB. POL'Y 759, 771–75 (2012) (describing new online technologies that improve court accessibility for low- and moderate-income litigants).

42. *See* Quinten Steenhuis & David Colarusso, *Digital Curb Cuts: Towards an Inclusive Open Forms Ecosystem*, 54 AKRON L. REV. 773, 777 (2020) ("Legal aid organizations, court service centers, and community organizations have a long history of producing written materials, including books and websites, that provide legal information and guidance."); *see generally* Legal Services Corporation, *About the Statewide Website Assessment Public Report*, https://lsc-live.app.box.com/s/zjv5cjuos1kbp43yhicc6pnhccjtboew (July 2017) (summarizing past efforts of state-specific legal aid websites).

43. For an overview, *see* Joint Technology Committee, *Getting Started with a Chatbot* (Apr. 2020), https://www.ncsc.org/__data/assets/pdf_file/0028/28567/2020-04-15-qr-getting-started-with-a-chatbot.pdf.

44. *See* ARIZONA JUDICIAL BRANCH, https://www.azcourts.gov/.

45. *See* JUSTICE COURT ACCESS PROGRAM, https://msjusticecourthelp.com/.

46. *See* NEW JERSEY COURTS, https://www.njcourts.gov/.

47. *See* TRAFFIC DIVISION – LA COURT, http://www.lacourt.org/division/traffic/traffic2.aspx.

48. *See* Thomas M. Clarke, *Building A Litigant Portal: Business and Technical Requirements,* STATE JUSTICE INST. & NAT'L CTR. FOR STATE COURTS 4 (Nov. 2015), https://ncsc.contentdm.oclc.org/digital/collection/accessfair/id/375 (describing portal idea).

49. *See* WELCOME TO MICHIGAN LEGAL HELP, https://michiganlegalhelp.org/.

legal documents and forms.[50] To be sure, many of these tools can seem downright primitive compared to cutting-edge uses of AI in medicine, finance, or transportation. At last count, only three states' courts—in New York, California, and Minnesota—boast integrated document assembly and e-filing that allow self-represented litigants (SRLs) to seamlessly generate needed legal documents and then transmit them to court.[51] Nor are most of these tools integrated into the case management systems of courts, legal aid programs, and other nonprofits.[52] Still, the dream of many access to justice advocates seems well within reach: court-connected portals that walk individuals with justice needs through plain-language interviews and then produce the specific tools or services they need to weigh legal options and file necessary court papers.[53] We return to these initiatives in Part II, as many comprise public/private collaborations that rely on non-profit partners and for-profit court vendors.[54] As a result, they implicate the emerging role of courts as data *curators* and *dispensers*, not just *users*.

The second use case that will increasingly press courts into the role of data users is ODR.[55] As already noted, ODR is steadily gaining

50. *See e.g.*, Illinois Courts Divorce, Child Support and Maintenance Page, Dissolution of Marriage/Civil Union (Divorce With Children), https://www.illinoiscourts.gov/forms/approved-forms/forms-circuit-court/divorce-child-support-maintenance (linking to "a guided interview that will ask you a series of questions related to this topic and then the program will complete the forms for you. It is free to use"); Supreme Court of Ohio and Ohio Judicial System, "Access to Justice Resources," https://www.supremecourt.ohio.gov/JCS/courtSvcs/justiceAccess/resources/ (linking to interactive self-help page with form-filling tools created by Ohio Legal Help, a non-profit nonlegal services provider).

51. *See* Johnson, *supra* note 32. For specific examples, *see* DIY Forms, https://nycourts.gov/courthelp//DIY/index.shtml; Minnesota Judicial Branch – Guide and File, https://mncourts.gov/guide-and-file; Change your Name in California, https://www.courts.ca.gov/selfhelp-namechange.htm.

52. *Id*.

53. Rebecca Love Kourlis, Natalie Anne Knowlton, & Logal Cornett, IIALS, *A Court Compass for Litigants* 4 (2016), https://iaals.du.edu/sites/default/files/documents/publications/court_compass_convening_report.pdf; Clarke, *supra* note 48 (describing a litigant portal that offers substantial assistance, from "help[ing] potential litigants identify legal problems" to "help[ing] litigants execute desired legal actions"). A concrete, high-profile example is a collaboration between the Legal Services Corporation (LSC) and Microsoft to provide statewide triage portals that elicit information from users, connect them to needed resources, and move them toward action, whether through referral to a legal service provider or document assembly and e-filing. *See* Legal Services Corp., *LSC Moves Forward with Legal Navigator Project* (2017), https://www.lsc.gov/media/highlights/simplifying-legal-help.

54. Among the non-profit providers is LawHelp Interactive, a project of ProBono Net. *See* Law Help Interactive, https://lawhelpinteractive.org/. Chief among the for-profit providers is Tyler Tech, which claims that its Odyssey Guide and File system is available in 17 states. *See* Odyssey Guide & File, http://www.guideandfile.com/.

55. For an impressive and regularly updated bibliography on all things ODR, *see* The National Center for Technology and Dispute Resolution, *Publications* (last visited Jan. 17, 2023), http://odr.info/publications/.

momentum, fueled by growing alarm about the abysmal state of access to justice and the emergence of multiple software vendors vying for market share. In addition, an expanding portfolio of successful pilots — among states like Michigan,[56] Ohio,[57] and Utah,[58] plus globally, from Canada[59] to the U.K. and beyond[60] — have spurred ODR growth.[61] ODR penetration has deepened substantially in the pandemic, and there is good reason to expect continued growth.[62]

ODR is also rapidly advancing in capability. For the moment, most court-linked ODR platforms remain virtual gathering places where disputants can engage, typically asynchronously, and bargain their way to settlement without costly trips to court—"pajama courts," as some call them.[63] These systems are relatively straightforward technologically, providing a 24/7 forum and, in some instances, access to human facilitators who help disputants classify their problems and then inform them about prospective options.[64] To that extent, existing ODR systems are part bargaining space, part legal help desk.

With data and analytic capacity growing fast, however, technologists are designing what might be called "ODR 2.0," which automates both tasks. These ODR systems incorporate a variety of algorithmic tools that prime the parties with case-relevant information without requir-

56. *See generally* Office of Dispute Resolution, *Resolve a Dispute Online with MI-Resolve*, MICH. CTS., https://www.courts.michigan.gov/administration/offices/office-of-dispute-resolution/mi-resolve/.

57. *See generally* Prescott & Sanchez, *supra* note 17.

58. Paula Hannaford-Agor, Kathryn J. Genthon, Susanne Mitchell, & Divya Mathew, *Impact of the Utah Online Dispute Resolution (ODR) Pilot Program: Final Report* ii (Dec. 10, 2020), https://ncsc.contentdm.oclc.org/digital/collection/adr/id/66/.

59. *See* BC CIVIL RESOLUTION TRIBUNAL, https://civilresolutionbc.ca/; *see* ETHAN KATSH & ORNA RABINOVICH-EINY, DIGITAL JUSTICE: TECHNOLOGY AND THE INTERNET OF DISPUTES 151 (2017).

60. *See* SUSSKIND, *supra* note 1, at 166.

61. *Id*.

62. *See* Michelle Casady, *Texas Judges See Lasting Benefits From Pandemic Practices*, LAW360 (Mar. 11, 2021, 9:30 AM), https://www.law360.com/articles/1362923/; Scott Dodson et al., *The Zooming of Federal Civil Litigation*, JUDICATURE, Fall 2021, at 12, 14–15 (predicting that "[s]ome categories of adversarial events . . . are likely to migrate permanently to online platforms," including discovery conferences, oral hearings, and possibly appellate arguments). *See also* David Freeman Engstrom, *Post-COVID Courts*, 64 UCLA L. REV. DISCOURSE, 246, 248 (2020) (arguing that the uptake of "new technologies that will transform how legal work is done" will accelerate post-pandemic).

63. Claire Osborn & Taylor Goldenstein, *Area Judges Make Plans to Try Out "Pajama" Court*, STATESMAN NEWS NETWORK (June 18, 2018), https://www.statesman.com/news/20180618/area-judges-make-plans-to-try-out-pajama-court. *See also* R.J. Vogt, *Can Online 'Pajama Courts' Reshape Justice?* LAW360 (May 12, 2019), https://www.law360.com/articles/1158405.

64. *See* SUSSKIND, *supra* note 1, at 153 (noting "tools and methods to help lay people organize and classify their cases (turning a grievance into a justiciable problem) and to analyze and reason (coming to a legal view)").

ing flesh-and-blood dispute handlers.[65] The most basic versions, long deployed in private ODR systems such as eBay, feature blind bidding to find overlap in the parties' reservation prices.[66] Slightly more complex versions draw on the confidential party preferences or similar prior cases in order to present disputants with optimized settlement "packages" for the issues in controversy.[67] A still more sophisticated version incorporates a predictive analytics engine to arm disputants with a richer array of information about their case.[68] That might be an outcome prediction based on where past disputants like them came out in the bargaining process on that same platform.[69] Or it might be a full-on prediction as to how a court will ultimately rule if the case were litigated to a judgment—*i.e.*, a best alternative to a negotiated agreement (BATNA)—so the disputants can bargain in the shadow of what a court is likely to decide.[70]

---

65. *See* SUSSKIND, *supra* note 1, at 6; KATSH & RABINOVICH-EINY, *supra* note 1, at 36–38, 46 (noting ODR's shift from "a process that simply facilitates communication of information to one that processes it"). *See also* John Zeleznikow, *Can Artificial Intelligence and Online Dispute Resolution Enhance Efficiency and Effectiveness in Courts*, 8 INTERNAT'L J. CT. ADMIN. 30, 30 (2017); Darin Thompson, *Creating New Pathways to Justice Using Simple Artificial Intelligence and Online Dispute Resolution*, 1 INTERNAT'L J. ONLINE DISP. RES. 4, 4 (2015).

66. This approach, pioneered by Cybersettle in 1998 and now a dominant method, requires the claimant and defendant to submit their highest and lowest settlement numbers in search of overlap. *See* Diane J. Levin, *Cybersettle Makes the Case for Resolving Disputes Online*, MEDIATION CHANNEL (Feb. 20, 2008), https://mediationchannel.com/2008/02/20/cybersettle-makes-the-casefor-resolving-disputes-online/.

67. Ernest Thiessen, Paul Miniato, & Bruce Hiebert, *ODR and eNegotiation, in* ONLINE DISPUTE RESOLUTION: THEORY AND PRACTICE, A TREATISE ON TECHNOLOGY AND DISPUTE RESOLUTION 329, 333; KATSH & RABINOVICH-EINY, *supra* note 1, at 35–36, 49 (2020) (describing system and noting that the "software examines the way in which the parties ranked their interests and analyzes whether at least one of the parties' interests can be better met without making the other party worse off. If there is an alternative solution, the parties are presented with it; they can then either choose the proposed agreement offered by the software or remain with the resolution they originally negotiated."); Arno R. Lodder & John Zeleznikow, *Artificial Intelligence and Online Dispute Resolution*, *in* ONLINE DISPUTE RESOLUTION: THEORY AND PRACTICE, A TREATISE ON TECHNOLOGY AND DISPUTE RESOLUTION 73, 77 (2020) (describing a game-theory-based ODR system that has disputants "rank and value each issue in dispute, by allocating the sum of one hundred points amongst all the issues" and then uses the information to propose a settlement package).

68. *See* Zeleznikow, *supra* note 65, at 39.

69. *See* Thiessen, Miniato, & Hiebert, *supra* note 67, at 329, 333; *see also* KATSH & RABINOVICH-EINY, *supra* note 1, at 35–36, 49 (describing system).

70. *See* Zeleznikow, *supra* note 65, at 39; Lodder & John Zeleznikow, *supra* note 67, at 80–81. A more colloquial term for the BATNA concept is "bargaining in the shadow of the law." *See generally* Robert Mnookin & Lewis Kornhauser, *Bargaining in the Shadow of the Law: The Case of Divorce*, 88 YALE L.J. 950 (1979).

B. How Digitizing Could Impact Access to Justice

The great power and promise of court uses of data is that they might serve as a solution, albeit a partial one, to the access to justice crisis. This is perhaps easiest to see with ODR 2.0. Automated information about case prospects can, after all, fill some or even much of the informational role that would otherwise be filled by a lawyer, providing an automated, non-human bridge from legal claim to resolution or remedy.[71]

This lawyer-substituting aspect of the more advanced forms of ODR also makes clear why ODR, if it is to be a fixture of the civil justice landscape, *must* come from courts, at least under current legal ethical rules governing the unauthorized practice of law (UPL).[72] The automated document assembly tools noted above are sometimes embedded in court-hosted litigant portals, but they can also take the form of standalone, privately offered services by for-profit and nonprofit companies like LegalZoom, RocketLawyer, and Upsolve, among many others, or non-profit groups such as ProBono Net, which operates LawHelp Interactive.[73] Standalone document assembly service providers tip-toe around UPL restrictions, and thus avoid regulatory scrutiny, by incorporating regular lawyer review to ensure that the forms provided are up-to-date and by scrupulously avoiding anything resembling legal advice.[74] In stark contrast, a standalone, private version of ODR 2.0 that primed self-represented litigants with information about case merits would quickly draw challenges on UPL grounds. Absent a sea-change in the rules of professional responsibil-

71. SUSSKIND, *supra* note 1, at 298.

72. Drew Simshaw, *Ethical Issues in Robo-Lawyering: The Need for Guidance on Developing and Using Artificial Intelligence in the Practice of Law*, 70 HASTINGS L.J. 173, 178–79 (2018) (noting how AI may constitute unauthorized practice of law even as AI in ODR pilots appears to increase access to justice).

73. For an overview, *see* Rebecca Sandefur, *Legal Tech for Non-Lawyers: Report of the Survey of US Legal Technologies*, AM. BAR FOUND. (2019), http://www.americanbarfoundation.org/uploads/cms/documents/report_us_digital_legal_tech_for_nonlawyers.pdf.

74. This was the essence of the settlement that LegalZoom entered into with the North Carolina bar by way of a consent agreement. *See* LegalZoom.com, Inc. v. N.C. State Bar, No. 11 CVS 15111, 2015 WL 6441853, at *1 (N.C. Super. Ct. Oct. 22, 2015). In August 2019, the American Bar Association adopted the best practices for online legal document providers that included similar constraints. *See generally Resolution 10A: ABA Best Practice Guidelines for Online Legal Document Providers,* ABA HOUSE OF DELEGATES (Aug. 12–13, 2019), https://www.americanbar.org/content/dam/aba/directories/policy/annual-2019/10a-annual-2019.pdf. For a concise overview of UPL challenges in this area, *see id.* at 2–5 (noting pattern of UPL challenges, only some successful, following by legislative overrides). Part III considers how court data governance can spur and stymie private-sector development of these tools by affecting their scalability. *Id.* at 5–7. And Part IV considers the possibility that a relaxation of lawyer rules, from UPL to the bar on nonlawyer ownership of firms and fee-splitting, may be in the offing, thus welcoming more software-based legal services providers into the system. *Id.* at 7–8.

ity, ODR 2.0 will not gain traction outside of the formal court system, making its development by courts imperative to its future growth.[75]

Importantly, ODR will not just *use* data. It will also *generate* it. Long-time ODR evangelist Richard Susskind is especially lucid on this point. The standard knock on ODR is its opacity—both the "black box" algorithmic sort, but also the simple fact that ODR happens behind closed (virtual) doors. ODR thus threatens to erode the public elaboration of legal norms, as others have noted with analog forms of alternative dispute resolution and the wider settlement-focused civil litigation system.[76] But ODR, by virtue of being an online platform, makes data easier to collect, use, and publish than in analog systems. As Susskind puts it, conventional courts have always had a high level of "real-time transparency," but a low level of "information transparency."[77] ODR's continued proliferation could flip this state of affairs. Indeed, the increased flow of case-related data could leave courts better off than before, so long as the scrutability lost when rules are embedded in code is offset by gains in more actionable, system-wide information.[78]

---

75. Maximilian A. Bulinski & J.J. Prescott, *Online Case Resolution Systems: Enhancing Access, Fairness, Accuracy, and Efficiency*, 21 MICH. J. RACE & L. 205, 221 (2016) ("Importantly, we do not imagine these systems as providing legal advice as a lawyer might; rather, the software would empower the court—i.e., the judge—to communicate more clearly with citizens about the law, their rights, and the consequences of exercising certain options.").

76. *See* Carrie Menkel-Meadow, *Whose Dispute Is It Anyway?: A Philosophical and Democratic Defense of Settlement (in Some Cases)*, 83 GEO. L.J. 2663, 2663–64 (1995) (reviewing the merits and demerits of settlement and ADR); J. Maria Glover, *The Federal Rules of Civil Settlement*, 87 N.Y.U. L. REV. 1713, 1718–20 (2012) (canvassing the ways civil procedure is designed to facilitate settlement).

77. *See* SUSSKIND, *supra* note 1, at 194–95.

78. For instance, many courts currently lack basic information about the contours of debt collection lawsuits despite the fact that they are the most numerous type of case in many jurisdictions. *See* Erika Rickard & Qudsiya Naqui, *Effects of Debt Lawsuits on Civil Courts – and Consumers – Obscured by Lack of Data*, PEW CHARITABLE TRUSTS (June 5, 2020), https://www.pewtrusts.org/en/research-and-analysis/articles/2020/06/05/effects-of-debt-lawsuits-on-civil-courts-and-consumers-obscured-by-lack-of-data ("In 2018, only New Mexico and Texas reported a cross-section of debt claims cases and outcomes for at least one court level.") With ODR, we might get closer to real information about default judgment rates, the percent of cases that reach settlement, the percent of cases that end in a consumer's favor (*i.e.*, dismissal), settlement amounts, relative to the amount initially sought, the rate of satisfaction of judgments or settlements; and rates of wage or bank account garnishment. For more "case" versus "system" transparency, *see id*. at 199 (suggesting that increases in "information transparency" might offset reductions in "'real-time' transparency"). For a jurisprudential version of the argument, *see generally* Brian Sheppard, *Warming Up to Inscrutability: How Technology Could Challenge Our Concept of Law*, 68 Univ. TORONTO L.J. 36 (2018). A nice way to capture some of the same ideas is to note that ODR is a *system* rather than a *tool.* KATSH & RABINOVICH-EINY, *supra* note 1, at 35, 52 (noting how ODR "lift[s] the onus" of obtaining justice from individual to entity); *id*. at 163 ("As courts rely on digital technology and ODR systems, they will learn to view data as a central feature in dispute resolution.").

ODR thus offers a window into some of the wider ways courts will use new data streams to self-scrutinize—that is, to spot equity issues, improve overall performance, and better tailor the adjudicative services they provide.[79] Most obviously, ODR-generated data will inform continuous process improvement of ODR systems themselves. But that same data might also be used in the service of wider reforms, whether in courts or other branches of government, aimed at preventing disputes from arising in the first place.[80] An especially interesting possibility that judges and advocates are beginning to think about—and long a dream of A2J advocates—is "federating" (*i.e.*, linking) civil justice data with data elsewhere in the government, particularly social welfare agencies.[81] Federating data could facilitate targeted public assistance to help individuals avoid entering the civil justice system in the first place.[82] Data-based transparency can also shape policies designed to reduce a litigant's resort to other public institutions or public programs *after* engagement with the civil justice system.[83]

Court use of data may be the wave of the future, but it will not be uncontroversial. An argument can be made that data federation and the sorts of system-level inquiries it permits will erode conventional notions of judicial autonomy and independence.[84] There is a reason why our theories of justice, and both our substantive and procedural law, favor *ex post* intervention.[85] Traditionally, courts intervene after

79. *See generally* Tanina Rostain & Amy O'Hara, *The Civil Justice Data Gap*, in LEGAL TECH AND THE FUTURE OF CIVIL JUSTICE 368, 369 (Engstrom ed., 2023); *see also* Naomi Burstyner, Tania Sourdin, Chinthaka Liyanage, Bahadorreza Ofoghi, John & Zeleznikow, *Using Technology to Discover More About the Justice System*, 44 RUTGERS COMPUTER & TECH. L.J. 1, 3–4 (2018).

80. SUSSKIND, *supra* note 1, at 194–95; *see also* Rebecca A. Johnson & Tanina Rostain, *Tool for Surveillance or Spotlight on Inequality? Big Data and the Law*, 16 ANN. REV. LAW. SOC. SCI. 453, 466 (2020) (noting use of big data and computational harms to predict harms, potentially obviating the need for legal process).

81. *See* SUSSKIND, *supra* note 1, at 194–95; Johnson & Rostain, *supra* note 80, at 466. *See* AM. ACAD. ARTS & SCI., MEASURING CIVIL JUSTICE FOR ALL 3 (2021), https://www.amacad.org/sites/default/files/publication/downloads/2021-Measuring-Civil-Justice-for-All.pdf ("In addition to understanding how courts operate, researchers are interested in linking civil justice data with other data sets to investigate the economic, demographic, and social antecedents of civil justice involvement and its downstream consequences for health, housing security, education, and economic security.").

82. *See* SUSSKIND, *supra* note 1, at 194–95; Johnson & Rostain, *supra* note 80, at 466; *See* AM. ACAD. ARTS & SCI., MEASURING CIVIL JUSTICE FOR ALL 20–21 (2021).

83. *Id.*

84. *See, e.g.*, Judith Resnik, *Managerial Judges*, 96 HARV. L. REV. 374, 382 (discussing the traditional model of detached, impartial judges).

85. *See* Spaulding, *supra* note 17, at 281–84. For instance, requests for ex ante intervention must meet higher procedural and substantive requirements. Thus, courts will not grant provisional injunctive relief unless the remedy at law is inadequate. Under the First Amendment, prior restraints are presumptively invalid.

specific wrongdoing, not before, and in response to specific acts, not the collective social and economic conditions that help produce them.[86] As a result, the main selling point of data federation—that better data-based transparency can provide a more synoptic understanding of the upstream determinants and downstream consequences of civil legal problems—stands in significant tension with liberal democratic notions of justice. Judicial efforts to tap new data streams risks bringing courts more fully into the realm of policymaking, embroiling them in divisive social policy debates while eroding the legitimacy benefits that come from the courts' traditional commitment to blind justice.

C. Digital Court Services and the Judicial Governance Challenge

Given these complexities, realizing the considerable promise of court-linked, data-intensive reforms will require courts to surmount numerous governance and oversight challenges. Here again, ODR platforms offer an excellent case study for examining the most salient challenges and concerns.

A threshold challenge of designing and implementing ODR platforms will be straightforward procedure-making. As one of us has written elsewhere, new virtual fora demand "traffic rules" that determine which cases move online and which remain in-person.[87] This will not be easy or uncontroversial. Questions about which types of cases get pushed online—for instance, routing eviction or consumer debt cases to ODR platforms, but not other types of cases—raise deeply political judgments about case value, case complexity, and litigant types.[88] Many agree that ODR holds the most promise for high-volume, low-complexity, low-stakes cases.[89] Fewer agree whether evictions and consumer debt cases meet those criteria, particularly from the perspective of those caught up in them.[90] No less difficult or politically fraught are the precise routing rules: What mix of party consent, judicial discretion, or no-flex rules make the most sense for moving

86. *Id*.

87. *See* David Freeman Engstrom, *Digital Civil Procedure*, 169 U. PA. L. REV. 2177, 2196 (2021).

88. *See* Spaulding, *supra* note 17, at 282–84.

89. *See* Schmitz, *supra* note 2, at 155 (noting that complex business cases are likely not well situated for ODR).

90. *Compare* J.J. Prescott, *Improving Access to Justice in State Courts with Platform Technology*, 70 VAND. L. REV. 1993, 2001–06 (2017) (noting ODR's promise for "minor" disputes where the cost of adjudication makes case resolution suboptimal) *with* Spaulding, *supra* note 17. A related critique is that many of them are focused on enforcement, particularly tax and traffic enforcement, and so are better seen as an extension of the government's enforcement apparatus, not a litigant-autonomy-promoting means of adjudication.

litigants into ODR?[91] As just one example of the fraught policy concerns baked into that choice among routing rules, one might worry that party consent, if made the keystone of the system, could become a bargaining chip.[92] Litigation's "haves" could withhold consent to move the dispute online against a litigant who could not afford higher-cost, in-person proceedings, or they could condition consent on concessions on discovery or other matters.[93]

However complex these questions might be, the "information rules" at the heart of ODR 2.0 are downright daunting. What information must litigants on an ODR platform provide, and what information do they get in return? As already noted, a top-line decision to be made when designing an ODR 2.0 system is whether to arm litigants with an automated case-outcome prediction or other information about likely outcomes.[94] But that is only the beginning of the types of legal information that the system might dispense. A recent report of the National Consumer Law Center contemplates a richer set of informational goods:

> ODR platforms should be designed in conjunction with information portals to provide relevant legal information using an interactive interface to help consumers identify, understand, and raise potential defenses and counterclaims. Systems can suggest certain language if a particular defense is present or even discontinue the ODR process entirely. Systems should affirmatively screen for things, like exempt income and prior bankruptcy filings. . . .[95]

That said, more information may not always be better. A major fear expressed by critics of ADR when it gathered steam in the 1980s and 1990s was the creation of overly "directive" systems that would unduly shrink the space within which parties could fashion creative solutions.[96] Priming the disputants with too much information, the thinking went, could be counter-productive. The response was that the marginalized would often do worse in more open-ended, informal, un-

---

91. William H.J. Hubbard & Ronen Avraham, *The Spectrum of Procedural Flexibility*, 87 U. CHI. L. REV. 883, 883 (2020).

92. *See* Engstrom, *Digital Civil Procedure*, *supra* note 87, at 2201.

93. *Id*.

94. *See supra* notes 65–70 and accompanying text.

95. National Consumer Law Center, *Consumer Protection and Court-Sponsored Online Dispute Resolution in Collection Lawsuits* (June 2021), https://www.nclc.org/images/pdf/debt_collection/ib-odr-july2019.pdf; *see also* David Allen Larson, *Designing and Implementing A State Court ODR System: From Disappointment to Celebration*, 2019 J. DISP. RESOL. 77, 80 (2019) (describing New York's recent process for designing an ODR system for consumer credit disputes that includes information about available legal defenses).

96. For an excellent overview, *see* Yishai Boyarin, *Court-Connected ADR—A Time of Crisis, A Time of Change*, 95 MARQ. L. REV. 993, 993, 1007–10 (2012).

structured adjudicatory contexts.[97] A related concern is that some litigants might be better able than others to use information to their advantage. To compensate, how much should the system incorporate "reality checks" that help unsophisticated litigants grasp the real cost over time of a resolution?[98] Should the system permit cancelation of an agreement within, say, twenty-four hours, to guard against unwise resolutions taken under duress or without adequate understanding? These are only a few of the questions to be answered—some old, some new—as ODR continues its move into courts.

Information rules also bring more technical complexities than traffic rules. For starters, coding any of the above features into an ODR platform demands substantial technical know-how. It requires a data-based understanding of what information might prove most useful—for instance, which legal defenses most often go unasserted, or which types of income or assets are exempt from collection. That information will be dynamic and will require frequent updating as statutes, regulations, and case law change. Moreover, prudent design and implementation of more advanced ODR systems will require determining how best to extract relevant information from available data and which user interface most effectively presents that information to the disputants. On the latter, academic work on ODR "usability" has only just begun,[99] leaving open questions about technical design that courts will have to confront.

In addition to the traffic and information rule challenges, optimal design of ODR 2.0 platforms brings numerous data management hurdles.[100] As an example, if ODR systems are to produce ready-made, structured data for evaluation and planning purposes, the ODR pro-

97. *Id*.; *See also* David Allen Larson, *Designing and Implementing a State Court ODR System: From Disappointment to Celebration*, 2019 J. Disp. Resol. 77, 92 (2019) ("Power imbalances between parties can be exploited in virtual spaces. Therefore, the system only should permit structured negotiations rather than unmonitored direct communications between the parties.").

98. *See* Orna Rabinovich-Einy, *Going Public: Diminishing Privacy in Dispute Resolution in the Internet Age*, 7 Va. J.L. & Tech. 4, 62 (2002) (noting use of reality checks by human mediators).

99. *See generally* Stacy Butler, Sarah Mauet, Christopher L. Griffin, Jr., & Mackenzie S. Pish, *The Utah Online Dispute Resolution Platform: A Usability Evaluation and Report* (Sept. 8, 2020), https://law.arizona.edu/sites/default/files/i4J_Utah_ODR_Report.pdf; *see also* Ayelet Sela, *E-Nudging Justice: The Role of Digital Choice Architecture in Online Courts*, 2019 J. Disp. Resol. 127 (2019); David Allen Larson, *Designing and Implementing a State Court ODR System: From Disappointment to Celebration*, 2019 J. Disp. Resol. 77 (2019).

100. For a common articulation of data governance and data management challenges, *see* Sean La Roque-Doherty, *Toward Smarter Courts Artificial Intelligence Has Made Great Inroads-but Not As Far As Increasing Access to Civil Justice*, ABA J. at 20, 21 (Apr./May 2021) ("The difficulty with data includes the insufficient size of available data sets, the absence of data standards, data integration, and data privacy and security.").

cess must fit into a court's existing case management system.[101] Compounding the problem is the likelihood of low data quantity. It is well-known that courts will never have "big data," at least not like many private sector entities, with their internet-scale data collection capabilities. Think millions or even billions of Facebook clicks every minute. Instead, the challenge for courts is "[s]mall data, done well."[102]

For courts designing ODR 2.0 systems, the answer to small data might be integration across jurisdictions or within jurisdictions to expand the data pool. But integration presents its own set of challenges. The first is a lack of common data standards—that is, rules under which data is stored and formatted.[103] Another is the fact that, while many state courts are "unified" systems with centralized control over filing, case management, and data, many are not. These realities make it harder to create the metadata schema—a common vocabulary for collecting structured information—necessary to facilitate rigorous comparisons across jurisdictions.[104] This difficulty, in turn, impairs courts' ability to self-scrutinize and improve their own policies and practices.[105] With comparable data, one can construct rigorous research designs that generate reliable inferences about the impact of policy choices made in some jurisdictions but not others. Without that ability, a court cannot compare, say, its own legal-help portal approach and design choices to choices made in other states. The vaunted "laboratories of democracy" of American federalism, always less than lab-like from a purely social scientific perspective,[106] lose

101. Larson, *supra* note 97, at 81 ("One of the difficulties with designing a court–integrated ODR process is that the ODR process must be seamlessly integrated into the existing court management system.").

102. *See* Joint Technology Committee, *Big Data: What State Courts Should Know* (Dec. 2014), https://www.ncsc.org/__data/assets/pdf_file/0028/17938/big-data-1-0-1-23-2015-final.pdf ("State courts do not have Big Data. State courts do have small amounts of highly structured data, which can be organized to support specific business processes and services so that court cases can be resolved in a fair, impartial, and timely manner.").

103. *See* notes 235–38, *infra* and accompanying text (describing efforts to standardize court data).

104. Federal Enterprise Data Resources, *DCAT-US Schema v1.1 (Project Open Data Metadata Schema),* https://resources.data.gov/resources/dcat-us/ (Last visited Jan. 17, 2023) ("Making metadata machine readable greatly increases its utility, but requires more detailed standardization, defining not only field names, but also how information is encoded in the metadata fields.").

105. There are at least two dozen city-, county-, and state-level projects geared toward creating minimal metadata standards, particularly for local government data. Data.gov, *Get Your Local Government on Data.gov* (Mar. 3, 2015), https://data.gov/opendata/get-local-government-data-gov/.

106. Michael A. Livermore, *The Perils of Experimentation*, 126 YALE L.J. 636, 636, 638 (2017) (reviewing the long literature on the costs and benefits of experimentation, including the possibility that subfederal devolution of policymaking authority may not generate rigorous informa-

even more traction where clashing data structures permit only apples-to-oranges comparisons across jurisdictions. We treat these issues, and the more general problems of interoperability they raise, in more detail in Part II below.

Perhaps more importantly, data integration that is designed to feed data-hungry ODR systems will often require controversial use of adjudication data drawn from *other* jurisdictions. For data scientists, the resulting dilemma is sometimes framed as a "distribution shift"—the risk that a model trained on data in one context will perform differently when it encounters data in a different context.[107] More concretely, system designers in numerous areas of the economy face a choice between a "custom" system (in the ODR context, a system that uses only the limited data drawn from the specific jurisdiction in which the ODR platform sits) and a "prebuilt" one (a system built on a wider pool of data, including data from other jurisdictions that may differ in relevant ways from the operative one).[108] This choice requires system designers to trade off tailoring against statistical power in pursuit of optimal precision. Moreover, elective judiciaries can make data-sharing controversial even *within* a relatively homogeneous state. In places where judges are meant to reflect the will of voters, why should litigants in one state trial court jurisdiction be primed with outcomes achieved in others? Data governance questions begin with the technical, but often end in deeply political waters.

Of course, courts need not hire a stable of technologists to develop their own ODR systems or oversee their implementation. Like administrative agencies that are fast adopting new algorithmic governance tools, courts will face a standard make-or-buy decision when developing new technologies or upgrading existing ones.[109] Some courts may

---

tion about possible innovations and may merely empower opponents of that innovation); Edward L. Rubin & Malcolm Feeley, *Federalism: Some Notes on A National Neurosis*, 41 UCLA L. Rev. 903, 911 (1994) (noting that sub-federal experimentation is a public good that will be systematically underproduced and, further, that, as a research-design matter, state-initiated policy initiatives will generate far less rigorous and actionable information than initiatives that are coercively *assigned* to states from the center).

107. *See* Rishi Bommasani et al., *On the Opportunities and Risks of Foundation Models* 18 (July 12, 2022), https://arxiv.org/pdf/2108.07258 ("A major limitation of standard machine learning is that it produces models that are not robust to distribution shifts, where the training distribution does not match the test distribution.").

108. *See, e.g.*, Manish Raghavan et al., *Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices*, ACM Conference on Fairness, Accountability, & Transparency at 10 (2019), https://arxiv.org/pdf/1906.09208 (explaining the custom/pre-built choice in the context of hiring platforms—algorithmic tools designed to help companies make hiring decisions).

109. The classic account is Oliver E. Williamson, *Public & Private Bureaucracies: A Transaction Cost Perspective*, 15 J.L. Econ. & Org. 306, 319 (1999). Useful theoretical and empirical analyses include Jody Freeman & Martha Minow, Eds., Government by Contract: Out-

choose to invest in internal technical capacity, training personnel, and assembling the raw materials necessary to create their own systems. Others will source ODR technology through the procurement process, relying on the many vendors offering ready-made platforms.[110] By tapping private sector expertise, out-sourcing can yield more technologically sophisticated tools at lower costs.[111] It may also be practical: Budgetary limits, civil service salary caps, and political sensitivities mean that government entities of all kinds, including courts, are often priced out of labor markets for employees with top-notch technical skill sets, limiting internal capacity-building.[112]

Reliance on private procurement, however, brings clear costs. Evidence from the growing field of public sector technology suggests that procurement-sourced tools might be more technologically sophisticated than home-grown ones, but they are often ill-suited for complex tasks and less policy-compliant and accountable.[113] Internal technical expertise may prove especially valuable where tasks are dynamic, changeable, and hard to measure.[114] Indeed, one of the great threats in public sector technology is that procurement-based sourcing of technology tools will steadily hollow out the embedded technical capacity of public entities while the private sector races ahead. As the public-private technology gap widens, government will grow ever more reliant upon procurement even as public officials lose the ability to exercise meaningful oversight.[115] This threat may be particularly

---

sourcing and American Democracy (2009); Jonathan Levin & Steven Tadelis, *Contracting for Government Services: Theory and Evidence from U.S. Cities*, 58 J. Indus. Econ. 507 (2010); Paul R. Verkuil, Outsourcing Sovereignty: Why Privatization of Government Functions Threatens Democracy and What We Can Do About It (2007).

110. *See* Steenhuis & Colarusso, *supra* note 42, at 777–78 (2021) (noting that courts have tended "to rely largely on outsourced technical expertise" rather than building their own internal technical capacity). One list of companies providing ODR services contains over 100 entities. *See* National Center for Technology and Dispute Resolution, *Provider List*, http://odr.info/provider-list/ (last visited Jan. 17, 2023).

111. David Freeman Engstrom, Daniel E. Ho, Catherine M. Sharkey & Mariano-Florentino Cuéllar, Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies, Report for the Administrative Conference of the United States 7–8 (2020) (noting these trade-offs) (hereinafter ACUS Report). For other views on technology procurement, particularly around AI, *see generally* Cary Coglianese & Erik Lampmann, *Contracting for Algorithmic Accountability*, 6 Admin. L. Rev. Accord 175 (2021); Deirdre K. Mulligan & Kenneth A. Bamberger, *Procurement as Policy: Administrative Process for Machine Learning*, 34 Berkeley Tech. L.J. 781 (2019).

112. *See* Engstrom et al., ACUS Report, *supra* note 111, at 89.

113. *Id*.

114. This is a standard finding of the transaction-cost literature on contracting out. *See, e.g.*, Levin & Tadelis, *supra* note 109 (reviewing theory and evidence).

115. *See generally* Aneesh Chopra, The Innovative State: How New Technologies Can Transform Government (2014) (offering a book-length examination of the public-private technology gap); *see also* Engstrom et al., ACUS Report, *supra* note 111, at 7–8 (noting

acute for courts because, unlike administrative agencies, many have scarce technical capacity in the first place.

A related threat to procurement-sourced tech is vendor "lock-in."[116] A challenge for courts adopting new digital and data-based systems is to ensure that the court owns the process, not outside technology vendors. This can be hard. Vendors have powerful incentives to supply their own proprietary systems without clean application programming interface (API) documentation in order to reduce compatibility with other systems and evade competition.[117] Vendors may also demand and obtain exclusive contracts—and offer enticing discounts in return—in order to safeguard future revenue flows.[118] As a result, reliance on a single technology vendor can prove easier and cheaper at the start-up phase, but the resulting lock-in effect can stifle later innovation or mid-course changes by giving the original vendor gatekeeper power over potentially more creative competitors.[119] In the ODR space and well beyond, courts that partner with vendors to build needed tech must avoid capture of a public market by a private actor.

The many complexities of data governance raise a final and vitally important set of governance questions about the *process* of procedure-making. Traffic rules, though hardly uncontroversial, nonetheless lend themselves to conventional, court-supervised rulemaking processes.[120] That long-standing status quo, as overseen by advisory committees and high courts, is notoriously slow-moving, but it seems reasonably

---

concerns). For a brief, ODR-specific discussion, *see* Larson, *supra* note 97, at 101 ("If a jurisdiction is going to rely on an outside vendor to build the ODR platform, which many jurisdictions will do because of the cost and time commitment required to build and maintain an internal system, it is important to determine how to collect information and educate yourself about what is possible technologically while still ensuring an unbiased request for proposals bidding process.").

116. *See* David Colarusso & Erika J. Rickard, *Speaking the Same Language: Data Standards and Disruptive Technologies in the Administration of Justice*, 50 SUFFOLK U. L. REV. 387, 392, 410 (2017) (noting "vendor lock-in" in the context of legal reforms); Justin C. Colannino, *Free and Open Source Software in Municipal Procurement: The Challenges and Benefits of Cooperation*, 39 FORDHAM URB. L.J. 903, 914 (2012) (exploring "vendor lock-in" in the municipal procurement context).

117. *See generally* Justice Opara-Martins, Reza Sahandi, & Feng Tian, *Critical Analysis of Vendor Lock-In and Its Impact on Cloud Computing Migration: A Business Perspective.* J CLOUD COMP. (2016).

118. Wilson C. Freeman & Jay B. Sykes, *Antitrust and "Big Tech,"* CONG. RSCH. SERV. 17 (Sept. 11, 2019), https://sgp.fas.org/crs/misc/R45910.pdf.

119. Jason Tashea, *The Justice System as a Digital Platform*, THE COMMONS (Sept. 30, 2020), https://wearecommons.us/the-justice-system-as-a-digital-platform/.

120. *See* Engstrom, *Digital Civil Procedure*, *supra* note 87, at 2219.

well-equipped to make the determination, say, whether to switch on ODR for this or that type of claim or dispute.[121]

In sharp contrast, data information and management rules are radically incompatible with current modes of rulemaking and judicial administration. Indeed, the creation of ODR 2.0 systems will not be a one-shot, *ex ante* process. Digitization is rarely something that occurs system-wide, and nor is it complete at a distinct moment in time. Instead, the process of building, implementing, and scaling digital systems tends to play out over time, in an iterative and dynamic design process.[122] It follows that, when it comes to ODR, wise implementation and meaningful accountability is unlikely to be achieved via conventional, ex ante, all-at-once rulemaking. Instead, responsible adoption and implementation will require something more akin to ongoing, stakeholder oversight, with continued pressure-testing and auditing to ensure that new digital systems are still functioning as intended.[123]

Constructing effective oversight schemes to perform that work will be no mean feat. If we know anything about algorithmic systems at this stage in the digital revolution, it is that they are socio-technical, human-machine assemblages, not mere ones and zeros.[124] Working out ODR's many technical complexities—from the modeling approach to be used in forming case-outcome predictions to the design of user interfaces—will demand purely technical expertise currently missing in many courts.[125] However, responsible design, development, and deployment of new digital systems will *also* require continuous

121. *Id*.; *see also* Linda S. Mullenix, *Hope over Experience: Mandatory Informal Discovery and the Politics of Rulemaking*, 69 N.C. L. Rev. 795, 802 (1991) (describing the Advisory Committee rulemaking process as "painfully slow, deliberative, and dull").

122. *See* Jakob Nielsen, *Iterative User Interface Design*, originally published in IEEE Computer Vol. 26, No. 11, 32–41 (Nov. 1, 1993), available at https://www.nngroup.com/articles/iterative-design/.

123. *See generally* Leah Wing, Janet Martinez, Ethan Katsh, & Colin Rule, *Designing Ethical Online Dispute Resolution Systems: The Rise of the Fourth Party*, 37 Negot. J. 49 (2001).

124. Mike Ananny & Kate Crawford, *Seeing Without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability*, 20 New Media & Soc'y 973, 983 (2016) (referring to algorithmic systems as human-machine "assemblages"); Danielle Keats Citron, *Technological Due Process*, 85 Wash. U. L. Rev. 1249, 1264–66 (2008) (providing a taxonomy of "mixed systems"). In particular, programmers must make numerous decisions, about how to partition the data, model specifications, which data, target variables, and data features to use, and how to tune the model. David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, 51 U.C. Davis L. Rev. 653, 683–700 (2017).

125. *See* John M. Greacen, *Eighteen Ways Courts Should Use Technology to Better Serve Their Customers*, IAALS 1 (Oct. 2018) (discussing National Center for State Courts survey revealing most consumers find state courts "severely lacking" in technical abilities), https://iaals.du.edu/sites/default/files/documents/publications/eighteen_ways_courts_should_use_technology.pdf.

engagement by non-technical forms of expertise. As in other realms, from public policy to hard science, we need "diverse scrutiny" from a mix of experts and stakeholders, including those who will feel the direct and indirect effects of decision-making.[126] Without pluralistic design input, ODR's critics rightly worry that the innovation sold as a way to empower the marginalized will instead warp into a highly efficient, Fordist system of oppressive fee collection, wage garnishment, and evictions.[127] Ensuring "diverse scrutiny" or, at the very least, clear rules of engagement, is a first line of defense against that possibility.

Much work remains to be done to specify what such a multi-stakeholder design and governance scheme could or should look like. One might start by examining the various bodies, among them state access-to-justice commissions, that have pioneered innovations in recent decades, from pro se court forms and self-help centers to court navigator programs and limited scope lawyer and non-lawyer assistance programs.[128] Courts would also be wise to consider a growing academic literature, leavened by accumulating experience, that is coalescing around "open policymaking" as means to achieving efficacy *and* legitimacy.[129] In broad strokes, that might mean a design and implementa-

126. *See* Wendy Wagner & Martin Murillo, *Is the Administrative State Ready for Big Data?: Exploring the Accountability Challenges in Environmental and Public Health Regulation*, KNIGHT FIRST AMEND. INST. (2021) (using the "diverse scrutiny" formulation); *see also* Mireille Hildebrandt, *Algorithmic Regulation and Rule of Law*, PHIL. TRANS. R. SOC. 376 (2018) (framing the challenge as creating space for "agonistic debate" between data scientists, expert lawyers, and lay people, including those who will feel the direct and indirect effects of decisions about system design and implementation). *See generally* HENRY W. CHESBROUGH, OPEN INNOVATION: THE NEW IMPERATIVE FOR CREATING AND PROFITING FROM TECHNOLOGY (2003) (providing a foundational analysis of how to achieve "open innovation" by working across organizational and disciplinary boundaries).

127. *See* Spaulding, *supra* note 17, at 266–69; *see also* Paul Kiel & Jeff Ernsthausen, *Capital One and Other Debt Collectors Are Still Coming for Millions of Americans*, PROPUBLICA (June 8, 2020), https://www.propublica.org/article/capital-one-and-other-debt-collectors-are-still-coming-for-millions-of-americans (reporting evidence that debt collection plaintiffs were able to accelerate filings during the pandemic by using digital systems to bring lawsuits in bulk).

128. For more on state access to justice commissions created to craft responses to the pro se crisis, *see* American Bar Association, *Access to Justice Commissions: Increasing Effectiveness Through Adequate Staffing and Funding* 1, 8, 9 (Aug. 2018); *see also* American Bar Association, *Access to Justice Commissions*, https://www.americanbar.org/groups/legal_aid_indigent_defense/resource_center_for_access_to_justice/atj-commissions/.

129. *See* Beth Simone Noveck, Rose Harvey, & Anirudh Dinesh, *The Open Policymaking Playbook*, NYU GOVLAB 6 (Apr. 2019), https://thegovlab.org/static/files/publications/openpolicymaking-april29.pdf. For a court-specific version, *see, e.g.*, IAALS, *LISTEN > LEARN > LEAD: A Guide to Improving Court Services Through User-Centered Design (2019)*, https://iaals.du.edu/sites/default/files/documents/publications/listen-learn-lead_improving_court_services.pdf. A similar set of ideas is gaining traction in administrative law, including possible use of "citizen assemblies" and "citizen juries" to develop and oversee policymaking. The idea is often traced to work in the early 1970s at the Jefferson Center. *See* Center for New Democratic Processes, *How We Work / Citizen Juries* (2019), https://

tion process that taps both the collective intelligence of diverse experts and the distributed wisdom and lived experience of an equally diverse public.[130] These ideas travel under numerous headings, from "participatory design" to "human-centered design," but they all boil down to constructing decisional processes that can simultaneously engage both experts and end-users.[131] It might also mean developing a system of "social auditing" to collect feedback from stakeholders on an ongoing basis.[132] The most highly structured versions of this approach assign concrete and well-defined tasks to specific stakeholders—nominating this or that group to monitor, say, data integrity or platform usability.[133]

One can remain skeptical of some of the management-speak that too often creeps into the playbooks sketching these design approaches. To that extent, judicial governance will surely see the same healthy skepticism already directed at the invocation of participatory design in other technology areas, particularly pervasive calls to "democratize AI."[134] "Diverse scrutiny" is hardly a panacea, and its benefits will depend on the quality of its design and implementation. Still, it seems clear that chief judges and court administrators should not, and cannot, be the sole arbiters of the process of designing the new civil justice system, whether ODR or otherwise. Courts, long behind the digitization curve, have remained mostly aloof from brewing debates about which governance approaches can bring meaningful ac-

---

www.cndp.us/about-us/how-we-work/. The concept is infused by the Deweyian notion that ordinary people combine common sense with local knowledge that more distant regulators—or chief judges—will often lack. *See* Mark Tushnet, *Introduction: The Pasts & Futures of the Administrative State*, 150 DAEDULUS 5, 13 (2021). Members are chosen at random from the population, compensated for their time, provided with general information about the problem before them, empowered to call upon experts where appropriate, and then charged with making binding decisions about policy. *Id.* at 14.

130. *See* Noveck et al., *supra* note 129, at 6.

131. *See generally* Victor D. Quintanilla, *Human-Centered Civil Justice Design*, 121 PENN ST. L. REV. 745 (2017) (comparing a human-centered approach to client-centered lawyering and traditional design approaches that elevate the designer's expertise above the end user's experience); *see also* Daniel W. Bernal & Margaret D. Hagan, *Redesigning Justice Innovation: A Standardized Methodology*, 16 STAN. J. CIV. RTS. & CIV. LIBERTIES 335, 335 (2020) (calling for the integration of both "expert-oriented and user-centered approaches").

132. *Id*. at 48–49.

133. *Id*.

134. *See, e.g.*, Johannes Himmelreich, *Against "Democratizing AI," AI & Soc'y* (2022), https://johanneshimmelreich.net/papers/against-democratizing-AI.pdf. For older accounts of participatory design, *see generally* DOUGLAS SCHULER & AKI NAMIOKA, PARTICIPATORY DESIGN: PRINCIPLES AND PRACTICES (1993) (one of the foundational textbooks); TERRY WINOGRAD, BRINGING DESIGN TO SOFTWARE (1996) (applying participatory design to software development).

countability to potent new forms of automation. No longer. As the legal system digitizes, courts will be thrust into those debates.

II. Courts as Data Dispensers

Courts are data monopolists. They sit atop mountains of data that grow larger each year, fed by a ceaseless stream of litigant-filed pleadings and papers and court-issued orders and decisions. This ever-growing pile of court records is rapidly digitizing, driven in part by pandemic-era efforts to minimize contagion at the courthouse via expanded e-filing options.[135] As the sole holders of the ensuing data trove, courts get to decide the terms on which data is made available to others who might benefit from it. In other words, courts already serve, and will continue to serve, as data *dispensers* and even data *curators*. Here, then, is a second vitally important data governance role that courts will increasingly play: in a fast-digitizing civil justice system, data policies and practices will critically shape outcomes by determining who can access and use court data.

A. Legal Tech: Generating a Data Deluge

Data curation will be a vitally important mode of data governance because "legal tech" tools are increasingly integral to the civil justice system, and court data is their lifeblood. To see the role that data governance can and will play going forward, one must first understand three things: (i) the current capacities and trajectory of the legal-tech toolkit; (ii) the benefits it confers on its users and on the civil justice system as a whole; and (iii) the distributive concerns it raises given unequal access to its inputs.

Over the past decade or two, a growing kit of legal-tech tools has steadily transformed how lawyers serve their clients. An initial wave of technologies was not so different from the more generic digitization

135. It is difficult to gauge the uptick in e-filing. There has been no systematic effort to track it. As a result, and as noted previously, much of our evidence is anecdotal. *See* Rickard & Naqui, *supra* note 78. But anecdotal evidence is all around. For instance, a January 2021 report of the Structural Innovations Working Group of the Commission to Reimagine the Future of New York's Courts advocated legislation granting the Chief Administrative Judge of the New York courts power to expand mandatory e-filing to all of the state's trial courts in any case type. *See* Structural Innovations Working Group of the Commission to Reimagine the Future of New York's Courts, *The Expansion of Electronic Filing: A Report and Recommendations of the Structural Innovations Working Group* (Jan. 2021), https://www.nycourts.gov/LegacyPDFS/publications/pdfs/CommitteeReport-eFiling.pdf. In New York, that would mean expanding e-filing beyond commercial cases to family law courts and "lower" courts, which handle civil claims below a certain dollar threshold. Another indicator of the growing digitization of court records is the advent of machine learning techniques that scan and ingest documents to automatically docket them. *See* La Roque-Doherty, *supra* note 100, at 2, 20.

of business practices—*i.e.*, word processing, spreadsheets—that transpired at the dawn of the current digital age. For instance, early legal-tech tools helped lawyers manage documents, track team-based work products, and automate back-office administrative tasks, such as client intake and billing.[136]

A more recent wave injected legal tech directly into lawyers' workstreams. A family of e-discovery tools that now pass under the banner of "technology-assisted review" or "predictive coding" uses machine learning models to tag documents for relevance and privilege, automating the otherwise time-consuming task of performing manual, "eyes on" review before or after production.[137] Another booming area consists of tools that identify patterns and anomalies in large corpora of legal texts—for instance, searching thousands of contracts in order to identify where a proposed deal deviates from industry-standard provisions.[138]

A third wave, and the current state of the art, may soon eclipse these older tools through continuing advances in natural language processing, (NLP). Indeed, potent new NLP models, such as Google's BERT or OpenAI's GPT-3 (and its offspring, ChatGPT), have made it possible to treat law itself as a form of data.[139] Pushing past repetitive tasks such as e-discovery or contract analysis, newer legal-tech

136. A good contemporary example of such a tool is Clio. *See* Clio, https://www.clio.com/ (last visited Jan. 17, 2023).

137. For a description of TAR and some of its implications for the adversarial system, *see* Engstrom & Gelbach, *supra* note 10, at 1017–18. For a recent accounting of TAR debates, *see* Neel Guha, Peter Henderson, & Diego Zambrano, *Gamesmanship in Modern Discovery Tech*, in Legal Tech and the Future of Civil Justice 112, 126-33 (Engstrom ed., 2023). For seminal work on TAR, *see* Seth Katsuya Endo, *Technological Opacity & Procedural Injustice*, 59 B.C. L. Rev. 821, 822–24 (2018); Dana A. Remus, *The Uncertain Promise of Predictive Coding*, 99 Iowa L. Rev. 1691, 1701–06 (2014). For a recent technical analysis, *see* Eugene Yang, Sean MacAvaney, David D Lewis, & Ophir Frieder, *Goldilocks: Just-Right Tuning of BERT for Technology-Assisted Review* (2021), https://arxiv.org/pdf/2105.01044.

138. Kathryn D. Betts & Kyle R. Jaep, *The Dawn of Fully Automated Contract Drafting: Machine Learning Breathes New Life into a Decades-Old Promise*, 15 Duke L. & Tech. Rev. 216, 226, 229 (2017); Emad Elwany, Dave Moore, & Gaurav Oberoi, *BERT Goes to Law School: Quantifying the Competitive Advantage of Access to Large Legal Corpora in Contract Understanding* (2019), https://arxiv.org/abs/1911.00473; Marco Lippi, Przemysław Pałka, Giuseppe Contissa, Francesca Lagioia, Hans-Wolfgang Micklitz, Giovanni Sartor, & Paoli, Torroni, *CLAUDETTE: An Automated Detector of Potentially Unfair Clauses in Online Terms of Service*, 27 Artificial Intelligence and Law 117 (2019); Dan Hendrycks, Collin Burns, Anya Chen, & Spencer Ball, *CUAD: An Expert-Annotated NLP Dataset for Legal Contract Review* (2021), https://arxiv.org/abs/2103.06268.

139. *See* Michael A. Livermore, Peter Beling, Keith Carlson, Faraz Dadgostari, Mauricio Guim, & Daniel N. Rockmore, *Law Search in the Age of the Algorithm*, 2020 Mich. St. L. Rev. 1183, 2019–20 (2020) (describing the "law as data" movement). For a short but excellent overview of the use of "foundational models" such as BERT and GPT-3 in the law domain (as well as dozens of other domains), *see* Percy Liang et al., *On the Opportunities and Risks of Foundation*

tools apply NLP to judicial decisions, briefs, and case documents in order to automate a range of higher-order legal cognitions.[140] It is here that we can expect the most significant leaps forward, as artificial intelligence (AI) increasingly supplements and, at times, supplants core lawyerly work.

A full tour of the current legal-tech toolkit can be found elsewhere and is not discussed here.[141] For now, one might begin by noting that the most promising among the newer generation of legal-tech tools reduce to two core tasks.

The first task is legal research and analytics.[142] The most basic of these can be thought of as hunting-and-gathering tools that assist lawyers in corralling relevant legal materials, oftentimes accompanied by automated content summaries or annotations of their relevance to the case at hand.[143] More advanced versions permit more focused inquiries—a natural-language "Q&A" (question-answer) system for legal professionals.[144] These tools deliver faster and surer attorney traction at the outset of a legal matter by collecting a core set of materials for an attorney to consider—think Westlaw and Lexis-Nexis on steroids.[145] Other more advanced versions go beyond initial information-

Models 59–66, Center for Research on Foundation Models (CRFM) and Stanford Institute for Human-Centered AI (2021), https://arxiv.org/abs/2108.07258.

140. *See* Engstrom & Gelbach, *supra* note 10, at 1020–24.

141. *See id*. at 1011–12; *see also* Daniel N. Kluttz & Deirdre K. Mulligan, *Automated Decision Support Technologies and the Legal Profession*, 34 BERKELEY TECH. L.J. 853, 889 (2019) (offering an overview of lawyer-driven legal tech tools).

142. *See* Livermore et al., *supra* note 139, at 1183, 1189, 1207, 1238 (offering a foundational descriptive and normative analysis of "law search," including the systemic implications of new search technologies from the perspective of both social welfare and rule of law values—for instance, for the degree of "convergence" in a legal system; further arguing that law search is "an integral component of legal reasoning"); *see generally* Jens Frankenreiter & Michael A. Livermore, *Computational Methods in Legal Analysis*, 16 ANN. REV. L. & SOC. SCI. 39 (2020) (analyzing means of data-driven legal research).

143. A similar type of tool is increasingly plausible for judges and other adjudicators. Particularly in high-volume, mass adjudication contexts, AI-powered tools can build what one federal administrative law judge referred to as a "decisional shell" around a case by gathering together relevant materials and even beginning to fill out a template of a decision. *See* ENGSTROM ET AL., ACUS REPORT, *supra* note 111, at 85.

144. For a state-of-the-art analysis of Q/A systems, *see generally* Andrew Vold & Jack G. Conrad, *Using Transformers to Improve Answer Retrieval for Legal Questions* (2021), https://www.conradweb.org/~jackg/pubs/ICAIL21_Vold_Conrad.pdf.

145. Conventional digital case law search, including Westlaw and Lexis, now also relies on AI. A good example is Ross Intelligence, now shuttered as a result of a copyright suit by Westlaw, which used an NLP engine to identify legal authorities in response to a natural language query. *See* Ross Intelligence, *Legal Tech Corner*, https://blog.rossintelligence.com/ (last visited Jan. 17, 2023); *see also* David Houlihan*, Ross Intelligence and Artificial Intelligence in Legal Research,* BLUE HILL RSCH. 1–2 (Jan. 2017), https://www.iadclaw.org/assets/1/7/6.2-_Blue_Hill_Benchmark_Report_-_Artificial_Intelligence_in_Legal_Research.pdf.

gathering and caselaw research and evaluate attorney work product *after* it is generated. Document analyzers permit an attorney to input a draft brief and receive back an analysis of questionable caselaw that was cited, apposite caselaw that was missed, or even an argument left on the table.[146] Most advanced of all, but still very much in development, are tools that create work product themselves, generating initial drafts of pleadings, discovery requests, and even briefs.[147]

The second task is predicting case outcomes—in many ways the essence of lawyering.[148] As Justice Holmes long ago noted, law is, once shorn of its many pretensions, little more than a prediction as to how a court will rule.[149] Case-outcome prediction tools take various forms, but most operate by analyzing a defined set of legal and related materials—for instance, past cases and decisions, the identity of opposing counsel, the relevant jurisdiction, and even stylistic features of the briefs—and then returning a probability or set of probabilities over different possible outcomes.[150]

---

146. *See* Zihan Huang, Charles Low, Mengqiu Teng, Hongyi Zhang, Daniel E Ho, Mark S Krass, & Matthias Grabmair, *Context-Aware Legal Citation Recommendation using Deep Learning* (June 2021), https://arxiv.org/abs/2106.10776; Malte Ostendorff, Elliott Ash, Terry Ruas, Bela Gipp, Julian Moreno-Schneider, & Georg Rehm, *Evaluating Document Representations for Content-Based Legal Literature Recommendations* (2021), https://arxiv.org/abs/2104.13841. Casetext offers a pair of currently marketed examples. One is Casetext's CARA system, which ingests a whole document and outputs relevant legal authorities. *See* CASETEXT, http://www.casetext.com/cara-ai/ (describing the CARA process) (last visited Jan. 17, 2023); cf. Beth Hoover, *Introducing Clerk: Win More Motions with Intelligent Brief Analysis*, JUDICATA (Oct. 5, 2017), https://blog.judicata.com/introducing-clerk-848abbed8fd3 (discussing a similar product, Clerk, which is limited to California state law). The other is Casetext's Compose, which purports to draft legal arguments by recommending conceptually similar legal authorities that fit a case's fact pattern as inputted by a user. *See* COMPOSE, https://compose.law (describing Compose's process) (last visited Jan. 17, 2023).

147. *See* Rob Carty, *Computer-Written Legal Briefs Are Closer Than You Think*, ARTIFICIAL LAWYER (Apr. 11, 2019), https://www.artificiallawyer.com/2019/04/11/computer-written-legal-briefs-are-closer-than-youthink. *See also* note 193, *infra* and accompanying text (describing legal tool that is the joint effort of Walmart, a BigLaw law firm, and a tech company that can generate first drafts of pleadings and discovery requests).

148. Daniel Martin Katz, *Quantitative Legal Prediction—or How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry*, 62 EMORY L.J. 909, 910 (2013) (arguing that case-outcome prediction is a core lawyerly task and suggesting that quantitative analysis of big data can outperform analog lawyer predictions). Mark K. Osbeck, *Lawyer as Soothsayer: Exploring the Important Role of Outcome Prediction in the Practice of Law*, 123 PA. ST. L. REV. 41, 43–44 (2018).

149. Oliver Wendell Holmes, *The Path of the Law*, 10 HARV. L. REV. 457, 458 (1897) ("[A] legal duty so called is nothing but a prediction that if a man does or omits certain things he will be made to suffer in this or that way by judgment of the court. . . ."); *Id.* at 461 ("The prophecies of what the courts will do in fact, and nothing more pretentious, are what I mean by the law. . . .").

150. For an example of how this type of analysis manifests in traditional literature, *see, e.g.*, Elizabeth Chika Tippett et al., *Does Lawyering Matter? Predicting Judicial Decisions from Legal Briefs, and What That Means for Access to Justice*, 100 TEX. L. REV. 1157 (2022).

As with some of the more advanced legal (re)search and analytics tools, case-outcome prediction engines are very much under development. For the moment, the most advanced tools cluster in relatively narrow, self-contained, and often technocratic areas, such as tax or labor law.[151] They are also time-consuming and labor-intensive to create. Contrary to public depictions of machine learning systems as simply turned loose on oceans of data, case-outcome prediction engines require substantial lawyer engagement at the front-end. A key first step, for example, is creating legal ontologies—that is, structured mappings or "knowledge representations" of doctrine, such as the elements or factors that have driven past case outcomes. Lawyers must then manually label a pool of legal texts to train machines to identify those elements or factors on their own.[152] Against the puffery of tech entrepreneurs and a growing academic literature that imagines a future legal system populated by "robolawyers" and "robojudges,"[153] measured assessment is clearly in order. Still, these tools underscore legal tech's great potential and power. Once a model is built, it conceivably operates at scale and very low marginal cost.[154] It can generate sharp insights—for instance, that this or that doctrinal element, long thought to drive judicial determinations of a given issue, has be-

151. Benjamin Alarie, Anthony Niblett, & Albert H. Yoon, *Using Machine Learning to Predict Outcomes in Tax Law*, 58 CANADIAN BUS. L.J. 231, 253 (2016) (reporting success in the tax area); Nils Holzenberger, Andrew Blair-Stanek, & Benjamin Van Durme, *A Dataset for Statutory Reasoning in Tax Law Entailment and Question Answering* (2020), https://arxiv.org/abs/2005.05257 (same, though not with deep learning); Charlotte S. Alexander, Khalifeh al Jadda, Mohammad Javad Feizollahi, & Anne M. Tucker, *Using Text Analytics to Predict Litigation Outcomes*, in LAW AS DATA: COMPUTATION, TEXT, & THE FUTURE OF LEGAL ANALYSIS 310 (Michael A. Livermore & Daniel Rockmore eds., 2019) (describing prediction challenges in the employment discrimination context).

152. For a full description of this process, *see* Engstrom & Gelbach, *supra* note 10.

153. *See generally* Eugene Volokh, *Chief Justice Robots*, 68 DUKE L.J. 1135 (2019) (predicting a future with robot judges); *see also* R.J. Vogt, *DoNotPay Founder Opens Up on 'Robot Lawyers'*, LAW360, (Feb. 9, 2020), https://www.law360.com/articles/1241251; Richard M. Re & Alicia Solow-Niederman, *Developing Artificially Intelligent Justice*, 22 STAN. TECH. L. REV. 242, 242 (2019) ("[T]he prospect of 'robot judges' suddenly seems plausible–even imminent."); Aziz Z. Huq, *A Right to a Human Decision*, 106 VA. L. REV. 611, 635 (2020) (discussing current algorithmic decision-makers); Frank Pasquale, *A Rule of Persons, Not Machines: The Limits of Legal Automation*, 87 GEO. WASH. L. REV. 1, 3 (2019); Benjamin Chen, Alexander Stremitzer, & Kevin Tobia, *Having Your Day in Robot Court*, HARV. J. L. & TECH. at 3 (forthcoming 2023).

154. The ODR platform Matterhorn, for example, scaled rapidly from 2019 to 2020, completing double the cases and expanding from fewer than 50 courts to 115 courts and counting. *Compare* SUSSKIND, *supra* note 1, at 175 (referencing Matterhorn website as of April 2019), *with* University of Michigan Innovation Partnerships, *2020 Distinguished Innovator of the Year* (Nov. 18, 2020), https://innovationpartnerships.umich.edu/stories/2020-distinguished-innovator-of-the-year/.

come (or perhaps always was) insignificant or even irrelevant.[155] In close cases, those sorts of insights can be invaluable.

B. The Power of Data and Controlling Its Dissemination

A concrete understanding of legal tech's current capacities and trajectory helps us to see its transformational potential as well as why court data policies and practices will be so central to the health of the justice system going forward. Indeed, abstracting away from any particular part of the legal-tech toolkit, one can see that the newer wave of tools share two features in common. First, they are united by their ability to shift the distribution of litigation's most important quantities: costs and information.[156] Second, these tools raise acute distributional challenges because they depend on data either uniquely held by courts or unevenly distributed among litigants.

Begin with the value that the more advanced legal tech tools deliver to users: lower costs and better information.[157] The significance of reduced litigation costs is easily grasped by anyone familiar with American litigation. Several generations of thinking about civil litigation and procedure have been pre-occupied—some would say obsessed—with litigation costs as undue drivers of case outcomes, and for a good reason.[158] Basic litigation economics holds that costs shape the distribution of the settlement surplus at the bargaining table.[159] Lowering litigation costs might thus bring systemic benefits, weakening the connection between effort and outcome, and thus bringing the system closer to one in which a claim's merit, rather than its costs, determines case outcomes.

Legal tech's capacity to lower absolute litigation costs may also narrow cost asymmetries. As one of us has noted elsewhere, it might thereby reduce the significance of certain procedural rules, such as the proportionality rules central to recent rulemaking efforts.[160] The narrowing of cost asymmetries might even cause us to reconsider *Twombly*'s plausibility pleading rule, which is based on a concern,

155. *See* Engstrom & Gelbach, *supra* note 10, at 1025–26.

156. *Id*. at 1005.

157. *Id.* at 1004–05 ("[L]egal tech's proliferation is likely to alter two foundational aspects of any litigation system: the distribution of litigation costs and the distribution of information."); Livermore et al., *supra* note 139, at 1209 (discussing implications of lower cost law search).

158. *See* Engstrom & Gelbach, *supra* note 10, at 1045.

159. For an overview, *see* Andrew F. Daughety & Jennifer F. Reinganum, *Settlement*, in 8 Encyclopedia of Law and Economics 386, 386–71 (Chris W. Sanchirico ed., 2d ed. 2012).

160. *See* Engstrom & Gelbach, *supra* note 10, at 1051.

however well-founded, about one-sided, impositional discovery requests.[161]

Legal tech tools, and the systematic reduction in litigation costs they promise, also have the potential to reshape the legal services marketplace. Lower costs reduce the price of legal services, potentially putting those services within reach of a larger set of individuals with legal needs. One of the signal trends of the past several decades has been the steady decline of PeopleLaw—the segment of the legal services market that serves individuals and small businesses, as distinct from BigLaw's corporate focus.[162] If legal tech reduces the cost of providing legal services, PeopleLaw could rebound, thus mitigating pervasive access challenges within the system. Unlike ODR 2.0 and its technological substitution for counsel, legal tech might actually grow the supply of lawyers willing and able to meet the civil justice needs that currently go unmet.

Legal tech's informational benefits are even more varied than its cost-reducing benefits and serve a mix of efficiency and rule-of-law values that may not be apparent at first glance. From a societal perspective, better and cheaper legal information can yield quicker, more socially efficient settlements. More generally, better information allows attorneys to provide more precise counsel to clients about available courses of action.[163] "Law-abiding people," as Justice Jackson put it in *Hickman v. Taylor*, depend on "the lawyer and the law office" to "learn the ever changing and constantly multiplying rules by which they must behave and to obtain redress for their wrongs."[164] Michael Livermore and co-authors give Jackson's insight a modern twist, persuasively arguing that legal tech—particularly the newer generation of legal-search tools—serves core rule-of-law values.[165] While that concept can mean different things to different people, most accounts agree that, at its core, rule of law means laws that are clear, determi-

161. *Id*.

162. State Bar of California, *State Bar of California Task Force on Access Through Innovation of Legal Services* 7–8 (2020), https://www.calbar.ca.gov/Portals/0/documents/publicComment/ATILS-Final-Report.pdf; William D. Henderson, *Legal Market Landscape Report* 12–15 (2018), https://board.calbar.ca.gov/docs/agendaItem/Public/agendaitem1000022382.pdf.

163. *See* Hadfield, *Legal Markets*, *supra* note 11, at 1270–75 (reducing legal systems to three core tasks: rule production, rule enforcement, and legal services, with the latter consisting of information about rules and the services needed to develop strategy in light of those rules and their enforcement).

164. Hickman v. Taylor, 329 U.S. 495, 514–15 (1947).

165. *See* Livermore et al., *supra* note 139, at 1203, 1211–13.

nate, and predictable.[166] More precise and more efficient acquisition of legal information serves those values by facilitating a high degree of what Livermore and co-authors call "convergence"—that is, the amount of agreement about the content of legal rights and obligations.[167] Returning to Justice Jackson's view, legal tech's convergence-promoting power can ultimately increase law compliance in an ever-more complex regulatory state.[168]

More unsung but no less important is the salutary role that information plays in triage and intake, crucial moments when lawyers make business-critical decisions about which clients to take on in the first place. For instance, legal aid lawyers, who have scarce resources to offer their anything-but-scarce prospective clients, must think hard about which representations to undertake.[169] Likewise, intake is a constant struggle for the plaintiffs' bar, which must make sound bets in order to succeed, or just keep the lights on, when operating on a contingent-fee basis.[170] A growing litigation finance industry now assists with that task—and shares in the benefits of increasingly sophisticated outcome-prediction engines—but it is neither pervasive enough nor sufficiently aligned with the interests of plaintiffs' counsel to fully

166. Jeremy Waldron, *The Rule of Law*, THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., Summer 2020), https://plato.stanford.edu/archives/sum2020/entries/rule-of-law/.

167. *See* Livermore et al., *supra* note 139, at 1207. A classic law and economics literature on the evolution of the common law emphasizes the role of private actors in supplying information to courts as they engage in welfare-enhancing lawmaking. *See, e.g.*, RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW (1973); *see also* Paul Rubin, *Why Is the Common Law Efficient?*, 6 J. LEGAL STUDS. 51, 53 (1977); George L. Priest, *The Common Law Process and the Selection of Efficient Rules*, 6 J. LEGAL STUDS. 65, 72 (1977); Nicola Gennaioli & Andrei Shleifer, *The Evolution of Common Law*, 115 J. POL. ECON. 43, 46 (2007).

168. *See* Livermore et al., *supra* note 139, at 1230. An important part of the Livermore view is that, where legal search is costly, system convergence will be low because regulated parties will under-invest, relative to the social optimum, in the legal search necessary to generate a clear view of their legal rights and obligations. *Id*. at 1210. Viewed through this lens, legal search and the ensuing legal knowledge it provides is a public good that, if left to the private market alone, will likely be underproduced. *Id.* at 1230. Alternatively, widely available data—and even "public option" legal search—can be framed as the kind of public subsidy thought to produce valuable positive externalities. *Id.* at 1236.

169. *See, e.g.*, Minnesota Legal Services Advisory Committee, *Analysis of the Civil Legal Aid Intake Infrastructure in Minnesota Final Report* 8 (June 2017), https://www.mncourts.gov/mncourtsgov/media/scao_library/documents/Minnesota-intake-study-Final-Report-6-14-17.pdf (examining reasons for denial of service at legal aid organizations).

170. *See* John C. Coffee Jr., *Understanding the Plaintiff's Attorney: The Implications of Economic Theory for Private Enforcement of Law Through Class and Derivative Actions*, 86 COLUM. L. REV. 669, 706–12 (1986) (discussing the lean structure of plaintiffs' firms, which impairs their ability to diversify their case portfolios); Nora Freeman Engstrom, *Re-Re-Financing Civil Litigation: How Lawyer Lending Might Remake the American Litigation Landscape, Again*, 61 UCLA L. REV. DISCOURSE 110, 112 (2013) (noting that "changes to lawyers' financial, social, and business structures have the power to influence case outcomes").

substitute for a lawyers' own case evaluations.[171] Here, better and cheaper information serves important systemic efficiency ends by matching counsel and client.

All this is for the good. What, then, is the catch? Enter the second feature that unites the newer generation of legal-tech tools: They all depend in substantial part on data that courts control and that are unequally available to litigants. These two factors are critical, for legal-tech tools often derive their value from exclusivity: the fact that some litigants have it, and others do not.[172] Information is power in litigation. More and better information about claim value and likely litigation effort permits those who possess it to systematically extract higher settlements.[173] Better information also enables litigation's repeat-player "haves" to win out over one-shot "have-nots" with familiar tactics: settle the cases with strong claims, litigate the winners, and "play for rules" at the appellate level.[174] If some litigants have more ready access than others to legal tech's cost and informational advantages, legal tech might be just one more way that the "haves" of the litigation system come out ahead. Access to data will increasingly shape access to justice.

For data that courts control, a key implication, and the subject of an emerging "open court data" movement,[175] is that the decisions chief

171. Michael K. Velchik & Jeffrey Y. Zhang, *Islands of Litigation Finance*, 24 STAN. J.L. BUS. & FIN. 1, 41 (2019).

172. *See* Engstrom & Gelbach, *supra* note 10, at 1089.

173. The logic behind this is more complicated than it first appears to be. For a full review of how existing models of litigation selection and settlement would arrive at this conclusion, *see* Engstrom & Gelbach, *supra* note 10, at 1074–75.

174. *See* Mark Galanter, *Why the Haves Come Out Ahead: Speculations on the Limits of Legal Change*, 9 LAW & SOC'Y REV. 95, 100–04 (1974); Albert Yoon, *The Importance of Litigant Wealth*, 59 DEPAUL L. REV. 649, 651, 658–59 (2010).

175. *See* Charlotte S. Alexander & Mohammad Javad Feizollahi, *On Dragons, Caves, Teeth, and Claws: Legal Analytics and the Problem of Court Data Access*, *in* COMPUTATIONAL LEGAL STUDIES: THE PROMISE AND CHALLENGE OF DATA-DRIVEN LEGAL RESEARCH 97 (Ryan Whalen, ed. 2019). Alexander is hardly the first to note this, however. A long and increasingly bitter literature decries the inaccessibility of court data. *See, e.g.*, Peter W. Martin, *Online Access to Court Records – from Documents to Data, Particulars to Patterns*, 53 VILL. L. REV. 854, 870–71 (2008); Peter W. Martin, *District Court Opinions that Remain Hidden Despite a Long-standing Congressional Mandate of Transparency – The Result of Judicial Autonomy and Systemic Indifference* 110 Law Libr. J. 305, 323, 329 (2018); Lynn M. LoPucki, *The Politics of Research Access to Federal Court Data*, 80 TEX. L. REV. 2161, 2162 (2002); Elizabeth Y. McCuskey, *Submerged Precedent*, 16 NEV. L.J. 515, 520 (2016); Pauline T. Kim, Margo Schlanger, Christina L. Boyd and Andrew D. Martin, *How Should We Study District Judge Decision-Making?*, 29 WASH. U. J.L. & POL'Y 83, 102–03(2009); Stephen J. Schultze, *The Price of Ignorance: The Constitutional Cost of Fees for Access to Electronic Public Court Records*, 106 GEO. L.J. 1197, 1225–26 (2018); Jonah Gelbach, *Free Pacer* (Feb. 26, 2021) (unpublished manuscript, presented at Legal Tech and the Future of Civil Justice virtual conference at Stanford Law School), https://drive.google.com/file/d/1j6AkMmQJznzLZgxKqM4UzzSeEDhdhTPe/view;

judges and court administrators make about data infrastructure will shape the innovation ecosystem and, in turn, who the legal tech industry serves. For the moment, the picture is not a pretty one. While American constitutions, statutes, and rules at all levels of government are dotted with "open court" provisions,[176] court data has long been some of the most closely held among public sector data—and some of the hardest to access.[177] Charlotte Alexander, an expert on legal tech and a leader of the "open court data" movement, puts it best: court records, from the federal level on down, sit behind "a wall of cash and kludge."[178] Paywalls and clunky user interfaces put court records, particularly the bulk downloads needed to build vibrant legal tech applications, beyond the reach of all but the most well-heeled law firms and tech companies.[179] While a number of non-profit efforts have valiantly publicized large amounts of American "decisional" law,[180] legal empiricists have long known that opinions are only the tip of the ice-

---

Livermore et al., *supra* note 139 at 1226–27 (decrying PACER issues and calling for publicly subsidized or "public option" law search, perhaps via the creation of a data clearinghouse containing all of US law in easily extractable form); *see generally* Adam R. Pah et al., *How to Build a More Open Justice System*, 369 SCIENCE 134 (2020) (noting some results from the so-called SCALES initiative https://scales-okn.org/about-the-project).

176. *See, e.g.*, Fed. R. Civ. P. 43(a) ("At trial, the witnesses' testimony must be taken in open court unless a federal statute, the Federal Rules of Evidence, these rules, or other rules adopted by the Supreme Court provide otherwise."); U.S. CONST. amend. VI ("In all criminal prosecutions, the accused shall enjoy the right . . . to be confronted with the witnesses against him . . . ."). Numerous state constitutions have "open court" clauses. *See, e.g.*, Jonathan M. Hoffman, *By the Course of the Law: The Origins of the Open Courts Clause of State Constitutions*, 74 OR. L. REV. 1279, 1279 (1995) (canvassing the provisions). For a useful overview of the "open courts" doctrine, *see* Michael Pressman & Michael Shammas, *Memorandum: The Permissibility & Constitutionality of Jury Trial by Videoconference*, CIV. JURY PROJECT (May 4, 2020), https://civiljuryproject.law.nyu.edu/memorandum-the-permissibility-constitutionality-of-jury-trial-by-videoconference. So, too, case law is filled with paeans to open-ness. Access to court records, as Justice Holmes once put it, ensures "that those who administer justice should always act under the sense of public responsibility, and that every citizen should be able to satisfy himself with his own eyes as to the mode in which a public duty is performed." Cowley v. Pulsifer, 137 Mass. 392, 394 (1884).

177. For early discussion, *see* Daniel J. Solove, *Access and Aggregation: Public Records, Privacy and the Constitution*, 86 MINN. L. REV. 1137, 1140 (2002).

178. Alexander & Feizollahi, *supra* note 175, at 97; *see also* Engstrom & Gelbach, *supra* note 10, at 1063 ("One problem is that PACER's search interface, which has all the sophistication and user-friendliness of its mid-1990s design, makes it almost useless for data filtering.").

179. *See* AM. ACAD. ARTS & SCI., MEASURING CIVIL JUSTICE FOR ALL 12 (2021) ("Many courts make case-level data available through means only accessible to commercial data aggregators.").

180. *See, e.g.*, CASELAW ACCESS PROJECT, LIBRARY INNOVATION LAB ("The Caselaw Access Project is making all U.S. case law freely accessible online. With the Caselaw Access Project API (CAPAPI) and bulk data service, we can share 40 million pages of published U.S. court cases."), https://lil.law.harvard.edu/projects/caselaw-access-project/; CORNELL, FREE LAW PROJECT, LEGAL INFORMATION INSTITUTE, HTTPS://www.law.cornell.edu/; PUBLIC.RESOURCE.ORG. ("Making Government Information More Accessible"), https://public.resource.org/.

berg.[181] One must get below the surface, into the less visible trenches of dockets, in order to generate many of the most actionable litigation insights.

The relative open-ness of data, however, is only part of the problem. If it is to serve as a robust foundation for software solutions by innovators beyond the courthouse, court data must also be *uniform*.[182] That is, there must be clearly defined standards for how the data is collected, stored, and made available.[183]

As for data that the court doesn't uniquely possess, the legal tech industry may already be slanted toward certain types of litigants. High data costs mean providers cater to corporate interests. Worse, the "haves" of the litigation world, particularly large repeat players within the system, already enjoy privileged access to data.[184] Indeed, recent reports reveal that Walmart and other large companies facing recurring types of litigation—for instance, slip-and-falls and employment disputes—are actively working with large law firms and technology companies to develop a robust suite of legal-tech tools to leverage their own, in-house data.[185] These tools remain proprietary, so details are scarce. But they appear to perform the two core tasks described previously that confer cost and information advantages: predicting case outcomes and automatically generating pleadings (*e.g.*, an answer to a complaint) and papers (*e.g.*, initial discovery requests).

Empirical questions abound about the net effect these imbalances will have on the shape of the legal services industry.[186] Some commentators paint a rosy portrait. Even if the most potent legal tech tools sit for the moment in the hands of the privileged few, that state of affairs

181. *See, e.g.*, David Freeman Engstrom, *The Twiqbal Puzzle and Empirical Study of Civil Procedure*, 65 STAN. L. REV. 1203, 1204 n.7 (2013) (reviewing studies that analyze the systemic impact of Twombly and Iqbal and noting that many fail to go beyond published opinions). Since only a small fraction of cases yield court opinions, the disputes in publicly available data may not be representative of the disputes presented to a model. *Id*. at 1209 n.24 (noting that published opinions, as reflected in Westlaw or Lexis, are only the tip of the litigation iceberg).

182. *See, e.g.*, Colarusso & Rickard, *supra* note 116, at 393 (noting different focus of the open court data and open standards movements).

183. *See* notes 254–58, *infra* and accompanying text.

184. *See* Engstrom & Gelbach, *supra* note 10, at 1018, 1023 (making this point at length); Wolfgang Alschner, *AI and Legal Analytics*, *in* AI AND THE LAW IN CANADA (Teresa Scassa & Florian Martin-Bariteau eds., Mar. 16, 2021) (noting monopoly possession of large legal text corpora by "data haves," particularly "large legal service providers" such as Westlaw and Lexis).

185. *See* Alan Bryan, Patrick DiDomenico, Tariq Abdullah, & James Lee, *Using A.I. to Digitize Lawsuits to Perform Actionable Data Anal*ytics, Presentation at the 2019 Corporate Legal Operations Consortium Vegas Institute (May 15, 2019), https://perma.cc/B4C2- XY3K. For an analysis of the Walmart tool and what it says about the future of legal tech, *see* Engstrom & Engstrom, *supra* note 21.

186. For a masterful overview of the economics of the legal services industry, including discussion of how technology might reshape it, *see* Hadfield, *Legal Markets*, *supra* note 11, at 1265–68.

might not last. Continued proliferation of information-enriching and cost-cutting legal-tech tools might instead level the litigation playing field by allowing smaller law firms to do battle with corporate BigLaw firms.[187] As already noted, the steadily shrinking PeopleLaw industry might rebound.[188]

That said, it is just as easy to paint a darker portrait. Indeed, over the near- to medium-term, a convergence of factors—many of them already noted—may ensure that only litigation's "haves" will be able to develop potent legal-tech applications and gain their advantages. After all, large entities may uniquely have the resources and capital access necessary to build technical capacity.[189] More importantly, there may be no solution to the privileged access that large repeat players currently enjoy. They alone possess the holy grail of case-level outcome data otherwise unavailable within a civil litigation system where secret settlements are the norm.[190] Finally, recall that legal-tech tools derive their value from exclusivity.[191] It may be overly optimistic to think the best legal-tech tools will trickle down from BigLaw firms to smaller, less profitable ones. Providers might seek out new profit streams for their products by offering tiered pricing schemes keyed to the ability to pay, but one can just as easily imagine business models whereby legal-tech purveyors charge a steep premium for their tools, limiting their benefits to only the well-heeled players who can afford their high cost.

Understanding this wider innovation landscape brings the courts' roles as data curators and dispensers into sharp relief. Wise data governance will do more than just make the civil justice system more efficient and promote rule of law values: it also has the capacity to offset the system's predisposition to serving the better off. Court efforts to make data widely available and readily useable can be thought of as an equity-promoting public subsidy. Data governance is not quite

187. *See, e.g.*, Albert H. Yoon, *The Post-Modern Lawyer: Technology and the Democratization of Legal Representation*, 66 U. TORONTO L.J. 456, 457 (2016). For a more general review of assertions about legal tech's democratizing effect, *see* Engstrom & Gelbach, *supra* note 10, at 1031–35, 1037–41.

188. Bill Henderson, *The Decline of the PeopleLaw Sector (037)*, LEGAL EVOLUTION (Nov. 19, 2017), https://www.legalevolution.org/2017/11/decline-peoplelaw-sector-037/.

189. *See generally* Engstrom & Gelbach, *supra* note 10 (cataloging the technical limits of NLP and its requirements of significant technical capacity and significant manual lawyer inputs).

190. For an overview, *see* Christopher R. Drahozal & Laura J. Hines, *Secret Settlement Restrictions and Unintended Consequences*, 54 U. KAN. L. REV. 1457 (2006). As a result, most civil-side cases exit dockets with an uninformative voluntary dismissal under Rule 41 or state equivalents. *See generally* Merritt McAlister, *Missing Decisions*, 169 U. PA. L. REV. 1101 (2021).

191. *See supra* notes 172–74 and accompanying text.

"public option" legal tech, as many have begun to advocate.[192] But it is a critically important way to mitigate distributive concerns and ensure the burgeoning legal tech industry has the necessary materials to serve all segments of the legal services market, not just its upper, most profitable precincts.[193] Even if data governance cannot fully level the litigation playing field, access to data is quickly becoming a core access to justice issue—and a crucial way to counter the privileged position of litigation's "haves."

A final observation can round out our understanding of the stakes: Data accessibility could become essential to expanding access to justice if the legal services industry relaxes the usual regulatory constraints on law practice.[194] From unauthorized practice of law to prohibitions on non-lawyer ownership of firms, these constraints have long stymied innovation. But deregulatory efforts seem poised to spread beyond the relatively few states where they are presently in motion,[195] opening the door to a critical new role for court data dissemination policy. For the moment, nearly all of the new-wave legal tech tools reviewed above serve lawyers. Far fewer serve the self-represented because, as noted previously, they must stop shy of providing legal advice in order to avoid UPL.[196] Instead, legal tech tools that serve the self-represented remain focused on identifying legal issues and guiding litigants into legal help tracks, or providing document assembly services through a "wizard" or guided interview.[197] They do

192. For some, the growing importance of legal tech to the workings of the civil justice system and the significant distributional concerns that arise from uneven and privileged access to data and technical capacity both point in the direction of "public option" legal tech. *See* Livermore, et al., *supra* note 139, at 1228–29,1237 (advocating "government support for the production of public access law search tools—a 'public option' for law search" because law search is a form of public good that is likely to be underproduced if left to private actors, thus underprotecting rule of law values and inefficiently trading off costs and benefits as a welfarist matter); *Id.* at 1237 (advocating a "USLaw.gov" site that serves as a legal materials clearinghouse with a user-friendly API and easy data extraction); Richard M. Re & Alicia Solow-Niederman, *Developing Artificially Intelligent Justice*, 22 STAN. TECH. L. REV. 242, 285 (2019) (proposing "public option" legal tech).

193. Related arguments have been made regarding the Freedom of Information Act, which is overwhelmingly used by corporate interests. *See* Margaret B. Kwoka, *FOIA, Inc.*, 65 DUKE L.J. 1361, 1377–78 (2016).

194. For a full examination of that possibility, *see* Part III, *infra.*

195. *See* notes 267–70, *infra* and accompanying text. As detailed below, those states include Utah and Arizona, with a number of larger states (California, Michigan, Florida) actively considering reforms. For a recent overview, *see generally* ENGSTROM ET AL., *supra* note 15.

196. *See* note 74, *supra* and accompanying text.

197. For a good example of a privately provided (not court-linked) tool, *see* Suffolk Law School, *The Legal Innovation & Technology Lab's Spot API,* https://spot.suffolklitlab.org/ (last visited Jan. 17, 2023). For general discussion of these tools, *see* Cabral et al., *supra* note 41; *see generally* Marc Queudot, Éric Charton, & Marie-Jean Meurs, *Improving Access to Justice with Legal Chatbots*, 3 STATS 356 (2020).

not stray into providing anything that could be construed as providing case-tailored information and so might constitute the practice of law. Generic information outlining procedures and choices, not "effectuation" of legal rights, as some access to justice advocates put it, remains the clear emphasis.[198]

Were it not for regulatory concerns, one could imagine court-disseminated data powering a very different set of answer- and advice-centered tools. Think, for instance, of Q&A systems that offer automated answers to legal questions or lead individuals or entities with civil justice needs through wizards or guided interviews and then, crunching caselaw from the relevant jurisdiction, offer advice about legal options and case prospects.[199] Were the regulatory environment to change, legal tech could do far more than level the litigation playing field between BigLaw and PeopleLaw. It could unleash new software solutions that open entirely new markets, and entirely new types of legal help, for those who need them most.

C. Data Control and the Judicial Governance Challenge

Courts will face numerous governance challenges as their data curation role expands in step with the wider digitization of the legal system. The core tasks are easy enough to state: Courts must get data to the right people, at the right time, in the right amounts, and in the right format. Achieving each of these things, however, is far less simple. It will require beefed-up versions of existing data governance policies and practices plus some new ones, too.

An able but workmanlike literature has already begun to sketch a set of frameworks and best practices for how courts should approach these data dissemination tasks. Top-line governance decisions include whether to make data available at all and, if so, how much and to whom. The excellent *Data Governance Policy Guide* of the National Center for State Courts (NCSC) lays out a spectrum of options for court control over data access, interpretation, and publication.[200] At

198. *See* Sandefur, *supra* note 73, at 3 (surveying legal tech tools that serve the self-represented and finding that "a large component are simply repositories of information" and that only half assist users in taking legal action, such as creating documents, compiling evidence, diagnosing a legal problem, or actually resolving a dispute); Rebecca L. Sandefur, *Legal Advice from Nonlawyers: Consumer Demand, Provider Quality, and Public Harms*, 16 STAN. J. CIV. RTS. & CIV. LIBERTIES 283, 298 (2020) (defining "effectuation").

199. For a state-of-the-art analysis of Q/A systems, *see generally* Vold & Conrad, *supra* note 151. For an example of a more advanced A2J-focused tool in Canada, *see generally* Hannes Westermann, Vern R Walker, Kevin D Ashley, & Karim Benyekhlef, *Using Factors to Predict and Analyze Landlord-Tenant Decisions to Increase Access to Justice*, PROCS. OF THE 17TH INT'L CONF. ON A.I. AND L. 133 (2019).

200. *Data Governance Policy Guide*, *supra* note 24, at 4.

one polar extreme is deep and broad access, but only to a core internal court data team. At the other is a fully open model, with machine-readable data made fully available to internal court users and the public alike. In between are partial-access models, where data is dished out based on roles and permissions, or, as detailed below, with qualifications, such as limits on bulk downloads.[201]

A court's determination about the scope of permissible data access will turn in large part on its commitment to transparency. But there are other considerations, too. Control levels will also turn on data quality and the perceived data literacy of the audience. Low data quality and low literacy creates more opportunities for mistakes, misleading interpretations, and conflicting publication of results.

Privacy and cybersecurity concerns will also loom large.[202] Court documents contain sensitive information, ranging from people's health history to details of private misdeeds. The nefarious uses to which growing pools of court-controlled data can be put are bounded only by the considerable imagination of the surveillance capitalists who can access it.[203] An alarming example comes via widespread reports that a New York City landlord association scraped housing court records to create a blacklist of renters who dared to try to vindicate their rights in housing court.[204] The best way to safeguard privacy, of

---

201. *Id.* at 15.

202. For a sampling of privacy and cybersecurity issues, *see generally* Joint Technology Committee, *JTC Resource Bulletin: Cybersecurity Basics for Courts* (Dec. 4, 2019), https://www.ncsc.org/__data/assets/pdf_file/0014/15251/cybersecurity-2020-01-06.pdf; *see generally* Joint Technology Committee, *JTC Resource Bulletin: Responding to a Cyber Attack (Feb. 17, 2016),* https://www.ncsc.org/__data/assets/pdf_file/0034/18898/responding-to-cyber-attack-2-26-2016-final.pdf; *see generally* Joint Technology Committee, *JTC Resource Bulletin*: *GDPR for US Courts (Sept. 19, 2018)*, https://www.ncsc.org/__data/assets/pdf_file/0024/18726/2018-09-19-gdpr-for-us-courts-final.pdf.

203. The term was made popular in Shoshana Zuboff, The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power (2019).

204. *See* Emily Myers, *What is the Tenant Blacklist? Can it Prevent me from Renting in NYC?* Brick Underground (Oct. 17, 2022), https://www.brickunderground.com/blog/2014/05/tenant_blacklist (describing a tenant blacklist that is provided to landlords who are vetting tenants during the rental application process); Kim Barker & Jessica Silver-Greenberg, *On Tenant Blacklist, Errors and Renters With Little Recourse*, N.Y. Times (Aug. 16, 2016), https://www.nytimes.com/2016/08/17/nyregion/new-york-housing-tenant-blacklist.html (outlining how prospective landlords use the tenant blacklist to "weed out risky tenants"); Ronda Kaysen, *How to Escape the Dreaded 'Tenant Blacklist'*, N.Y. Times (Apr. 13, 2019), https://www.nytimes.com/2019/04/13/realestate/how-to-escape-the-dreaded-tenant-blacklist.html ("There are hundreds of tenant screening bureaus, collecting names from courthouses around the country and selling the information to landlords."). Scholars have only just begun to catalog the types of the sensitive information contained in court records, chief among them locational, identity, health, and financial information as well as past involvement in criminal or civil proceedings. *See* David S. Ardia & Anne Klinefelter, *Privacy and Court Records: An Empirical Study*, 30 Berkeley Tech. L.J. 1807, 1807 (2015). Some are predictable and perhaps even relatively benign in the grand scheme

course, is not to collect and store sensitive data in the first place. But that may not be possible as the system becomes pervasively digitized. Indeed, courts are facing a growing tide of public records requests, backed by an array of sunlight laws, for data they control.[205] While roughly a dozen states exempt courts from such laws, most do not.[206] Nor is it plausible anymore to permit "practical obscurity" to protect privacy interests. Commercial data aggregators have the motivation and the know-how to go digging.[207]

Decisions about control levels and privacy are in substantial part policy judgments—and, in that sense, akin to the traffic rules that will determine which litigants get pushed to ODR. Other core governance tasks, however, are almost entirely technical. For instance, courts must decide how much technology to apply to the process of data collection. New e-filing systems can be both a boon and a bane for data quality. Designed well, e-filing user interfaces can remove clerk or data-entry intermediaries from the equation and yield higher-quality data. Designed poorly, e-filing can result in just the opposite. Other forms of technology, particularly the NLP tools that some courts use to automate data entry by scraping filings and dockets, can improve matters or make them dramatically worse.[208] And even the best automated systems require clerk review and ongoing oversight through audits or data quality reports.

Still other mainline governance challenges have both a political and a technical aspect. Among the most difficult are how to build an infrastructure around data use and dissemination that ensures both the quality and the integrity of data. A first step will be designating personnel, whether individuals or workgroups, with responsibility for each of these things. The NCSC, for instance, recommends a "data governance committee" made up of representatives from the court's research and statistics and IT arms, as well as trial and appellate

---

of surveillance capitalism—for instance, the use of divorce records to market fitness services to newly single women. *See* Karen Gottlieb, *Using Court Record Information for Marketing in the United States: It's Public Information, What's the Problem?*, PRIVACYRIGHTS.ORG (Feb. 1, 2004), https://privacyrights.org/resources/using-court-record-information-marketing-united-states-its-public-information-whats. But others, such as the tenant blacklist, plainly cross lines.

205. *See* National Center for State Courts, *National Open Court Data Standards (NODS)*, https://www.ncsc.org/services-and-experts/areas-of-expertise/court-statistics/national-open-court-data-standards-nods ("Demands for court data are growing dramatically, particularly as courts implement electronic record systems. Both public and private organizations are aggressively putting pressure on courts to make court data and legal documents publicly accessible.").

206. *How Open Is Your Government? Find Out*, MUCKROCK, https://www.muckrock.com/place/ (Last visited Dec. 17, 2020).

207. *See* AM. ACAD. ARTS & SCI., MEASURING CIVIL JUSTICE FOR ALL 12 (2021).

208. *Data Governance Policy Guide*, *supra* note 24, at 11–12.

clerks, judges, court staff (who will be among the primary users), a public information officer, and a legislative liaison.[209] Building this governance structure may also entail designation of key leaders with authority over pieces of data governance: a chief information officer (whose focus is developing the systems necessary to capture, measure, and track data) and a chief data officer (CDO) (whose focus is finding meaning in data).[210] Additional personnel reporting to the CDO would include data quality analysts, who monitor and address data quality issues, and data stewards, who maintain the integrity of a particular dataset.

Also of critical importance is the cost of data access. For example, litigation is steadily opening up Public Access to Court Electronic Records (PACER), including a recent decision by the Fourth Circuit affirming a lower court determination that PACER charges more than necessary to cover operating costs, in violation of the E-Government Act.[211] Tea leaves can also be read in a recent decision of the Supreme Court re-affirming the pithy notion that "no one can own the law" and rejecting Georgia's argument that it could hold copyright to the official annotated state statutes.[212] These decisions can be thought of as a very early shot across the bow in what may soon ripen into a wider reckoning. State legislatures are getting into the act as well. Florida passed a novel law that requires the state's courts to collect data on pretrial release decisions, indigence, ethnicity of parties, and more, then standardize that data and publish it online for free.[213] Many other states are considering following suit.[214]

A particular flashpoint has been whether to make data available in bulk. The federal courts have refused to dispense data except on a fee-per-document basis.[215] Individually, document-based charges are not much, but when combined, they put the large datasets needed to power new legal-tech tools beyond the reach of all but the most well-

209. *Id.* at 6.

210. *Id.* at 8 (defining these roles and further describing the role of a CIO as "technology governance" and the role of a CDO as "data governance").

211. *See* Nat'l Veterans Legal Servs. Program v. United States, Nos. 2019-1081 & 2019-1083, 2020 WL 4516079 (Fed. Cir. Aug. 6, 2020).

212. Georgia v. Public.Resource.Org, Inc., 140 S. Ct. 1498, 1507 (2020).

213. Measures for Justice, *Florida Passes Historic Legislation to Help Close the Criminal Justice Data Gap,* (Mar. 11, 2018), https://www.measuresforjustice.org/news/2018-03-11-florida-passes-historic-legislation).

214. Jason Tashea, *Liberating Criminal Justice Data: How a Florida Law Provides a Blueprint for the Nation*, ABA J. (June 18, 2019), https://www.abajournal.com/web/article/liberating-criminal-justice-data-how-a-florida-law-offers-a-blueprint-for-the-nation.

215. *See* Gelbach, *Free Pacer*, *supra* note 175, 20-22.

heeled entities.[216] Worse, use of automated tools to cut the cost of document-by-document downloads are expressly forbidden, taking away an important way that less-well-heeled data users might economize.[217]

The story is much the same at the state level, where policies range from general availability, to blanket bans, to conditional acceptance.[218] Arizona, an outlier, makes case file records generally available for bulk download.[219] Some states prohibit the dissemination of bulk court records in electronic form, except where explicitly provided by a court rule or order.[220] At least one Florida court only disseminates bulk court data to commercial purchasers.[221] Other states, in contrast, bar the bulk download of court files for commercial gain[222] or charge a fee to cover the cost of bulk data provision,[223] or have terms of service prohibiting the scraping of court websites and/or use the Completely Automated Public Turing test to tell Computers and Humans Apart, commonly known as CAPTCHA, to prevent it.[224]

---

216. *Id*.

217. *Id*.

218. Council for Court Excellence, *Remote Public Access to Electronic Court Records: A Cross-Jurisdictional Review for the D.C. Courts* 10 (Apr. 2017), http://www.courtexcellence.org/uploads/publications/RACER_final_report.pdf ("Only one state allows [bulk downloading] without restriction; nine simply do not allow it at all; and the rest set limits such as only certain data elements, noncommercial users only, or only under contract with a vendor or the court directly.").

219. AZ ST Code of Jud. Admin., § 1-605; Ariz. R. Sup. Ct. 123.

220. Electronic Access Policy for Circuit Court Records of the Illinois Courts § 4.40.

221. *See* MIAMI-DADE COUNTY CLERK OF COURTS, https://www2.miami-dadeclerk.com/Developers/ ("This site is for commercial consumers of the Clerk of Courts data. The services provided include . . . access to bulk data files.").

222. Fine print of homepage, MICHIGAN COURTS RECORDS MANAGEMENT, https://www.courts.michigan.gov/administration/trial-court/trial-court-operations/records-management/ (last visited Feb. 26, 2023) ("Bulk data downloads and commercial uses of the data from this site are prohibited."); Missouri Court Operating Rule 2 Public Access to Record of the Judicial Department 2.10 ("bulk distribution of court records shall be made only upon the approval of the state judicial records committee. Under no circumstances shall bulk distribution of court records be made for commercial gain").

223. Hawaii Court Records Rules 2.5, 10.16, 10.17: "The Administrative Director may grant requests for bulk information from accessible electronic court records provided: (1) the bulk data distribution will not unreasonably interfere with the Judiciary's operations and/or governmental functions; (2) the requester pays all charges for programming the computers, linking systems, and transmitting the data; and (3) when required by law, access is approved by a court of competent jurisdiction."

224. *See Just One Look: Alabama's ON-DEMAND Public Access to Trial Court Records*, https://pa.alacourt.com/default.aspx?loc=Alacourt.gov (last visited Jan. 17, 2023) (illustrating that Alabama charges: $9.99 for a name search that includes one case detail; $9.99 for a case number search that includes one case detail; $5.00 for the first twenty pages of images and $0.50 per page thereafter; and $19.99 for case monitoring for the lifetime of the case if it is a district case and $29.99 if a circuit case); Harris County District Clerk, New User Registration, https://www.hcdistrictclerk.com/eDocs/Secure/Registration.aspx (Jan. 17, 2023); Philadelphia Courts

Even in states that do not bar bulk access to case data, local courts may impose their own rules that limit accessibility[225]—a species of a broader intra-jurisdictional challenge, born of the mix of unified and local control that characterizes most court systems, that we return to below.

Here is where the pandemic may have had a double-edged effect. On the one hand, and as already noted, the pandemic moved many states to adopt e-filing, thickening data flows.[226] At the same time, the pandemic generated budgetary shortfalls and powerful pressures, making it all too tempting for state judicial administrators to monetize their newly digitized records.

Another fundamental governance choice is whether to disseminate data in raw or processed form. Case management data generated and used to process court cases presents relatively few hurdles in this regard. Other types of data, however, are far more complex. For example, data extracted from one or more case management, e-filing, or other administrative systems might need to be manipulated and reformulated to be useful. An important question is the degree to which courts should perform that manipulation or instead make materials available in bulk form so that outside groups, or "justice partners," can put it towards desired purposes. This choice, as with many other data governance decisions, is at least in part a determination as to how much to publicly subsidize data production by putting it in user-ready forms for the benefit of others.

While these challenges—around security, structure, technical design, cost allocation, and form—are plenty difficult, at least two other governance puzzles sit atop them, rounding out the portrait of

---

First Judicial District of Pennsylvania, Philadelphia Municipal Court Electronic Filing System, https://fjdclaims.phila.gov/phmuni/login.do# (last visited Jan. 17, 2023); Washington Courts, Name Search, https://dw.courts.wa.gov/index.cfm?fa=home.namesearch&terms=accept&flash form=0 (last visited Jan. 17, 2023); Baker Clerk, Person Search, https://www.civitekflorida.com/ocrs/app/search.xhtml (last visited Jan. 17, 2023); Arizona Judicial Branch, Public Access to Court Information, https://apps.supremecourt.az.gov/publicaccess/(X(1)S(ncwwtm 45cm24di25fm5t0zvl))/caselookup.aspx?AspxAutoDetectCookieSupport=1 (last visited Jan. 17, 2023); Arizona Judicial Branch, eAccess Terms and Conditions, https://azcourtdocs.gov/arizona/publicTerms.admin (last visited Jan. 17, 2023).

225. For example, in Charleston and Greenville county, South Carolina, in order to access the court records you have to accept a disclaimer that reads in part "Access to the South Carolina Judicial Department Public Index web sites by a site data scraper or any similar software intended to discover and extract data from a website through automated, repetitive querying for the purpose of collecting such data is expressly prohibited." *See Public Index Search*, S.C. JUD. BRANCH, https://jcmsweb.charlestoncounty.org/publicindex/(X(1)S(0r4x5rk4l3rwu0tari1h1tnt))/Disclaimer.aspx?AspxAutoDetectCookieSupport=1, (last visited Jan. 17, 2023).

226. *See* note 62, *supra* and accompanying text.

problems courts must address in their data dispenser role. One is *intra*-jurisdictional. The other is *inter*-jurisdictional.

The *intra*-jurisdictional challenge extends from the architecture of American court governance itself. Many states feature non-unified court systems with local government funding and the devolution of administrative power and control that often come with it. An illuminating example is Michigan, where the state constitution vests administrative authority in the Michigan Supreme Court, exercised through the State Court Administrative Office.[227] Day-to-day operations at the trial court level, however, are given over to administrators acting under the supervision of a local chief judge, selected in nonpartisan elections.[228] Further complications arise from the fact that, by law, the local county clerk (a partisan elected position in the executive branch) is the designated clerk of the circuit court, despite not working for the court system at all.[229] The clerk maintains the court's records, but the court lacks any supervisory authority over the clerk, lending a separation-of-powers component to court data policy. Worse, while state funds power the Supreme Court and Court of Appeals, trial court funding comes from local governmental units and court operations (*e.g.*, costs assessed on litigants). In Michigan and many other states, court governance and court records are a polymorphic mix of unified and local data.[230]

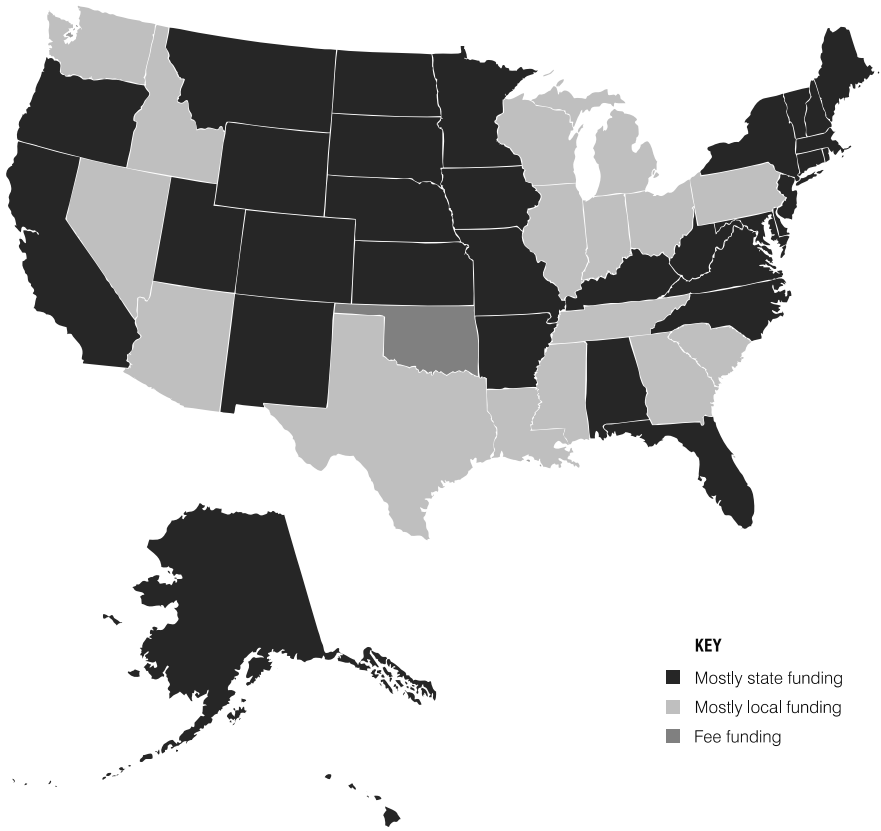
227. MICH. CONST. 1963, Art VI, § 3.

228. Mich. Comp. Laws § 168.412, § 168.467a (2021).

229. Mich. Comp. Laws § 168.200 (2021) (explaining county clerk's role as a county officer).

230. Michigan Courts, "Records Management" page, https://www.courts.michigan.gov/administration/trial-court/trial-court-operations/records-management/ (last visited Jan. 17, 2023).

EXHIBIT 8. Court Funding in the 50 States



Source: The Trial Court Funding Commission Final Report 19 (2019), https://www.michigan.gov/documents/treasury/TCFC_Final_Report_9-6-2019_665923_7.pdf

Such arrangements are not conducive to datafication or the rational development of information-technology systems more generally.[231] Data systems depend on uniformity and comparability, and clunky, ad hoc structures are already proving to be a significant barrier to reform in some states.[232] Variable data standards without common data fields or definitions complicate and even preclude state-level composite reports that spot trends and help craft interventions and responses.

231. Whether unified or non-unified judiciaries function better in other ways—for instance, case-processing efficiency—is an open question. *See* William Raftery, *Efficiency of Unified vs. Non-unified State Judiciaries: An Examination of Court Organizational Performance*, Ph.D. Dissertation, Va. Commonwealth Univ., 1, 44 (2015), https://scholarscompass.vcu.edu/cgi/viewcontent.cgi?article=5117&context=etd; *see also* G. Alan Tarr, *Court Unification and Court Performance: A Preliminary Assessment*, 64 Judicature 356, 356 (1981).

232. *See* Bridget McCormack, *The Disruption We Needed: COVID-19, Technology, and Access to Justice*, *in* Legal Tech and the Future of Civil Justice 307, 310-12 (Engstrom ed., forthcoming 2023).

Some states, particularly those without unified governance systems, are barely out of the blocks with even basic state-level data collection, let alone the standardized state-wide data infrastructures envisioned here. As of 2015, a shocking number of states lacked the capacity necessary to offer simple statistics about judicial administration, including how many cases were filed and disposed of each year.[233]

In Michigan, and elsewhere, workarounds have been found. Data warehouses—a set of databases created to compile, store, and analyze data—are under construction in numerous states.[234] But centralizing data in this way often requires monumental efforts to negotiate data-sharing agreements among local courts in order to coax data from them without impairing their sense of local control. The resulting Memorandums of Understanding (MOUs) often leave at least some constraints on data holdings and their dissemination.[235] As a result, new data-centered initiatives may require time-consuming and innovation-dampening renegotiation and reprogramming.

Interestingly, the pandemic-fueled surge in ODR may play a salutary role. As COVID-19 shuttered courthouses, Michigan courts significantly sped up efforts to implement the state's main ODR platform (called MI-Resolve) in the remaining sixty-six counties throughout the state where it was not yet in place.[236] Reports are that, as the list of courts using MI-Resolve has grown, system-wide change has become easier and more efficient.[237] Local jurisdictions have begun to see the advantages of a unified, state-wide case management system with real-time access to comprehensive court data. Even so,

---

233. Raftery, *supra* note 231, at 73.

234. *See, e.g.*, Proposed Amendment of Rule 2 and Proposed Addition of Rule 21 of the Rules Concerning the State Bar of Michigan and Proposed Amendment of Rule 9.119 and Proposed Addition of Rule 9.1XX of the Michigan Court Rules, http://www.michbar.org/file/barjournal/article/documents/pdf4article4177.pdf ("[T]he Court is considering the adoption of an Administrative Order that would require mandatory submission of case data to the Judicial Data Warehouse."). In addition to Michigan, Florida is developing a data warehouse. *See* Florida Courts Office State Courts Administrator, "Judicial Management Services" page, https://www.flcourts.org/Resources-Services/Court-Services/Judicial-Data-Management-Services-JDMS. Oregon courts have operated a centralized database since 2016. *See* Oregon Judicial Branch, "Oregon eCourt: The Implementation Process," https://www.courts.oregon.gov/programs/ecourt/Pages/default.aspx.

235. *Data Governance Policy Guide*, *supra* note 24, at 6 (noting, among possible complexities, a situation in which a data request for statewide data that resides in a data warehouse encompasses local-level data provided based on an agreement that places constraints on its dissemination).

236. Matterhorn, *MI-Resolve: Michigan Mediation Center ODR*, https://getmatterhorn.com/odr-solutions/civil/odr-for-mediation/mi-resolve-michigan-mediation-center-odr/ (last visited Feb. 6, 2023).

237. *See* McCormack, *supra* note 232.

the roots of existing governance arrangements run deep, and they may yet prove hard to fully dislodge.[238]

The picture is better and worse beyond Michigan. In Connecticut, for instance, the chief court administrator establishes statewide protocols for collecting and storing information.[239] California offers an opposite example; there, courts in different jurisdictions use different e-filing and case management system service providers—some 100 different ones in all.[240] This results in data with starkly incompatible terms siloed in multiple institutions.[241] Even beyond these two states, data accessibility is governed by a huge and bewildering hodgepodge of laws and policies. To be sure, the pandemic may have helped matters. For example, thirteen states did not permit e-filing for self-represented litigants (SRLs) before 2020, but ten of those have since created a mechanism for at least some SRLs to e-file.[242] Still, progress is haphazard: nine states never permitted SRL e-filing in eviction cases and eight have yet to allow SRL e-filing in debt collection cases.[243]

The *inter*-jurisdictional challenge is a spin on many of these same themes, but mapping its contours highlights a final and critically important dimension of the problem. As noted previously, court data, to be fully useable, must be both *open*, in the sense of accessibility, and

238. The hurdles that non-unified control puts in the way of digitization and datafication are certainly not new. Roscoe Pound, in the 1909 Report of Special Committee to Suggest Remedies and Formulate Proposed Laws to Prevent Delay and Unnecessary Cost in Litigation that is widely seen as a founding document for modern judicial administration, decried the lack of empirical policy research on courts. "When "[e]ach clerk's office is independent of every other," he noted, "[it] is no one's duty to study the system, suggest improvements, or enforce them when made." *See* Charles W. Eliot et al., *Preliminary Report on Efficiency in the Administration of Justice, reprinted in* JUDICIAL ADMINISTRATION: TEXT AND READINGS 52 (Russell R. Wheeler & Howard R. Whitcomb eds., 1977). For an excellent overview of struggles over American judicial administration from the Progressive Era to the present, *see* Russell R. Wheeler, *Roscoe Pound and the Evolution of Judicial Administration*, 48 S. TEX. L. REV. 943, 966 (2007). Pound indeed brought American law a general approach to judicial administration with several core principles: minimal court levels and specialization; state-wide supreme court administrative direction; and non-judge administration to superintend statewide court funding, personnel management, and data collection. *Id*. at 958.

239. *About Connecticut Courts: Administration and Operation of the Courts*, CT JUD. BRANCH, https://www.jud.ct.gov/ystday/adminop.html (last visited Jan. 17, 2023).

240. *See Service Providers*, ODYSSEY EFILE CA, http://www.odysseyefileca.com/service-providers.htm. (last visited Jan. 17, 2023).

241. California's many different state courts have different technology systems, and in fact some still use paper systems. *See* Carol A. Corrigan & William R. McGuinness, *Commission on the Future of California's Court System* 214 (Apr. 26, 2017).

242. The Pew Charitable Trusts, *How Courts Embraced Technology, Met the Pandemic Challenge, and Revolutionized Their Operations*, methodological appendix (Dec. 1, 2021), https://www.pewtrusts.org/-/media/assets/2021/11/clsm-court-tech-methodological-appendix.pdf.

243. *Id.* at 3.

also *open-sourced* with *open and uniform standards*, in the sense that it is collected, stored, and made available in a fully public, knowable, and standardized format. Standardization can take simple forms, such as basic tables setting forth common predefined column names, or more complex ones, such as Extensible Markup Language (XML) entities linked with attached data files.[244] Neither approach, however, prevails in the current system. Instead, American courts feature a Babel-like mix of data standards and infrastructure that squelch salutary innovation. The end result is a pervasive collection action problem that strategic court data governance will need to solve.

Data standards that are publicly shared, explained, and available for anyone to use can fuel innovation. In determining the scalability and portability of external interventions, standards sharpen upstream incentives for innovators to invest in building new tools in the first place.[245] Without a set of common standards, innovators must either limit themselves to clusters of similar jurisdictions or else pay a "decipherment tax" of costly data manipulation in order to craft their own interoperable structures.[246] Given this reality, it's little surprise that some of the most successful recent efforts to tap new public sector data streams have come from large tech companies, with their significant technical capacity and their own strategic outlook as data monopolists.[247]

A lack of standardization has particular implications for legal-tech tools that serve the self-represented. When it comes to serving a client base with a limited or nonexistent ability to pay, scale is everything. For-profit and nonprofit entities alike will not invest scarce resources in high-quality software development unless those tools can reach large numbers of people. When it comes to tools serving the worst off, disuniformity barriers extend well beyond data standards. A good example is filing requirements. The fact that many of the roughly 3,000 trial-level courts in the United States impose different filing requirements and forms makes it hard to create apps that generate filing-ready documents for self-represented litigants. The lack of uniform filing requirements—and the tendency of local courts to add ad hoc requirements here and there to serve various ends—is a point-of-im-

244. *See* Colarusso & Rickard, *supra* note 116, at 392.

245. *Id.*

246. *Id.* at 391.

247. *Id.* at 394–95 (recounting successful effort to adopt common data standards for urban transit navigation applications, but perhaps only because of Google's market-dominating involvement and its willingness to invest in a large-scale, jurisdiction-spanning effort).

plementation, "last mile" problem that stymies upstream innovation as much as disuniformity in data standards.[248]

Amidst growing calls for standardization,[249] a range of projects are actively working to solve these challenges in both their intra-jurisdictional and inter-jurisdictional guises. The Suffolk Law Document Assembly Line Project has been collaborating with Massachusetts state courts to provide litigants with interactive court forms coupled with advice about whether, say, a tenant qualifies for eviction protection under federal or state moratoria.[250] A key component of the project is an effort to create a permissive, open-source system that allows remixing of code and common shared elements.[251] Further versions of the project are reportedly going forward in Illinois and Louisiana.[252]

Another marquee effort is the National Open Court Data Standards (NODS) project, a joint initiative of the NCSC and the Conference of State Court Administrators. The NODS project aims to develop a commonly defined set of variables to be included in court data collection and also more technical standards unifying data structure, variable formats, and values.[253] The aim, as the project landing page notes, is "to support the creation, sharing, and integration of court data by ensuring a clear understanding of what court data represent[s] and how court data can be shared in a user-friendly format."[254]

Finally, a project run out of Stanford Law School, the Filing Fairness Project, is mounting an ambitious, multi-jurisdictional effort to simplify and standardize filing systems in areas where unmet civil justice needs are most acute, including evictions, so that scalable technology tools can develop to assist self-represented litigants and legal aid organizations.[255] The goal is to commit several jurisdictions to standardize key parts of their technical infrastructure and simplify their fil-

248. Rochelle Klempner, *The Case for Court-Based Document Assembly Programs: A Review of the New York State Court System's "DIY" Forms*, 41 FORDHAM URB. L.J. 1189, 1221 (2014).

249. *See* Wilf-Townsend, *supra* note 11, at 3 (advocating, among other things, federal funding that brings some degree of uniformity in the storing, accessing, and reporting of case information via standards).

250. *See* Steenhuis & Colarusso, *supra* note 42, at 793–95.

251. *Id*.

252. *Id*. at 801.

253. *See* National Center for State Courts, *National Open Court Data Standards (NODS),* https://www.ncsc.org/services-and-experts/areas-of-expertise/court-statistics/national-open-court-data-standards-nods (last visited Jan. 17, 2023).

254. *Id*.

255. *See* Stanford Law School*, Filing Fairness Project,* https://law.stanford.edu/filing-fairness-project/ (last visited Jan. 17, 2023).

ing systems in order to provide a proof of concept, and an empirical test, of the potential returns to uniformity.

None of these efforts is a sure thing. Standardization—whether of the intra- or inter-jurisdictional variety—is easier said than done, and many prior efforts have tried but failed.[256] The current set of projects, as promising as they may seem, may not come out any better. The Document Assembly Line Project has been a brute-force effort, built on the backs of volunteer coders and spirited along by the pandemic, but it only involves a single state. Without a greater scale, there may be insufficient interest and investment to maintain it. NODS is, like any standard, purely voluntary; courts may or may not adopt it. Finally, the Stanford Project must solve what is ultimately a people problem, and a change-management problem, but coordination costs are high and potentially insurmountable, especially given the problem of vendor "lock-in" noted previously. Indeed, court vendors have powerful incentives to maintain systems built atop their own proprietary technologies and to obstruct efforts that might facilitate interoperability or competitive offerings from other innovators.[257]

The good news is that well-designed standardization efforts may offer a substantial value proposition to courts seeking to build out their data infrastructure. This is because interactive legal applications of various sorts can be integrated into a court's electronic filing and case management systems such that they not only produce PDF forms or provide information to litigants, but also add a digital entry into a

256. *See* Steenhuis & Colarusso, *supra* note 42, at 27 ("Pilot projects have succeeded without yet causing wide-spread lasting change."). A good example of a valiant and longstanding standardization effort that has not substantially moved the dial is the Legal XML project, a non-profit effort to set national data standards that are specific to legal systems, including e-filing. *See* About LegalXML, LEGALXML, http://www.legalxml.org/about/index.shtml (introducing legal consortium sector). LegalXML's plan piggybacks on the architecture developed for the National Information Exchange Model, a successful federal-level collaboration between the Department of Homeland Security and the Department of Justice to promote interoperability of data sets across the courts and federal and state agencies. *See* NIEM's History, NIEM, https://www.niem.gov/about-niem/history (recounting formation of NIEM); Electronic Court Filing Version 4.01 Plus Errata 01, OASIS 9-11 (July 14, 2014), http://docs.oasis-open.org/legalxml-courtfiling/specs/ecf/v4.01/ecf-v4.01-spec/ecf-v4.01spec.pdf (explaining how LegalXML worked with NIEM when developing ECF 4.0, its model for electronic court filing). *See generally* Colarusso & Rickard, *supra* note 116, at 396.

257. Jason Tashea, *The Justice System as a Digital Platform*, THE COMMONS (Sept. 30, 2020), https://wearecommons.us/the-justice-system-as-a-digital-platform/. Think here of TurboTax, which has spent years developing a product that can manage the extraordinary complexity of the nation's tax system (both federal and state), but now has a powerful interest in maintaining a tax system that requires them.

court's electronic docket.[258] As a concrete example, litigant help portals that elicit information from self-represented litigants via expert systems, guided interviews, or "walk throughs" may be most effective if that information can be seamlessly shared with a court's e-filing and case management system. If standardization can be hitched to a mobile-first litigant portal and document assembly service with fully integrated e-filing, internal satisfaction and political support could well prove strong enough to overcome opposition—even where exclusive contracts and proprietary technology have given court vendors a stranglehold on the system.

III. Courts as Data (and Data-Use) Regulators

Post-pandemic courts will be thrust into a final data governance role: as *regulators* of data and, perhaps more accurately, of data *use* by an array of actors beyond court walls. This is keenly apparent in the context of regulating legal services, a duty courts have long delegated to bar associations. Those groups have historically used self-regulating licensing schemes to keep their memberships profitably in demand, but a growing number of experiments in industry oversight may change that. Loosening UPL rules and limits on nonlawyer ownership may enable more data-powered providers to enter the market, improving economics and efficiency for courts and users. Courts will control how these new tools are used, and their choices will have a dynamic effect on the justice system's trajectory.

A. Deregulating Legal Services: How Legal Tech Could Grow

All indications are that one of the mega-trends within the American civil justice system over the next decade will be the steady erosion of the professional monopoly that lawyers have long enjoyed over the delivery of legal services—"[t]he last vestige of the medieval guild system," as a prominent Silicon Valley general counsel put it.[259] In some

258. *See E-filing with Assembly Line*, Document Assembly Line Project, (https://suffolklitlab.org/docassemble-AssemblyLine-documentation/docs/efiling/overview/) (last visited Jan. 17, 2023).

259. For an influential statement, *see* Gillian K. Hadfield & Deborah L. Rhode, *How to Regulate Legal Services to Promote Access, Innovation, and the Quality of Lawyering*, 67 Hastings L.J. 1191, 1192 (2016). *See also* Renee Newman Knake, *The Legal Monopoly*, 93 Wash. L. Rev. 1293, 1293 (2018); Deborah L. Rhode, *Policing the Professional Monopoly: A Constitutional and Empirical Analysis of Unauthorized Practice Prohibitions*, 34 Stan. L. Rev. 1, 6 (1981) (offering an earlier view of the Bar's monopolist tendencies, including its unauthorized practice campaign). For the "medieval" quip, *see* Peter Lattman, *Law Firms: "The Last Vestige of the Medieval Guild System,"* Wall. St. J. (Jan. 29, 2007), https://blogs.wsj.com/law/2007/01/29/ciscos-gc-on-law-firms-the-last-vestige-of-the-medieval-guild-system/.

states, that has meant small steps toward opening the legal system to *human* nonlawyer legal services providers—for instance, permitting non-lawyers, akin to nurse practitioners, to offer legal services in particular substantive legal silos, such as housing, domestic violence, or evictions.[260] Perhaps more importantly, the erosion of the lawyer monopoly will also welcome non-human providers—that is, software applications—more fully into the system. These latter providers run the gamut from Rocket Lawyer and LegalZoom to the previously-noted array of apps that serve the self-represented by offering legal information, constructing filing-ready documents, and creating and preserving evidence.

The most prominent deregulatory efforts to this point have come on other shores. The most well-known is United Kingdom's Legal Services Act (LSA), enacted in 2007. Glossing quite a bit of detail, the LSA relaxed rules prohibiting UPL, non-lawyer ownership of firms, and fee-sharing with nonlawyers, aiming to generate more competition within the legal services market plus innovation in service delivery models.[261] The law created an independent, government-appointed regulatory body, the Legal Services Board, to approve and oversee a phalanx of smaller regulators with the power to license individual providers as well as entities pursuing business models previously barred under the old self-regulatory scheme. The change does not appear to have revolutionized access to justice or innovation in the U.K., with the Board itself concluding consumer price and quality transparency are insufficient.[262] Still, even having a centralized Board-like entity that regularly audits its regulation with empirical surveys[263] is a leap from the disparate, ex post disciplinary and liability frameworks that currently dominate the U.S. market.

---

260. Chambliss, *supra* note 27, at 336–37, 347 (cataloging reform efforts); Gillian K. Hadfield & Deborah L. Rhode, *How to Regulate Legal Services to Promote Access, Innovation, and the Quality of Lawyering*, 67 HASTINGS L.J. 1191, 1220–23 (2016) (same); *see* State Bar of California, *Task Force on Access Through Innovation of Legal Services*, http://www.calbar.ca.gov/About-Us/Who-We-Are/Committees-Commissions/Task-Force-on-Access-Through-Innovation-of-Legal-Services. *See also* Stephen R. Crossland & Paula C. Littlewood, *Washington's Limited License Legal Technician Rule and Pathway to Expanded Access for Consumers*, 122 Dick. L. Rev. 859, 862 (2018).

261. *See* Hadfield & Rhode, *supra* note 260, at 1203 (recounting and assessing the UK reforms); *see also* Hadfield, *Legal Markets*, *supra* note 11, at 1301–02 (offering an updated account).

262. Legal Services Board, *The State of Legal Services 2020,* at 42, https://legalservicesboard.org.uk/wp-content/uploads/2020/11/The-State-of-Legal-Services-Narrative-Volume_Final.pdf.

263. *Id.* at 6.

In 2004, a different method of deregulation occurred in Australia when New South Wales lawmakers liberalized rules permitting lawyers to co-venture with other professionals. Whereas the U.S. regulatory structure seeks to protect consumers by holding individual attorneys to high standards and then preventing anyone else from providing services, the Down Under version relies on an "appropriate management system" in which firm-level management is held to objective standards.[264] These standards cover issues like conflicts of interests, records management, and supervision of non-lawyer staff, among others. Firms must appoint legal practitioner directors that monitor compliance, and failure to do so constitutes professional misconduct. Crucially, this entity-level oversight does not replace individual lawyer regulation. Instead, individuals and entities alike must adhere to the code of conduct and are subject to discipline. Both the Law Society of New South Wales and a legislatively created government office monitor the market.[265]

On U.S. soil, the most salient example of broad-scale deregulation is the Utah Supreme Court's 2019 establishment of a regulatory "sandbox," a space where entities can pilot and evaluate new legal services models under the careful, data-intensive watch of an administrator.[266] Utah's Supreme Court created a regulatory agency, the Office of Legal Services Innovation, to dish out what amount to waivers of the Utah equivalents of Rule 5.4's prohibition on non-lawyer ownership and fee-sharing. It can also waive the bar on UPL for legal services providers who propose new delivery models, subject to case-

264. Susan Sab Fortney, Tahlia Gordon, *Adopting Law Firm Management Systems to Survive and Thrive: A Study of the Ausralian Approach to Management-Based Regulation*, 10 U. OF ST. THOMAS L.J. 152, 153–54 (2012).

265. Section 140(3) Legal Profession Act 2004 (NSW).

266. For an overview of the sandbox model, *see* JTC Quick Response Bulletin, *Fostering Innovation in Legal Services: Testing Legal Regulatory Changes in a Protected "Sandbox"* (2020), https://www.ncsc.org/__data/assets/pdf_file/0018/42813/2020-07-27-QR-sandbox_final.pdf. For Utah-specific details, *see* Utah Work Group on Regulatory Reform, *Narrowing the Access-to-Justice Gap by Reimagining Regulation* (Aug. 2019), https://www.utahbar.org/wp-content/uploads/2019/08/FINAL-Task-Force-Report.pdf. The sandbox approach has begun to appear in numerous policy areas, not just lawyer regulation and access to justice, in response to a growing view that, in an increasingly complex world, policymaking should be iterative and flexible—and specifically structured, at least in part, to answer empirical questions. *See* Hilary J. Allen, *Regulatory Sandboxes*, 87 GEO. WASH. L. REV. 579, 580–81 (2019); *see generally* Chang-Hsien Tsai et. al., *The Diffusion of the Sandbox Approach to Disruptive Innovation and Its Limitations*, 53 CORNELL INT'L L.J. 261 (2020); *see generally* Brian R. Knight, & Trace E. Mitchell, *The Sandbox Paradox: Balancing the Need to Facilitate Innovation with the Risk of Regulatory Privilege*, 72 S.C. L. REV. 445, 446 (2020). For more general overviews of the role of information and learning in policymaking, *see* Yair Listokin, *Learning Through Policy Variation*, 118 YALE L.J. 480, 483 (2008); Michael C. Dorf & Charles F. Sabel, *A Constitution of Democratic Experimentalism*, 98 COLUM. L. REV. 267, 338 (1998).

by-case approval by the state justices.[267] Another recent and much-discussed example is Arizona, where the state's high court opted for a more permanent, non-experimental loosening of the rules governing nonlawyer ownership of law firms by abolishing Rule 5.4 but then mandating that entities with nonlawyer owners apply to the state's high court to be "alternative business structures."[268] As these experiments spread, more software-based providers will gain admission to the legal system. Courts will, as the primary regulators of UPL, increasingly regulate these new legal services providers' use of data and will find themselves routinely evaluating digital and data-based models.

B. Regulatory Reform as a Digitizing Force

Lawyer de-regulation—or, in a more PR-sensitive formulation, "re-regulation"[269]—is moving onto public agendas in light of a growing body of evidence that existing markets for law and legal services are not functioning well in the U.S. and beyond.[270] That is so at least in part, evidence suggests, because of near-exclusive reliance on self-regulation of the legal profession.[271] Self-regulation tends to be stringent, which imposes an inefficient business model on law practice over the short-term and chokes off innovation over the longer-term.[272]

Conventionally understood, professional licensing, particularly the self-regulating variety, brings hard tradeoffs. On the benefit side of

---

267. *See* ENGSTROM ET AL., *supra* note 15, at 16–18 (providing an overview of these regulatory reform efforts). For a concrete example, *see The Office of Legal Services Innovation: An Office of the Utah Supreme Court* (2020), https://utahinnovationoffice.org/.

268. Arizona law defines an ABS as "a business entity that includes nonlawyers who have an economic interest or decision-making authority in the firm and provides legal services in accord with [various Supreme Court-issued rules]." Arizona Code of Judicial Administration ("ACJA") 7-209. Arizona's adopted reforms also approved a new category of nonlawyer licensee, called "Legal Paraprofessionals," who will be able to represent clients in court. For an account, *see* Lyle Moran, Arizona *Approves Nonlawyer Ownership, Nonlawyer Licensees in Access-to-Justice Reforms*, ABA J. (Aug. 28, 2020), https://www.abajournal.com/web/article/arizona-approves-alternative-business-structures-as-part-of-access-to-justice-reforms.

269. This is the chosen term of deregulation's advocates because it de-emphasizes the loosening of constraints. *See* Jayne Reardon, *Re-regulating Lawyers for the 21st Century*, 2CIVILITY (July 18, 2019), https://www.2civility.org/lawyer-regulation-re-regulating-lawyers-for-the-21st-century/.

270. Daniel J. Siegel, *Playing in the Regulatory Sandbox: A Survey of Developments*, AM. BAR ASS'N MAGAZINE (July 1, 2021), https://www.americanbar.org/groups/law_practice/publications/law_practice_magazine/2021/ja21/siegel/.

271. Elizabeth Chambliss, *Evidence-Based Lawyer Regulation*, 97 WASH. U. L. REV. 297, 320 (2019).

272. *See* Hadfield, *Legal Markets*, *supra* note 11, at 1.

the ledger is a potential solution to the lemons problem.[273] Extending from Akerlof's classic studies of used cars and the effect of uncertainty in secondary markets for consumer goods, legal services are classic credence goods—that is, their quality is largely opaque to their consumers.[274] Where consumers cannot reliably police quality, regulators can set minimal skill and competency requirements for market entry and then police the ensuing service delivery in the face of persistent market pressure to shirk on quality or deliver unneeded services. Licensing of the self-regulation sort might be especially good at performing these core regulatory tasks because it is more expertly administered by those with professional training and experience. This may also make it less costly than outsider-controlled regulation.[275]

Sitting on the cost side of the ledger, however, is a long and growing bill of particulars. Current regulation in the United States is so restrictive that the market for legal services is almost entirely a market for lawyers, as maintained by a formidable set of command-and-control-style constraints on market entry. For instance, UPL constraints ensure that a lawyer license is necessary to perform virtually any law-related service, defined as any advice requiring the application of legal knowledge to a person's or entity's particular circumstances.[276] As noted previously, only "scrivener" duties—for instance, document assembly—avoids UPL's clutches. [277] In addition, the bar on nonlawyer ownership of firms (and, relatedly, "fee-splitting" between lawyers and nonlawyers) limits the types of organizational forms that provide legal services to the law firm partnership.[278] Legal services are not, and may not be, delivered by corporations, by so-called "multi-disciplinary practices" that combine different types of professionals (law-

273. *See* George A. Akerlof, *The Market for "Lemons": Quality Uncertainty and the Market Mechanism*, 84 QUART. J. ECON. 488, 488–89 (1970).

274. *See* Nora Freeman Engstrom, *Attorney Advertising and the Contingency Fee Cost Paradox*, 65 STAN. L. REV. 633, 673 (2013) (explaining "credence good" concept and applying it to law).

275. Michael J. Trebilcock, *Regulating Service Quality in Professional Markets*, in THE REGULATION OF QUALITY (Donald N. Dewees, ed., 1983).

276. Deborah L. Rhode & Lucy Buford Ricca, *Protecting the Profession or the Public? Rethinking Unauthorized-Practice Enforcement*, 82 FORDHAM L. REV. 2587, 2589 (2014).

277. Even document assembly services have faced UPL criticism. *See generally* Susan Saab Fortney, *Online Legal Document Providers and the Public Interest: Using a Certification Approach to Balance Access to Justice and Public Protection*, 72 OKLA. L. R. 91, 92–94 (2019) (discussing early Texas state court ruling that a 1990s-era document assembly provider violated UPL laws).

278. *See, e.g.*, Ohio Prof. Cond. R. 5.4(b) (lawyer shall not form a partnership with a nonlawyer if any of the activities of the partnership consist of the practice of law).

yers, accountants, social workers, etc.) under a single roof,[279] or even by non-profit entities if not operated and financed by lawyers. Also inhibiting are bar restrictions on recommendations, referrals, and advertising—contributing to the striking finding that the majority of an average lawyer's time is currently spent on administrative tasks and client acquisition, not on performing billable legal work.[280]

The consequences of these command-and-control restrictions on market entry are manifold. For starters, it remains at least possible that monopolized legal markets can help solve the lemons problem by enforcing minimal competence. However, the current self-regulatory scheme, with its hefty entry requirements (seven years of higher education, bar licensure, continuing education requirements) and pervasive constraints on practice, financing, and advertising have catapulted the price at which lawyers can provide legal services well beyond the price most Americans' can pay.[281] As already noted at length, most Americans cannot even afford services for what are thought to be fairly straightforward litigation matters, from real property to family to employment disputes.[282] PeopleLaw—to return to Part II's slicing and dicing of the legal services industry—is shrinking at least in part because its lawyers have no plausible business model for serving the overwhelming majority of potential clients with civil justice needs that that cluster in low- to middle-income demographics and have only very limited ability to pay.

Even more worrying in a fast-digitizing litigation system are the effects on innovation. Limits on nonlawyer ownership of firms and fee-splitting choke off capital flows by prohibiting incentive contracts and revenue-sharing with nonlawyers, requiring that investments in technology instead come from withheld profits, partner capital contributions, and conventional loans.[283] Constrained capital access, when

279. *See* Louise Lark Hill, *Alternative Business Structures for Lawyers and Law Firms: A View from the Global Legal Services Market*, 18 OREGON REV. OF INT'L L. 135, 136 (2017).

280. Recent studies suggest that lawyers perform substantive, billable work for as little as 2 or 3 hours per day. The rest of their time is spent on administrative tasks, some of them directly related to regulatory requirements and client acquisition, which is made harder by the pervasive constraints on legal advertising and referrals and recommendations. *See* Hadfield, *Legal Markets*, *supra* note 11, at 1293 (cataloguing these constraints); Victor Li, *Lawyers are only billing a fraction of their time; how can they be more efficient?*, ABA JOURNAL (Sept. 20, 2016), https://www.abajournal.com/news/article/lawyers_are_only_billing_a_fraction_of_their_time_how_can_they_be_more_effi/.

281. The best reading of a large empirical literature is that occupational licensing "raises prices but may or may not shore up quality." Hadfield, *Legal Markets*, *supra* note 11, at 34 (offering extensive review of the literature).

282. *See* notes 11–14, *supra* and accompanying text.

283. *See, e.g.*, Hadfield, *Legal Markets*, *supra* note 11 at 37 ("By cutting law off from capital markets—especially from venture capital—professional regulation cuts law off from innovation.

combined with prohibitions on non-compete agreements, also contribute to the growing fragility of the leveraged-partner BigLaw business model. The highly portable nature of legal talent—star partners can "grab and go" with impunity—creates an "elastic tournament" for human capital that leads risk-averse law firms to emphasize short-term partner pay-outs over long-term investment in new and potentially transformative technologies.[284] Finally, standardized—and, indeed, ruthlessly homogenized—training and licensure processes bring efficiencies via shared knowledge and norms but also sacrifice the innovative spark that comes from diverse viewpoints.[285]

With so few actual experiments outside of the various deregulatory efforts in the United Kingdom, Australia, Utah, and Arizona noted previously, courts lack the critical mass of data necessary to test key empirical premises about the degree to which technological and organizational innovations could lower the effective cost of legal services.[286] Weighing the conventionally-voiced benefits of a self-regulating legal profession (expert and lower-cost setting and enforcement of quality standards) against its costs (high prices, supply restrictions, capital

---

The problem is not that lawyers are inherently risk averse; it is that risk aversion makes sense for undiversified investors"). As just one example, law firms may not enter into profit-sharing contracts with technology companies and instead must pursue technological innovations through only fee-for-service contracts. The exception to this general state of affairs, of course, is a growing litigation finance industry which funds lawyers and clients by covering litigation costs in exchange for a piece of the recovery, typically structured as a non-recourse loan. *See generally* Velchik & Zhang, *supra* note 171.

284. *See* Marc Galanter & William D. Henderson, *The Elastic Tournament: A Second Transformation of the Big Law Firm*, 60 STAN. L. REV. 1867, 1867, 1871–72 (2008); Larry E. Ribstein, *The Death of Big Law*, 2010 WISC. L. REV. 749, 751–52, 759; Jonathan Levin & Steven Tadelis, *Profit Sharing and The Role of Professional Partnerships*, QUART. J. OF ECON. 131, 133 (2005); James B. Rebitzer & Lowell J. Taylor, *When Knowledge is an Asset: Explaining the Organizational Structure of Large Law Firms*, 25 J. LAB. ECON. 201, 203 (2007); John Morley, *Why Law Firms Collapse*, 75 BUS. LAWYER 1399, 1412–13 (2020). A further innovation-suppressing aspect of the partnership model is that senior partners disproportionately hold decision-making power but have systematically shorter time horizons for realizing profits from investments in innovation. *See id.* at 1414.

285. *See* Hadfield, *Legal Markets*, *supra* note 11, at 1296. *But see* Catherine Beaudry & Andrea Schiffauerova, *Who's right, Marshall or Jacobs? The localization versus urbanization debate*, 38 RESEARCH POLICY 318, 319 (Dec. 30, 2008) (finding mixed evidence regarding the relationship of specialization and diversity to innovation, in that returns to specialization might be highest in low-tech mature industries, which may most fairly describe the legal services market). The partnership form into which all legal practice is shoehorned also brings inherent limits on scale—in a nutshell, smaller firms are preferred because they limit the liability risk of low-quality partners—and, by limiting scale, also caps the all-important scalability of innovations. Branding limits scalability as well because a partnership cannot offer one-to-many services through a branded website unless it is owned entirely by lawyers. Hadfield, *Legal Markets*, supra note 11, at 1297–99.

286. For a full discussion of the current state of empirical evidence, *see* ENGSTROM ET AL., *supra* note 15, at 19–21.

constraints, and partnership-based risk-aversion and short-termism) is hard enough.[287] Doing so in a rapidly changing and fast-digitizing civil justice system is still harder and will not yield anything resembling definitive conclusions anytime soon.

Still, a growing consensus asserts that deregulation can help dent the access crisis—or, at the very least, is worth trying. One reason the tide may be turning is the increasing viability of software-based delivery models. As already noted, a growing menu of legal-tech tools stops shy of providing anything resembling legal advice, focusing instead on moving individuals and small businesses into legal help tracks or helping them assemble filing-ready legal documents.[288] Even given these limitations, legal-tech companies have begun to achieve impressive scale. LegalZoom, for example, has steadily cornered the market on small-business formation after carving out a truce in its regulatory battles with state bar associations, as detailed previously.[289] The company recently completed an initial public offering based on what the prospectus claimed is $50 billion in reachable markets—and plenty more in the event of widescale regulatory reform.[290] Just as important, technology continues to leap forward—particularly, the natural language machine learning models at legal tech's core—increasing software's capacity to perform higher-order legal cognitions.[291] This progress may soon render the line between human and non-human legal services more a creature of regulation than of legitimate consumer protection concerns.

Political constraints hardly help matters. Washington State's high-profile effort to create a non-lawyer licensure system, the Limited License Legal Technicians program, quickly faltered and was unable to attract significant enough numbers to move the dial on access is-

287. Even before the disruptions of the current digital era, the academic literature on self-regulation was largely theoretical. *See, e.g.*, Peter M. DeMarzo, Michael J. Fishman, & Kathleen M. Hagerty, *Self-Regulation and Government Oversight*, 72 REV. OF ECON. STUDS. 687, 687–88 (2005) (building a theory of why a monopolist self-regulating organization will choose a lax and suboptimal level of enforcement effort to root out fraud than the framework that consumers would choose); Avner Shaked & John Sutton, *The Self-Regulating Profession*, 48 REV. ECON. STUDS. 217, 233 (1981) (same—using economic modeling to show that a self-regulating profession will impose quality requirements above the socially optimal level, yielding a profession is too small and generates insufficient supply).

288. *See* notes 71–75 *supra* and accompanying text.

289. *See* Teresa J. Schmid, *LegalZoom Continues to Redefine Legal Services Market*, YOURABA (July 5, 2021), https://www.americanbar.org/news/abanews/publications/youraba/2021/0705/legal-zoom/.

290. *Id.*

291. *See* Engstrom & Gelbach, *supra* note 10, at 1020–26.

sues.[292] The program sunsetted in 2020, drawing howls from access to justice advocates that a turf-protecting bar had set entry requirements far too high and, in any event, that its premature termination had come well before it could yield a reliable empirical portrait of its potential.[293] Similarly, a California State Bar effort to design a regulatory sandbox was unceremoniously shut down after it drew vociferous criticism from powerful lawmakers voicing conventional worries that deregulation will lead to low-quality services as well as darker concerns over the creeping colonization of the legal services industry by Big Tech.[294] And in New York, litigation asserting a First Amendment right to provide legal advice has faced stout opposition from the state attorney general.[295]

One thing, however, seems certain. In a world where self-regulation of the legal profession is beginning to loosen, courts will increasingly be thrust into the role of evaluating and certifying alternative legal services delivery models.[296] In the process, courts will not just be data users or data dispensers—they will increasingly regulate *others'* use of data and data analytics.

C. Legal Services Regulation and the Judicial Governance Challenge

These new regulatory duties present daunting governance challenges beyond what courts will face in their roles as data users and data dispensers. Those new governance challenges will begin with the institutional design of the new regulatory structures themselves. For

292. Lyle Moran, *How the Washington Supreme Court's LLLT Program Met Its Demise*, ABA JOURNAL (July 9, 2020), https://www.abajournal.com/web/article/how-washingtons-limited-license-legal-technician-program-met-its-demise.

293. Stephen Daniels & James Bowers, *Alternative Legal Professionals and Access to Justice: Failure, Success, and the Evolving Influence of the Washington State LLLT Program (The Genie is Out of the Bottle)*, 71 DEPAUL L. REV. 227, 258–261 (2022). Interestingly, deregulation may be happening in any event, at the hands of overwhelmed state and local judges. *See* Steinberg et al., *Judges and the Deregulation of Lawyers*, *supra* note 27, at 1315–16 (describing judge practices of explicitly tapping nonlawyer advocates from the bench and describing it as an implicit form of deregulation).

294. David Freeman Engstrom & Nora Freeman Engstrom, *Why Do Blue States Keep Prioritizing Lawyers Over Low-Income Americans?*, SLATE (Oct. 17, 2022), https://slate.com/news-and-politics/2022/10/blue-states-legal-services-lawyers-fail.html; Letter re: *Legislative Concerns Regarding the Closing the Justice Gap Working Group* from Mark Stone, Chair, Assembly Comm. on Judiciary, & Tom Umberg, Chair, Senate Comm. on Judiciary, to Ruben Duran, Chair, Board of Trustees, State Bar of California (Dec. 7, 2021), https://s3.documentcloud.org/documents/21151650/state-bars-ctjg-concerns-12-7-21.pdf.

295. In early 2022, a federal judge granted an injunction to allow Upsolve, a nonprofit bankruptcy legal services firm, train nonlawyers to give legal advice. *See* Upsolve, Inc. v. James, No. 22-cv-627 (PAC), 2022 WL 1639554, at *17-18 (S.D.N.Y. May 24, 2022).

296. Hadfield & Rhode, *supra* note 260, at 1215.

instance, if a regulatory entity is created to exercise gatekeeper power over which sandbox or alternative business structure applicants gain entry to the system, should it sit within the judicial branch and report directly to the Supreme Court, or should it instead be its own legislatively enacted governmental entity, with a board that serves as an auxiliary to the Supreme Court, as with many state bar associations?[297] Those choices will determine how direct the judicial regulatory role is—and how much regulators can build on the technical capacity that already exists.

A second flashpoint in the design process in California and elsewhere has been the appropriate scope of reforms. For instance, should a sandbox take all comers, or should it instead be limited to specific types of legal services providers (*e.g.*, non-profits rather than for-profits), specific areas of law (*e.g.*, evictions or debt collections), or provider models that are explicitly designed to serve underserved or disadvantaged groups? "Wide open" adherents emphasize innovation and information. They see the sandbox as, at bottom, an effort to spur desired innovation—and, perhaps more importantly, to generate more and better information than currently known about *how* to spur desired innovation and where consumer harm most likely occurs. If current restrictions are relaxed, would the best innovation take the form of human nonlawyer providers, software, or a mix? Just how much do nonlawyer ownership rules hinder capital flows that might underwrite innovation efforts? In what substantive legal silos—evictions, consumer credit, family law, or something else—will innovations tend to cluster? How much innovation will issue from the non-profit sector and how much from the for-profit sector? Will for-profit entities develop delivery models that steadily filter down from better-heeled client bases to needier ones? Which delivery models and what substantive legal areas will yield the most worrying forms of consumer harm? A sandbox with narrow entry requirements may not give policymakers a view across the full spectrum of potential providers or generate useable information about where innovation is likely to arise and where customer harm is most likely to present.

A third and related design debate is the relative costs and benefits of front-end, prescriptive entry requirements to shape innovation and guard against consumer harm, as compared with ongoing oversight and policing of new legal services providers admitted through a less

297. *See generally* ENGSTROM ET AL., *supra* note 15.

stringent screening process.[298] Back-end oversight might be thought to be more pro-innovation. After all, upfront, prescriptive rules make assumptions about where innovation will come from and where and how the most worrying types of consumer harm will present. The sandbox model is predicated, at least in part, on the notion that consumer harm is both a concern to be mitigated and also one of the core empirical questions that a sandbox is designed to test. But back-end oversight also depends on a robust regulatory scheme to perform needed monitoring and enforcement in order to limit consumer harm and, where necessary, to adjudicate ouster from the system.

These are important design choices. Sandbox design, or the creation of new ABS licensure processes, will bring many more. It remains to be seen, of course, which of these approaches, and which institutional designs, will ultimately win out. But lurking beneath the surface of debates over how to build a sandbox or ABS licensure system sit a deeper and more novel set of governance challenges—how best to build the regulatory apparatus necessary to oversee new digital legal services providers that will define the civil justice system's next iteration.

For starters, courts will be more regularly brought into regulation of the business of law, and this new regulatory role will very often be both entity-based and technical. As Professor Bill Henderson, a leading voice on lawyers and technology, aptly observes, the U.S. system has long been "designed to guard against individual lawyer impropriety."[299] Under the current system of self-regulated bar discipline, courts are only episodically involved as final arbiters in a process that individually reviews legal representations for consistency with legal-ethical duties—that is, whether an attorney breached her duties under the rules of professional responsibility while delivering legal services to her clients. Going forward, however, regulatory oversight will, as in the U.K. and Australian systems, increasingly focus on consumer welfare, and it will increasingly police lawyer impropriety through entity-level regulation.[300] That regulation will also be direct and regular, not episodic. Indeed, many of the proposed deregulatory frameworks—including those in place in Utah and Arizona as well as those under consideration in California, Florida, and Michigan—require the court

298. For a useful examination of the choice between ex ante, prescriptive regulation of entry and ex post regulation of effects, *see* Samuel Issacharoff, *Regulating After the Fact*, 56 DEPAUL L. REV. 375, 377 (2007).

299. *See* William D. Henderson, *Legal Market Landscape Report*, STATE BAR OF CALIFORNIA 27 (July 2018), https://board.calbar.ca.gov/docs/agendaItem/Public/agendaitem1000022382.pdf.

300. *Id.*

to directly review all legal services providers seeking sandbox entry or ABS status.

This new entity-focused gatekeeper role will put courts in a regulatory posture that is very different from their current role of overseeing and regulating law practice. Above all else, it will require an entirely new regulatory apparatus from the one that has been built up over time within the lawyer disciplinary system. Entry determinations will necessarily require that courts understand and assess a legal services provider's organizational structure and its bureaucratic routines and practices, as will ongoing, post-entry oversight. Indeed, after new providers have begun to populate the system, monitoring and enforcement efforts will most likely be a mix of outcome- and output-focused regulation, as built around efforts to gauge the incidence of consumer harm by a given provider, and more process-oriented approaches, as focused on a provider's own managerial and other systems for identifying and resolving complaints.

The shift in focus to organizational routines and practices will also mean that the courts' new gatekeeper role will be far more technical in nature. This is true in at least two senses. First, because many digital providers will operate at scale, their evaluation and regulation by courts will require new data-based metrics for gauging provider quality and consumer harm. In Utah, entrants to the sandbox must make disclosures on the front end, including information about the technical guts of their software systems and the data on which it relies. From there, entrants must turn over data on representations so that the sandbox entity—and, once again, ultimately the Utah Supreme Court—can gauge the degree, if any, of consumer harm.

Second, entity-level analysis, whether *ex ante* entry or *ex post* oversight, will necessarily bring courts into dialogue with a host of data infrastructure and cybersecurity practices like never before. Of course, even without the current reform efforts, courts are destined to perform more technical modes of analysis anyway. It is inevitable in a fast-digitizing legal system that questions about legal ethics and professional duties will extend into technical domains. A wide range of rules, from Rule 1.6's duty of confidentiality and Rule 1.1's duty of competence to the various rules governing the use of information from past clients (Rule 1.9), current clients (Rule 1.8), or prospective clients (Rule 1.18), already extend to data privacy and cybersecurity practices.[301] This trend will only increase as the legal system digitizes.

301. *See, e.g.*, Model Rules of Pro. Conduct r. 1.6 cmt. 18 (Am. Bar Ass'n 1983) (discussing data privacy); Model Rules of Pro. Conduct r. 1.1; Model Rules of Pro. Conduct r. 1.9; Model Rules of Pro. Conduct r. 1.8; Model Rules of Pro. Conduct r. 1.18.

The ever-present threat of hacks and data breaches means that courts, even in their more conventional role as final arbiters of bar disciplinary proceedings, will at times sit in judgment on technical matters.

Moreover, many tech-forward BigLaw firms are expressly trying to move what they do along the spectrum from a *service*, provided as highly bespoke legal counsel, to a *product*, using their goodwill and expertise to develop software platforms that provide one-to-many legal services. Wilson Sonsini, as just one example, recently formed SixFifty, a technology company that offers product-based legal services to mid-sized businesses.[302] Even without deregulation, the years to come will see increasing consideration of firm practices, particularly around the use of client data, that will press on existing ethical rules. Courts will determine how those developments are monitored and improved over time.

Somewhat ironically, deregulation promises to put judicial oversight and regulation of law practice on steroids. For instance, ongoing oversight of new legal services providers may well require courts, whether as an initial matter or sitting in review of a court-created regulator's actions, to determine which business practices amount to "dark patterns"—that is, the use of design interfaces to manipulate consumer preferences.[303] The larger problem, it should not be hard to see, is that most courts currently lack anything resembling a conventional policy apparatus that might perform these tasks.[304]

Conclusion

This Article has toured the ways courts are already evolving into powerful data governors by using data, controlling access to it, and regulating its use by others. As the COVID pandemic has receded, multiple states have put out data-driven reports analyzing their recent transformations. They all reflect the importance data governance will play—and, indeed, already plays—in the civil justice system. For ex-

302. *SixFifty Releases Free Tool To Help Renters Avoid Eviction by Generating Necessary Declaration To Invoke Trump Administration's New Order,* SixFifty (Sept. 10, 2020), https://www.sixfifty.com/sixfifty-releases-free-tool-to-help-renters-avoid-eviction-by-generating-necessary-declaration-to-invoke-trump-administrations-new-order/.

303. For more on dark patterns, *see generally* Jamie Luguri & Lior Jacob Strahilevitz, *Shining a Light on Dark Patterns*, 13 J. Leg. Analysis 43 (2021). Another and perhaps better label is "digital market manipulation." *See* Ryan Calo, *Digital Market Manipulation*, 82 Geo. Wash. L. Rev. 995, 995 (2014). This alternative label makes clear that internet-scale data and a growing science of human-computer interaction makes possible new forms of manipulation that may or may not be of a piece with past, analog forms of mass manipulation, from Madison Avenue to the advent of mass-produced political propaganda at the middle of the last century.

304. *See* Hadfield & Rhode, *supra* note 260, at 1217 (discussing the challenges of American court organization).

ample, New York's Trials Working Group recently called for extensive data collection and review "for the purpose of assessing the success or failure of any efforts to reduce judicial backlog."[305] An Arizona Supreme Court working group highlighted the data it collected through new technology that uses optical character recognition (OCR) and AI to extract text from document images and transform them into metadata, automating workflows and integrating with case management systems.[306] Ohio's Supreme Court Task Force on Improving Court Operations all but begged for some form of intra-jurisdictional "data standardization."[307]

While such reports show promising awareness about the growing importance of court data governance, they only scratch the surface of what a fully digitized civil justice system may someday look like. The rapid transformation that the pandemic accelerated requires a diligent effort to transform accordingly; the stakes for the system's legitimacy are high. Courts need strong governance policies to plan accordingly without wasting scarce time and resources by missing key reforms. Worse, bad governance might stymie market-driven innovation from beyond the court walls, creating a vicious cycle whereby the civil justice system antiquates rather than innovates. This would leave courts out-gunned by the superior and ever-increasing technical capacities of many of the litigants before them, who will be able to predict how judges will rule, perhaps even before the judges themselves know.[308]

With smart data governance, however, courts can leverage growing streams of data to administer and manage the system more empirically. Good data governance can also tap the ingenuity and resources of justice partners beyond court walls, facilitating private development of tools that assist the self-represented and a growing legal-tech tool-

305. Future Trials Working Group of the Commission to Reimagine of the Future of New York's Courts, *Report and Recommendations of the Future Trials Working Group* 6 (Apr. 2021), https://www.nycourts.gov/whatsnew/pdf/future-trials-working-grp-april2021.pdf.

306. Arizona Supreme Court, *Post-Pandemic Recommendations: COVID-19 Continuity of Court operations During a Public Health Emergency Working Group* 24 (June 2, 2021), https://www.azcourts.gov/Portals/216/Pandemic/2021/Post-PandemicRecommendations.pdf?ver=2021-06-08-192520-583.

307. The comments referred to the Ohio Courts Network (OCN), which allows courts to query a centralized statewide database of records and provides users with a "one-stop shop for person-level data stored in other state databases." Identifying the local court level's "lack of data standardization" as the key barrier, the task force called for an assessment of the data elements being collected, expansion of collection efforts, and collaboration with external partners. The Supreme Court of Ohio, *Task Force on Improving Court Operations Using Remote Technology* 39 (2021), https://www.supremecourt.ohio.gov/docs/Boards/iCourt/ReportVolumeI.pdf.

308. *See* Engstrom & Gelbach, *supra* note 10, at 1075 (sketching several possible futures for legal tech, including a scenario in which a core set of litigants enjoy superior technical capacities relative to judges and their litigation opponents).

kit that augments the work of lawyers. A digitized system with data as its lifeblood might yet be a system that operates more efficiently and delivers a more equitable civil justice experience. Commentators have long seen the public-private technology gap as a core problem for the regulatory state,[309] and this is no less true for courts. Hopefully this Article's tour of the courts' new data governance roles can help lay a foundation for some of the vital work to come.

309. *See, e.g.*, CHOPRA, *supra* note 115.