

Litigation Outcome Prediction, Access to Justice, and Legal Endogeneity

Charlotte S. Alexander

The United States has a serious and persistent civil justice gap. In 1994, an American Bar Association study found that half of low- and moderate-income households had faced at least one recent civil legal problem, but only one-quarter to one-third turned to the justice system.¹ Twenty-four years later, a 2017 study by the country's largest civil legal aid funder found that 71 percent of low-income households surveyed had experienced a civil legal need in the past year, but 86 percent of those problems received "inadequate or no legal help."² Studies in individual states tell a similar story.³

Unmet civil legal needs include a variety of high-stakes case types that affect basic safety, stability, and well-being: domestic violence restraining orders; health insurance coverage disputes; debt collection and relief actions; evictions and foreclosures; child support and custody cases; and education- and disability-related claims.⁴ There is generally no legal right to counsel in these cases, and there are too few lawyers

Thanks to Madison Gibbs for her excellent research assistance and to Albert Yoon, Dan Linna, David Freeman Engstrom, Peter Molnar, and Sanjay Srivastava for insightful conversations on these topics.

¹ AM. BAR ASS'N, CONSORTIUM ON LEGAL SERVICES AND THE PUBLIC, LEGAL NEEDS AND CIVIL JUSTICE: A SURVEY OF AMERICANS, MAJOR FINDINGS FROM THE COMPREHENSIVE LEGAL NEEDS STUDY 27 (1994), <https://legalaidresearch.org/2020/03/03/legal-needs-and-civil-justice-a-survey-of-americans-major-findings-from-the-comprehensive-legal-needs-study/>.

² LEGAL SERVS. CORP., THE JUSTICE GAP: MEASURING THE UNMET CIVIL LEGAL NEEDS OF LOW-INCOME AMERICANS 6 (2017), <https://www.lsc.gov/our-impact/publications/other-publications-and-reports/justice-gap-report>; see also generally Rebecca L. Sandefur & James Teufel, *Assessing America's Access to Civil Justice Crisis*, 11 U.C. IRVINE L. REV. 753 (2021).

³ See, e.g., N.C. Equal Access to Justice Comm'n & NC Equal Justice All., *In Pursuit of Justice, An Assessment of the Civil Legal Needs of North Carolina* 4 (2021), <https://nccqualjusticealliance.org/assessment/>; VICTOR D. QUINTANILLA & RACHEL THELIN, INDIANA CIVIL LEGAL NEEDS STUDY AND LEGAL AID SYSTEM SCAN 6 (2019), <https://www.repository.law.indiana.edu/facbooks/206/>; LEGAL SERVS. CORP., THE JUSTICE GAP 53 n.6 (collecting additional state studies).

⁴ LEGAL SERVS. CORP., THE JUSTICE GAP, at 7.

willing and able to offer representation at prices that low- and middle-income clients can afford.⁵ In my home state of Georgia, for example, five or six rural counties – depending on the year – have *no resident attorneys*, and eighteen counties have only one or two.⁶ These counties’ upper-income residents travel to the state’s urban centers for legal representation. Lower-income residents seek help from rotating legal aid lawyers who “ride circuit,” meeting clients for, say, two hours at the public library on the first Wednesday of the month.⁷ Or they go without.

Can computationally driven litigation outcome prediction tools fill the civil justice gap? Maybe.

This chapter reviews the current state of outcome prediction tools and maps the ways they might affect the civil justice system. In Section 7.1, I define “computationally driven litigation outcome prediction tools” and explain how they work to forecast outcomes in civil cases. Section 7.2 outlines the theory: the potential for such tools to reduce uncertainty, thereby reducing the cost of civil legal services and helping to address unmet legal needs. Section 7.3 surveys the work that has been done thus far by academics, in commercial applications, and in the specific context of civil legal services for low- and middle-income litigants. Litigation outcome prediction has not reached maturity as a field, and Section 7.4 catalogs the data, methodological, and financial limits that have impeded development in general and the potential to expand access to justice in particular.

Section 7.5 steps back and confronts the deeper effects and the possible unintended consequences of the tools’ continued proliferation. In particular, I suggest that, even if all the problems identified in Section 7.4 can be solved and litigation outcome prediction tools can be made to work perfectly, their use raises important endogeneity concerns. Computationally driven tools might reify previous patterns, lock out litigants whose claims are novel or boundary-pushing, and shut down the innovative and flexible nature of common law reasoning. Section 7.6 closes by offering a set of proposals to stave off these risks.

Admittedly, the field of litigation prediction is not yet revolutionizing civil justice, whether for good or ill. Empirical questions remain about the way(s) that outcome prediction might affect access to justice. Yet if developments continue, policy makers and practitioners should be ready to exploit the tools’ substantial potential to fill the civil justice gap while also guarding against the harms they might cause.

⁵ Kathryn A. Sabbeth, *Housing Defense as the New Gideon*, 41 HARV. J.L. & GENDER 55, 56–57 (2018); see also Pamela Bookman & Colleen F. Shanahan, *A Tale of Two Civil Procedures*, 122 COLUM. L. REV. (forthcoming 2022) (describing state courts as “lawyerless”).

⁶ *Legal Profession*, NEW GA. ENCYCLOPEDIA (Aug. 11, 2020), <https://www.georgiaencyclopedia.org/articles/government-politics/legal-profession/>; Kathryn Hayes Tucker, *Here Are the Six Georgia Counties That Have No Lawyers*, THE DAILY REPORT (Jan. 8, 2015), <https://www.law.com/dailyreportonline/almlID/1202714378330/Here-Are-the-Six-Georgia-Counties-That-Have-No-Lawyers?/>.

⁷ Tucker, *Six Georgia Counties*.

7.1. LITIGATION OUTCOME PREDICTION DEFINED

I define “computationally driven litigation outcome prediction tools” as statistical or machine learning methods used to forecast the outcome of a civil litigation event, claim, or case. A litigation event may be a motion filed by either party; the relevant predicted outcome would be the judge’s decision to grant or deny, in full or in part. A claim or case outcome, on the other hand, refers to the disposition of a lawsuit, again in full or in part. My scope is civil only, though much of the analysis that follows could apply equally to criminal proceedings.

“Computationally driven” here refers to the use of statistical or machine learning models to detect patterns in past civil litigation data and exploit those patterns to predict, and to some extent explain, future outcomes. Just as actuaries compute the future risk of loss for insurance companies based on past claims data, so do outcome prediction tools attempt to compute the likelihood of future litigation events based on data gleaned from past court records.

In broad strokes, such tools take as their inputs a set of characteristics, also known as predictors, independent variables, or features, that describe the facts, legal claims, arguments, and authority, the people (judge, lawyers, litigants, expert witnesses), and the setting (location, court) of a case. Features might also come from external sources or be “engineered” by combining data. For example, the judge’s gender and years on the bench might be features, as well as the number of times the lawyers in the case had previously appeared before the same judge, the judge’s caseload, and local economic or crime data. Such information might be manually or computationally extracted from the unstructured text of legal documents and other sources – necessitating upstream text mining or natural language processing tasks – or might already be available in structured form.

These various features or case characteristics then become the inputs into one of many types of statistical or predictive models; the particular litigation outcome of interest is the target variable to be predicted.⁸ When using such a tool, a lawyer would plug in the requested case characteristics and would receive an outcome prediction along with some measurement of error.

7.2. THEORY: ACCESS-TO-JUSTICE POTENTIAL

In theory, computationally driven outcome prediction, if good enough, can supplement, stretch, and reduce the cost of legal services by reducing outcome uncertainty. As Gillian Hadfield summarizes, uncertainty comes from several sources.⁹ Sometimes the law is simply unclear. Other times, actors, whether police

⁸ See Chapter 3 in this volume.

⁹ Gillian K. Hadfield, *Weighing the Value of Vagueness: An Economic Perspective on Precision in the Law*, 81 CAL. L. REV. 541 (1994).

officers, prosecutors, regulators, or courts, have deliberately been given discretion. Further, an individual may subjectively discount or increase the probability of liability due to “mistakes in the determination of factual issues, and errors in the identification of the applicable legal rule.”¹⁰ One way to resolve these uncertainties is to pay a lawyer for advice – in particular, a liability estimate.

Given a large enough training set, a predictive model may detect patterns in how courts have previously resolved vagueness and how officials have previously exercised discretion. Further, such a tool could correct the information deficits and asymmetries that may produce mistaken liability estimates. Outcome prediction tools might also obviate the need for legal representation entirely, allowing potential and actual litigants to estimate their own chances of success and proceed *pro se*. This could be a substantial boon for access to justice. Of course, even an outcome-informed *pro se* litigant may fail to navigate complex court procedures and norms successfully.¹¹ Fully opening the courthouse doors to self-represented litigants might also require simplification of court procedures. Still, outcome prediction tools might go a long way toward expanding access to justice, whether by serving litigants directly or by acting as a kind of force multiplier for lawyers and legal organizations, particularly those squaring off against better-resourced adversaries.¹²

A second way outcome prediction tools could, in theory, open up access to justice is by enhancing the ability of legal services providers to quantify, and manage, risk. Profit-driven lawyers, as distinguished from government-funded legal services lawyers, build portfolios of cases with an eye toward managing risk.¹³ Outcome prediction tools may allow lawyers to allocate their resources more efficiently, wasting less money on losing cases and freeing up lawyer time and attention for more meritorious cases, or by constructing portfolios that balance lower- and higher-risk cases.

In addition, enterprising lawyers with a higher-risk appetite might use such tools to discover new areas of practice or potential claim types that folk wisdom would advise against.¹⁴ To draw an example from my previous work, I studied the boom in wage-and-hour lawsuits in the early 2000s and identified as one driver of the

¹⁰ *Id.*

¹¹ Bookman & Shanahan, *A Tale of Two Civil Procedures*, at 16–17; Rebecca L. Sandefur, *Elements of Professional Expertise: Understanding Relational and Substantive Expertise through Lawyers’ Impact*, 80 AM. SOC. REV. 909, 915–16 (2015).

¹² David Freeman Engstrom & Jonah B. Gelbach, *Legal Tech, Civil Procedure, and the Future of Adversarialism*, 169 U. PA. L. REV. 1001, 1072 (2020); John O. McGinnis & Russell G. Pearce, *The Great Disruption: How Machine Intelligence Will Transform the Role of Lawyers in the Delivery of Legal Services*, 82 FORDHAM L. REV. 3041, 3049 (2014).

¹³ Herbert M. Kritzer, *Contingency Fee Lawyers as Gatekeepers in the Civil Justice System*, 81 JUDICATURE 22, 23 (1997) While Kritzer focuses on contingency-free practice, his core insight extends to cases brought under fee-shifting statutes or flat-fee arrangements as well, where lawyers are likewise balancing outlay of resources against probable recovery.

¹⁴ Thanks to David Freeman Engstrom for suggesting this possibility.

litigation spike an influx of enterprising personal injury attorneys into wage-and-hour law.¹⁵ One early mover was a South Florida personal injury attorney named Gregg Shavitz, who discovered his clients' unpaid wage claims by accident, became an overtime specialist, and converted his firm into one of the highest-volume wage-and-hour shops in the country. This was before the wide usage of litigation outcome prediction tools. However, one might imagine that more discoveries like Gregg Shavitz's could be enabled by computationally driven systems, rather than by happenstance, opening up representation for more clients with previously overlooked or under-resourced claim types.¹⁶

I return to, and complicate, this possibility in Section 7.5, where I raise concerns about outcome prediction tools' conservatism in defining winning and losing cases, which may reduce, rather than increase, access to justice – empirical questions that remain to be resolved.

7.3 PRACTICE: WHERE ARE WE NOW?

From theory, I now turn to practice, tracing the evolution and present state of litigation outcome prediction in scholarship, commercial applications, and tools developed specifically to serve low- and middle-income litigants. This Section also begins to introduce these tools' limitations in their present form, a topic that I explore more fully in Section 7.4.

7.3.1 Scholarship

Litigation outcome prediction is an active scholarly research area, characterized by experimentation with an array of different data sets, modeling approaches, and performance measures. Thus far, no single dominant approach has emerged.

In a useful article, Kevin Ashley traces the history of the field to the work of two academics who used a machine learning algorithm called k-nearest neighbors in the 1970s to forecast the outcome of Canadian real estate tax disputes.¹⁷ Since then, academic work has flourished. In the United States, academic interest has focused,

¹⁵ Charlotte S. Alexander, *Litigation Migrants*, 56 AM. BUS. L.J. 235 (2019).

¹⁶ To carry the thought exercise further, perhaps third-party litigation financing firms could fund these sorts of risky case-selection strategies, which solo lawyers or small firms might otherwise be hesitant to adopt. Center on the Legal Profession, Harvard Law School, *The Practice, Investing in Legal Futures*, *Litig. Fin.*, Sept.–Oct. 2019.

¹⁷ Kevin Ashley, *A Brief History of the Changing Roles of Case Prediction in AI and Law*, 36 LAW IN CONTEXT: A SOCIO-LEGAL JOURNAL 93, 96 (2019) (citing Ejan Mackaay & Pierre Robillard, *Predicting Judicial Decisions: The Nearest Neighbor Rule and Visual Representation of Case Patterns*, 3 DATENVERARBEITUNG IM RECHT 302 (1974)).

variously, on decisions by the US Supreme Court,¹⁸ federal appellate courts,¹⁹ federal district courts,²⁰ immigration court,²¹ state trial courts,²² and administrative agencies.²³ Case types studied include employment,²⁴ asylum,²⁵ tort and vehicular,²⁶ and trade secret misappropriation.²⁷ Other scholars outside the United States have, in turn, developed outcome prediction tools focused on the European Court of Human Rights,²⁸ the International Criminal Court,²⁹ French appeals courts,³⁰ the Supreme Court of the Philippines,³¹ lending cases in China,³² labor cases in Brazil,³³ public morality and freedom of expression cases in Turkey's Constitutional Court,³⁴ and Canadian employment and tax cases.³⁵ Some of this research has spun off into commercial products, discussed in the next section.

¹⁸ See, e.g., Daniel Martin Katz, Michael J. Bommarito II & Josh Blackman, *A General Approach for Predicting the Behavior of the Supreme Court of the United States*, 12 PLoS ONE (2017).

¹⁹ See, e.g., Sergio Galletta, Elliott Ash & Daniel L. Chen, *Measuring Judicial Sentiment: Methods and Application to U.S. Circuit Courts* (Aug. 19, 2021) (unpublished manuscript), <https://ssrn.com/abstract=3415393>.

²⁰ Elizabeth C. Tippet et al., *Does Lawyering Matter? Predicting Judicial Decisions from Legal Briefs, and What That Means for Access to Justice*, 101 TEXAS L. REV. (forthcoming 2022).

²¹ See, e.g., Matthew Dunn et al., *Early Predictability of Asylum Court Decisions*, 2017 PROC. ACM CONF. ON AI & LAW.

²² See, e.g., Devin J. McConnell et al., *Case-Level Prediction of Motion Outcomes in Civil Litigation*, 18 PROC. INT'L CONF. ON A.I. & LAW 99 (2021).

²³ See, e.g., Karl Branting et al., *Semi-Supervised Methods for Explainable Legal Prediction*, 17 PROC. INT'L CONF. ON A.I. & LAW 22 (2019).

²⁴ Tippet et al., *Does Lawyering Matter?*.

²⁵ Dunn et al., *Early Predictability of Asylum Court Decisions*.

²⁶ McConnell et al., *Case-Level Prediction of Motion Outcomes*.

²⁷ Kevin D. Ashley & Stefanie Brünighaus, *Automatically Classifying Case Texts and Predicting Outcomes*, 17 A.I. L. 125 (2009).

²⁸ See, e.g., Nikolaos Aletas et al., *Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective*, 2(93) PEERJ COMPUT. SCI. (2016).

²⁹ Fabien Tarissan & Raphaëlle Nollez-Goldbach, *Analysing the First Case of the International Criminal Court from a Network-Science Perspective*, 4 J. COMPLEX NETWORKS 616 (2016).

³⁰ Paul Boniol et al., *Performance in the Courtroom: Automated Processing and Visualization of Appeal Court Decisions in France*, ARXIV (2020), <https://arxiv.org/abs/2006.06251>.

³¹ Michael Benedict L. Virtucio et al., *Predicting Decisions of the Philippine Supreme Court Using Natural Language Processing and Machine Learning*, 42 IEEE INT'L CONF. ON COMPUT. SOFTWARE & APPLICATIONS 130 (2018).

³² Luyao Ma et al., *Legal Judgment Prediction with Multi-Stage Case Representation Learning in the Real Court Setting*, 44 PROC. INT'L ACM SIGIR CONF. ON RSCH. & DEV. IN INFO. RETRIEVAL 993 (2021).

³³ Andre Lage-Freitas et al., *Predicting Brazilian Court Decisions*, ARXIV (2019), <https://arxiv.org/abs/1905.10348>.

³⁴ Mehmet Fatih Sert, Engin Yıldırım & İrfan Haslak, *Using Artificial Intelligence to Predict Decisions of the Turkish Constitutional Court*, 2021 SOC. SCI. COMP. REV. 1.

³⁵ Maxime C. Cohen et al., *The Use of AI in Legal Systems: Determining Independent Contractor vs. Employee Status* (Jan. 28, 2022) (unpublished manuscript) <https://ssrn.com/abstract=4013823>; Yifei Yin, Farhana Zulkemine & Samuel Dahan, *Determining Worker Type from Legal Text Data Using Machine Learning*, 2020 IEEE INT'L CONF. ON DEPENDABLE, AUTONOMIC & SECURE COMPUTING; Benjamin Alarie et al., *Using Machine Learning to Predict Outcomes in Tax Law*, 58 CAN. BUS. L.J. 231 (2016).

This scholarly work reflects all the strengths and weaknesses of the wider field. Though direct comparison among studies can be difficult given different datasets and performance measures, predictive performance has ranged from relatively modest marginal classification accuracy³⁶ to a very high F1 score of 98 percent in one study.³⁷

That said, some high-performing academic approaches may suffer from research design flaws, as they appear to use the text of a court's description of the facts of a case and the laws cited to predict the court's ruling.³⁸ This is problematic, as judges or their clerks often write case descriptions and choose legal citations with pre-existing knowledge of the ruling they will issue. It is no surprise that these case descriptions predict outcomes. Further, much academic work is limited in its generalizability by the narrow band of cases used to train and test predictive models. This is due to inaccessible or missing court data, especially in the United States, a problem discussed further in Section 7.4. Finally, some researchers give short shrift to explanation, in favor of prediction.³⁹ Though a model may perform well in forecasting results, its practical and tactical utility may be limited if lawyers seeking to make representation decisions do not know what drives the predictions and cannot square them with their mental models of the world. As discussed further in Section 7.4, explainable predictions are becoming the new norm, as interpretations are now available for even the most "black box" predictive models. For the moment, however, explainability remains a sticking point.

7.3.2 Commercial Applications

The commercial lay of the land is similar to the academic landscape, with substantial activity and disparate approaches focused on particular case types or litigation events.

³⁶ See, e.g., McConnell et al., *Case-Level Prediction of Motion Outcomes*, at 104 (reporting "maximum classification accuracy of 0.644" as compared to a naïve baseline of 0.501 using adaBoost, a decision-tree-based classification method and a variety of preprocessing steps applied to the input text).

³⁷ See, e.g., Octavia-Maria Şulea et al., *Exploring the Use of Text Classification in the Legal Domain*, ARXIV (2017), <https://arxiv.org/abs/1710.09306> (reporting "results of 98% average F1 score in predicting a case ruling" of the French Supreme Court).

³⁸ See *id.*

³⁹ Compare, e.g., Katz et al., *A General Approach for Predicting the Behavior of the Supreme Court* (focusing exclusively on predictive performance), with Ma et al., *Legal Judgment Prediction*, at 8–9 (presenting interpretability strategy for "black box" neural network predictions). This may be an unfair critique, as prediction and explanation can be two entirely separate goals. A classic example illustrates the difference: A ruler who wants to know whether to spend money for a rain dance to break a drought cares about causation. The same ruler who wants to know whether it will rain tomorrow so she can bring an umbrella cares only about prediction. Will it rain or not? For more discussion, see Jon Kleinberg et al., *Prediction Policy Problems*, 105 AM. ECON. REV. 491 (2015).

The Big Three legal research companies – LexisNexis, Westlaw, and Bloomberg Law – have all developed outcome prediction tools that sit within their existing suites of research and analysis tools. LexisNexis offers what they label “judge and court analytics” as well as “attorney and law firm analytics.” In both spaces, the offerings are more descriptive than predictive – showing, for example, “a tally of total cases for a judge or court for a specific area of law to approximate experience on a motion like yours.”⁴⁰ The predictive jump is left to the user, who decides whether to adopt the approximation as a prediction or to distinguish it from the case at hand. LexisNexis provides further predictive firepower in the form of an acquired start-up, LexMachina, which provides, among other output, estimates of judges’ likelihood of granting or denying certain motions in certain case types.⁴¹ Westlaw offers similar options in its litigation and precedent analytics tools,⁴² as does Bloomberg Law in its litigation analytics suite.⁴³ Fastcase, a newer entrant into the space, offers a different approach, allowing subscribers to build their own bespoke predictive and descriptive analyses, using tools and methodologies drawn from a host of partner companies.⁴⁴

A collection of smaller companies offers litigation outcome prediction focused on particular practice areas or litigation events. Docket Alarm, now owned by Fastcase, offers patent litigation analytics that produce “the likelihood of winning given a particular judge, technology area, law firm or party.”⁴⁵ In Canada, Blue J Tax builds on the scholarly work described above to offer outcome prediction in tax disputes,⁴⁶ while in the United Kingdom companies like CourtQuant “predict [case] outcome and settlement probability.”⁴⁷

A final segment of the industry are law firms’ and other players’⁴⁸ homegrown, proprietary tools. On the plaintiffs’ side, giant personal injury firm Morgan & Morgan has developed “a ‘Google-style’ operation” in which the firm “evaluate[s] ‘actionable data points’ about personal injury settlements or court proceedings” and uses the insight to “work up a case accordingly – and . . . do that at scale.”⁴⁹ Defense-

⁴⁰ *Litigation Analytics*, LEXISNEXIS, <https://www.lexisnexis.com/en-us/products/lexis-plus/litigation-analytics.page>.

⁴¹ LEX MACHINA, <https://lexmachina.com/legal-analytics/>.

⁴² *Litigation Analytics*, WESTLAW EDGE <https://legal.thomsonreuters.com/en/products/westlaw-edge/litigation-analytics#compare>; *Precedent Analytics*, WESTLAW EDGE, <https://legal.thomsonreuters.com/en/insights/articles/introducing-precedent-analytics>.

⁴³ *Legal Analytics*, BLOOMBERG L., <https://pro.bloomberglaw.com/legal-analytics/>.

⁴⁴ *AI Sandbox*, FASTCASE, <https://www.fastcase.com/sandbox/>.

⁴⁵ DOCKET ALARM, <https://www.docketalarm.com/>.

⁴⁶ BLUE J TAX, <https://www.bluej.com/blue-j-tax>.

⁴⁷ COURTQUANT, <https://www.courtquant.com/about>.

⁴⁸ Here, legal operations service providers like Ernst and Young and other accounting and consulting firms, third-party litigation finance companies, and insurance companies that insure against litigation costs are all players that are invested in tech- and often AI-fueled outcome prediction. See, e.g., APEX LITIG. FIN., <https://www.apexlitigation.com/>.

⁴⁹ Christine Schiffner, *Inside the “Google-Style” Tech Hub Driving Plaintiffs Firms’ Growth*, NAT’L L.J. (Nov. 15, 2021), <https://www.law.com/nationallawjournal/2021/11/15/inside-the-google-style-tech-hub-driving-plaintiffs-firms-growth/>.

side firms are doing the same. Dentons, the world's largest firm, even spun off an independent analytics lab and venture firm to fund development in outcome prediction and other AI-enabled approaches to law.⁵⁰

It is difficult to assess how well any of these tools performs, as access is expensive or unavailable, the feature sets used as inputs are not always clear, and the algorithms that power the predictions are hidden. I raise some concerns about commercial model design in Section 7.4 – in particular, reliance on lawyer identity as a predictor – and, as above, return to the perpetual problem of inaccessible and missing court data.

7.3.3 Outcome Prediction for Low- and Middle-Income Litigants

For reasons explored further below, examples are scarce of computationally driven litigation outcome prediction tools engineered specifically for the kinds of cases noted in this chapter's opening. Philadelphia's civil legal services provider, Community Legal Services, uses a tool called Expungement Generator (EG) to determine whether criminal record expungement is possible and assist in completing the paperwork.⁵¹ The EG does not predict outcomes, but its automated approach enables efficiency gains for an organization that prepares thousands of expungement petitions per year.⁵² Similarly, an application developed in the Family Law Clinic at Duquesne University School of Law prompts litigants in child support cases to answer a set of questions, which the tool then evaluates to determine "if there is a meritorious claim for appeal to be raised" under Pennsylvania law.⁵³ As with the EG, the Duquesne system does not appear to use machine learning techniques, but rather to apply a set of mechanical rules. The clinic plans prediction as a next step, however, and is developing a tool that analyzes winning arguments in appellate cases in order to guide users' own arguments.⁵⁴

7.4. PRESENT LIMITS

Having surveyed the state of the outcome prediction field, I now step back and assess its limits. As David Freeman Engstrom and Jonah Gelbach rightly concluded in earlier work: "[L]egal tech tools will arrive sooner, and advance most rapidly, in legal

⁵⁰ *Dentons Launches Nextlaw Labs and Creates Legal Business Accelerator*, DENTONS (May 19, 2015) <https://www.dentons.com/en/about-dentons/news-events-and-awards/news/2015/may/dentons-launches-nextlaw-labs-creates-legal-business-accelerator>.

⁵¹ NateV, *Expungement-Generator*, GITHUB, <https://github.com/NateV/Expungement-Generator>.

⁵² *Id.*; Rana Faye, *Meet the Disruptor: Michael Holland*, PHILA. CITIZEN (May 3, 2016), <https://thephiladelphiacitizen.org/disruptor-michael-hollander-expungement-generator/>.

⁵³ Katherine L. Norton, *Mind the Gap: Technology as a Lifeline for Pro Se Child Custody Appeals*, 58 DUQ. L. REV. 82, 91 (2020).

⁵⁴ *Id.*

areas where data is abundant, regulated conduct takes repetitive and stereotypical forms, legal rules are inherently stable, and case volumes are such that a repeat player stands to gain financially by investing.”⁵⁵ Many of the commercial tools highlighted above fit this profile. Tax-oriented products exploit relatively stable rules; Morgan & Morgan’s internal case evaluation system exploits the firm’s extraordinarily high case volumes.

Yet, as noted above, data’s “abundance” is an open question, as is data quality. Methodological problems may also hinder these tools’ development. In the access to justice domain, the questions of investment incentives and financial gains loom large as well. The remainder of this Section addresses these limitations.

7.4.1 *Data Limitations*

Predictive algorithms require access to large amounts of data from previous court cases for model training, but such bulk data is not widely or freely available in the United States from the state or federal courts or from administrative agencies that have an adjudicatory function.⁵⁶ The Big Three have invested substantial funds in compiling private troves of court documents and judicial decisions, and jealously guard those resources with high user fees, restrictive terms and conditions, and threatened and actual litigation.⁵⁷

Data inaccessibility creates serious problems for outcome prediction tools designed to meet the legal needs of low- and middle-income litigants.⁵⁸ Much of this litigation occurs in state courts, where data is sometimes poorly managed and siloed in multiple systems.⁵⁹ Moreover, there is little money in practice areas like eviction defense and public benefits appeals, in which clients, by definition, are poor. Thus, data costs are high, and financial incentives for investment in research and development are low.

Even the products offered by the monied Big Three, however, suffer from data problems. With large companies separately assembling their own private data repositories, coverage varies widely, producing remarkable disagreement about basic facts. A recent study revealed that the answers supplied to the question “How many opinions on motions for summary judgment has Judge Barbara Lynn (N.D. Tex.) issued in patent cases?” ranged from nine to thirty-two, depending on the legal

⁵⁵ Engstrom & Gelbach, *Legal Tech*, at 1029.

⁵⁶ 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 (Ryan Whalen ed., 2020).

⁵⁷ Alaina Lancaster, *Judge Rejects ROSS Intelligence’s Dismissal Attempt of Thomson Reuters Suit over Westlaw Content*, LAW.COM (Mar. 29, 2021), <https://www.law.com/therecorder/2021/03/29/judge-rejects-ross-intelligences-dismissal-attempt-of-thomson-reuters-suit-over-westlaw-content/>.

⁵⁸ See Chapters 6, 14, and 15 in this volume.

⁵⁹ See Chapter 13 in this volume.

research product used.⁶⁰ This is an existential problem for the future of litigation outcome prediction, as predictions are only as good as the data on which they are built.⁶¹

A final data limitation centers on the challenges of causal explanation. Even if explainable modeling approaches are used, the case characteristics that appear to be the strongest predictors of outcomes may not, in fact, be actionable. For instance, when a predictive tool relies on attorney identity as a feature, the model's prediction may actually be free-riding on the attorney's own screening and selection decisions. In other words, if the presence of Lawyer A in a case is strongly predictive of a win for her clients, Lawyer A's skills as a litigator may not be the true cause. The omitted, *more* predictive variable is likely the strength of the merits, and Lawyer A's skill at assessing those merits up-front. Better data could enable better model construction, avoiding these kinds of proxy variable traps.

7.4.2 Methodological Limitations

Sitting atop these data limitations are two important methodological limitations. First, as noted above, even if predictive tools do a good job of forecasting the probable outcome of a litigation event, they may only poorly explain *why* the predicted outcome is likely to occur. Explanation is important for a number of related reasons, among them engendering confidence in predictions, enabling bias and error detection, and respecting the dignity of people affected by prediction.⁶² Indeed, the European Union's General Data Protection Regulation (GDPR) has established what some scholars have labeled a "right to an explanation," consisting of a right "not to be subject to a decision based solely on automated processing" and various rights to notice of data collection.⁶³ Though researchers are actively developing explainable AI that can identify features' specific importance to a prediction and generate counterfactual predictions if features change value,⁶⁴ the field has yet to converge on a single set of explainability practices, and commercial approaches vary widely.

⁶⁰ Sean La Roque Doherty, *Not All Litigation Analytics Products Are Created Equal*, A.B.A. J. (Aug. 1, 2020), <https://www.abajournal.com/magazine/article/analytics-products-offer-different-results-depending-on-data-sources-quality-and-the-types-of-analytics-and-reports-they-provide>.

⁶¹ Engstrom and Engstrom's contribution to this volume identifies yet another data limitation: the absence of reliable data on cases and claims that are settled, where the contents of the settlement are unavailable. See Chapter 6 in this volume.

⁶² Margot E. Kaminski, *The Right to Explanation, Explained*, 34 BERKELEY TECH. L.J. 189 (2019).

⁶³ *Id.*

⁶⁴ See, e.g., Scott M. Lundberg et al., *From Local Explanations to Global Understanding with Explainable AI for Trees*, 2 NATURE MACH. INTELLIGENCE 56 (2020); *What Is Explainability?* ALIBI, https://docs.seldon.io/projects/alibi/en/stable/overview/high_level.html#what-is-explainability.

Second, outcome prediction is limited by machine and deep learning algorithms' inability to reason by analogy. Legal reasoning depends on analogical thinking: the ability to align one set of facts to another and guess at the likely application of the law, given the factual divergences. However, teaching AI to reason by analogy is a cutting-edge area of computer science research, and it is far from well established. As computer scientist Melanie Mitchell explains, "'Today's state-of-the-art neural networks are very good at certain tasks . . . but they're very bad at taking what they've learned in one kind of situation and transferring it to another' – the essence of analogy."⁶⁵ There is a famous analogical example in text analytics, where a natural language processing technique known as word embedding, when trained on an enormous corpus of real-world text, is able to produce the answer "queen" when presented with the formula "king minus man plus woman."⁶⁶ The jump from this parlor trick to full-blown legal reasoning, though, is substantial.

In short, scaling up computationally driven litigation outcome prediction tools in a way that would fill the civil justice gap would require access to more and better data and methodological advances. Making bulk federal and state court and administrative agency data and records freely and easily accessible would be a very good step.⁶⁷ Marshaling resources to support methods and tool development would be another. Foundation funding is already a common ingredient in efforts to fill the civil justice gap. I propose that large law firms pitch in as well. All firms on the AmLaw 100 could pledge a portion of their pro bono budgets toward the development of litigation outcome prediction tools to be used in pro bono and low bono settings. The ABA Foundation might play a coordinating and convening role, as it is already committed to access-to-justice initiatives. Such an effort could have a much broader impact than firms' existing pro bono activities, which tend to focus on representation in single cases. It might also jump-start additional interest from the Big Three and other commercial competitors, who might invest more money in improving algorithms' predictive performance and spin off free or low-cost versions of their existing suites of tools.

7.5 UNINTENDED CONSEQUENCES

Time will tell whether, when, and how the data and methodological problems identified in Section 7.4 will be solved. Assuming that they are, and litigation

⁶⁵ John Pavlus, *The Computer Scientist Training AI to Think with Analogies*, QUANTA MAG. (July 14, 2021), <https://www.quantamagazine.org/melanie-mitchell-trains-ai-to-think-with-analogies-20210714/>; see also Katie Atkinson & Trevor Bench-Capon, *Reasoning with Legal Cases: Analogy or Rule Application?* 17 PROC. INT'L CONF. ON ARTIFICIAL INTELLIGENCE & L. 12 (June 2019).

⁶⁶ Emerging Technology from the arXiv, *King – Man + Woman = Queen: The Marvelous Mathematics of Computational Linguistics*, MIT TECH. REVIEW (Sept. 17, 2015), <https://www.technologyreview.com/2015/09/17/h66211/king-man-woman-queen-the-marvelous-mathematics-of-computational-linguistics/>.

⁶⁷ Adam R. Pah et al., *How to Build a More Open Justice System*, 369(6500) SCIENCE (2020).

outcome prediction tools can reliably generate highly reliable forecasts, there still may be reason for caution.

This Section identifies two possible unintended consequences of outcome prediction tools, which could develop alongside the salutatory access to justice effects described in Section 7.2: harm to would-be litigants denied representation whose claims are novel or less viable according to predictive tools, and harm to the common law system as a whole.

Here, the assumption is that such tools have access to ample data, account for all relevant variables, and are transparent and explainable – in other words, the tools work as intended to learn from existing patterns in civil litigation outcomes and reproduce those patterns as outcome predictions. Yet it is this very reproductive nature that is cause for concern.

7.5.1 Harms to Would-Be Litigants

Consider the facts of *Elisa B. v. Superior Court*,⁶⁸ a case decided by the California Supreme Court in 2005. Emily B. sought child support from her estranged partner, Elisa B., for twins whom Emily had conceived via artificial insemination of her eggs during her relationship with Elisa. If Emily walked into a lawyer's office seeking help with her child support action, the lawyer might be interested in the case's viability: How often have similar fact patterns come before California courts, and what was their outcome? The answers might inform the lawyer's decision about whether to offer representation.

In real life, this case was one of first impression in California. The governing law, the Uniform Parentage Act, referred to "mother" and "father" as the potential parents.⁶⁹ Searching for relevant precedent, the *Elisa B.* court reasoned by analogy from previous cases that involved, variously, three potential parents (one man and two women), non-biological fathers, non-biological mothers, and a woman who raised her half-brother as her son.⁷⁰ From this and other precedent, the court cobbled together a new legal rule that required Elisa B. to pay child support for her and Emily B.'s children.

I am doubtful that an outcome prediction tool would have reached this same conclusion. The number of analogical jumps that the court made would seem to be outside the capabilities of machine and deep learning, even assuming methodological advancement.⁷¹ Further, judges' decisions about what prior caselaw to draw upon and how many analogical leaps to make may be influenced by factors like

⁶⁸ *Elisa B. v. Sup. Ct.*, 117 P.3d 660 (Cal. 2005).

⁶⁹ *Id.* at 664 ("The UPA defines the '[p]arent and child relationship' as 'the legal relationship existing between a child and the child's natural or adoptive parents' The term includes the mother and child relationship and the father and child relationship.").

⁷⁰ *Id.* at 667.

⁷¹ Atkinson & Bench-Capon, *Reasoning with Legal Cases*.

ideology and public opinion, which could be difficult to model well. Emily B.'s claim would likely receive a very low viability score.⁷²

A similar cautionary tale comes from my own previous work with Camille Gear Rich and Zev Eigen on attorneys' non-computational assessments of claim viability. We documented plaintiffs' employment attorneys' dim view of the likelihood of success for employment discrimination claims and their shifting of case selection decisions away from discrimination and toward easier-to-prove wage-and-hour claims.⁷³ One result of this shift, we observed, was that even litigants with meritorious discrimination claims were unable to find legal representation. That work happened in 2014 and 2015, before litigation outcome prediction tools were widely available, and I am not aware of subsequent empirical studies on the effect of such tools on lawyers' intake decisions. Yet if lawyers were already using their intuition to learn from past cases and predict future outcomes, pre-AI, machine and deep learning tools could just cement these same patterns in place.

Thus, in this view, as civil litigation outcomes become more predictable, claims become commoditized. Outlier claims and clients like Emily B. may become less representable, much like high-loss risks become less insurable. While access to justice on the whole may increase, the courthouse doors may be effectively closed to some classes of potential clients who seek representation for novel or disfavored legal claims or defenses.⁷⁴

Further, to the extent that representation is denied to would-be litigants because of their own negative personal histories, ingested by a model as data points, litigation outcome prediction tools can reduce people to their worst past acts and prevent them from changing course. Take as an example a tenant with an old criminal record trying to fight an eviction, whose past conviction reduces her chance of winning according to an algorithmic viability assessment. This may be factually accurate – her criminal record may *actually* make eviction defense more challenging – but a creative lawyer might see other aspects of her case that an algorithmic assessment might miss. By reducing people to feature sets and exploiting the features that are most predictive of outcomes, but perhaps least representative of people's full selves, computational tools enact dignitary harm. In the context of low-income

⁷² Prediction tools become like Oliver Wendell Holmes' Vermont justice: "There is a story of a Vermont justice of the peace before whom a suit was brought by one farmer against another for breaking a chum. The justice took time to consider, and then said that he has looked through the statutes and could find nothing about chums, and gave judgment for the defendant." Oliver Wendell Holmes, Jr., *The Path of the Law*, 10 HARV. L. REV. 457 (1897).

⁷³ Charlotte S. Alexander, Zev Eigen & Camille Gear Rich, *Post-Racial Hydraulics: The Hidden Dangers of the Universal Turn*, 91 N.Y.U. L. REV. 1 (2016).

⁷⁴ A future of "legal singularity," in which all outcomes are perfectly predictable, is not necessary for my argument here. Benjamin Alarie, *The Path of the Law: Towards Legal Singularity*, 66 U. TORONTO L.J. 443 (2016). Even prediction that works well for some subclass of cases will change lawyers' preferences for those cases over other, less certain cases. This has consequences for those clients' civil legal needs.

litigants facing serious and potentially destabilizing court proceedings, and who are algorithmically denied legal representation, such tools can also cause substantial economic and social harm, reducing social mobility and locking people into place.

Indeed, machine and deep learning methods are inherently prone to what some researchers have called “value lock-in.”⁷⁵ All data is historical in the sense that it captures points in time that have passed; all machine and deep learning algorithms find patterns in historical data as a way to predict the future. This methodological design reifies past practices and locks in past patterns. As machine learning researcher Abeba Birhane and her collaborators point out, then, machine learning is not “value-neutral.”⁷⁶ And as AI pioneer Joseph Weizenbaum observed, “the computer has from the beginning been a fundamentally conservative force which solidified existing power: in place of fundamental social changes . . . the computer renders technical solutions that allow existing power hierarchies to remain intact.”⁷⁷ It is no accident that the anecdotes above involve a lesbian couple, employment discrimination claimants, and a tenant with a criminal record: the fear is that would-be litigants like these with the least power historically become further disempowered at the hands of computational methods.

Yet as Section 7.2 suggested, a different story might also be possible: More accurate predictions might enable lawyers to fill their case portfolios with low-risk sure winners as hedges when taking on riskier cases like *Elisa B.*, or might help them discover and invest in previously under-resourced practice areas. At this stage, whether predictive tools would increase or decrease representation for outlier claims and clients is an open empirical question, which researchers and policy makers should work to answer as data and methods improve and outcome prediction tools become more widely used.

7.5.2 Harms to the System

I turn now to the second potential harm caused by computationally driven litigation outcome prediction: harm to the common law system itself.⁷⁸ As Charles Barzun explains, common-law reasoning “contains seeds of radicalism [in that] the case-by-case process by which the law develops means it is always open to revision. And even though its official position is one of incremental change . . . doctrine [is] constantly

⁷⁵ See, e.g., Laura Weidinger et al., *Ethical and Social Risks of Harm from Language Models*, ARXIV (2021), <https://arxiv.org/abs/2112.04359>.

⁷⁶ Abeba Birhane et al., *The Values Encoded in Machine Learning Research*, ARXIV (2021), <https://arxiv.org/abs/2106.15590>.

⁷⁷ *Id.* (citing JOSEPH WEIZENBAUM, *COMPUTER POWER AND HUMAN REASON: FROM JUDGMENT TO CALCULATION* (1976)).

⁷⁸ David Freeman Engstrom, *Private Litigation’s Pathways: Lessons from Qui Tam Enforcement*, 114 COLUM. L. REV. 1913, 1934 (2014) (“[P]rivate enforcers will tend to push into statutory and regulatory interstices.”).

vulnerable to being upended.”⁷⁹ Barzun points to Catharine MacKinnon’s invention of sexual harassment doctrine out of Title VII’s cloth as an example of a “two-way process of interaction” between litigants, representing their real-world experience, and the courts, interpreting the law, in a shared creative process “in which the meaning and scope of application of the statute changes over time.”⁸⁰

If lawyers rely too heavily on litigation outcome prediction tools, which reproduce past patterns, the stream of new fact presentations and legal arguments flowing into the courts dries up. Litigation outcome prediction tools may produce a sort of super stare decisis by narrowing lawyers’ case selection preferences to only those case, claim, and client types that have previously appeared and been successful in court. Yet stare decisis is only one aspect of our common law system. Another competing characteristic is flexibility: A regular influx of new cases with new fact patterns and legal arguments enables the law to innovate and adapt. In other words, noise – as differentiated from signal – is a feature of the common law, not a bug. Outcome prediction tools that are too good at picking up signals and ignoring noise eliminate the structural benefits of the noise, and privilege stare decisis over flexibility by shaping the flow of cases that make their way to court.

Others, particularly Engstrom and Gelbach, have made this point, suggesting that prediction

comes at a steep cost, draining the law of its capacity to adapt to new developments or to ventilate legal rules in formal, public interpretive exercises The system also loses its legitimacy as a way to manage social conflict when the process of enforcing collective value judgments plays out in server farms rather than a messy deliberative and adjudicatory process, even where machine predictions prove perfectly accurate.⁸¹

The danger is that law becomes endogenous and ossified. “Endogenous,” to repurpose a concept introduced by Lauren Edelman, means that the law’s inputs become the same as its outputs and “the content and meaning of law is determined within the social field that it is designed to regulate.”⁸² “Ossified,” to borrow from Cynthia Estlund, means that the law becomes “essentially sealed off . . . both from democratic revision and renewal from local experimentation and innovation.”⁸³

⁷⁹ Charles L. Barzun, *The Common Law and Critical Theory*, 92 COLO. L. REV. 1, 13 (2021).

⁸⁰ *Id.* at 8.

⁸¹ Engstrom & Gelbach, *Legal Tech*, at 1036–37.

⁸² Lauren B. Edelman, Christopher Uggen & Howard S. Erlanger, *The Endogeneity of Legal Regulation: Grievance Procedures as Rational Myth*, 105 AM. J. SOCIO. 406 (1999).

⁸³ Cynthia L. Estlund, *The Ossification of American Labor Law*, 102 COLUM. L. REV. 1527, 1530 (2002). The same points have been made in connection with grant funding for scientific research, where the fear is that innovation is stifled because researchers hew too closely to the example of previous successfully funded proposals. See, e.g., Scott O. Lilienfeld, *Psychology’s*

7.6 NEXT STEPS

As noted above, whether any of the unintended consequences outlined above will come to pass – and, indeed, whether access to justice improvements will come to pass as well – turns on empirical questions. Given the problems and limitations identified in Section 7.4, will litigation outcome prediction tools actually work well enough to achieve either their potential benefits or cause their potential harms? My assessment of the present state of the field suggests there is a long way to go before we reach either set of outcomes. But as the field matures, we can build in safeguards against the endogeneity risks and harms I identify above through technical, organizational, and policy interventions.

First, on the technical side, computer and data scientists, and the funders who make their work possible, should invest heavily in improving algorithmic analogical reasoning. Without the ability to reason by analogy, outcome predictors not only will miss an array of possible positive predictions, but they will also be systematically biased against fact patterns like Emily B.’s, which present issues of first impression.

Further on the technical front, developers could purposefully over-train predictive algorithms on novel, but successful, fact patterns and legal arguments in order to nudge the system off its path and make positive predictions possible even for cases that fall outside the norm. This idea is adapted from OpenAI’s work in nudging its state-of-the-art language model, GPT-3, away from its “harmful biases, such as outputting discriminatory racial text” learned from its training corpus, by over-exposing it to counter texts.⁸⁴

Technical fixes focus on outcome prediction tools’ production side. Organizational fixes target the tools’ consumers: the lawyers, law firms, and other legal organizations that might use them to influence case selection. I propose here that no decision should be made exclusively on the basis of algorithmic output. This guards against the dignitary and other real harms described above, as would-be litigants are treated as full people rather than feature sets. This also parallels the GDPR’s explanation mandate, though I suggest it here as an organizational practice that is baked into legal organizations’ decision-making processes.⁸⁵

Finally, I turn to policy. The story above assumes a profit-driven lawyer as the user of outcome prediction tools. Of course, there are other possible motivations for a lawyer’s case selection decisions, such as seeking affirmatively to establish a new interpretation of the law or right a historic wrong. These cause lawyers, from all points on the ideological spectrum, may be particularly likely to take on seemingly

Replication Crisis and the Grant Culture: Righting the Ship, 12 PERSPECTIVES ON PSYCH. SCI. 660 (2017).

⁸⁴ Irene Solaiman & Christy Dennison, *Process for Adapting Language Models to Society (PALMS) with Values-Targeted Datasets*, ARXIV (2021), <https://arxiv.org/abs/2106.10328>.

⁸⁵ Kaminski, *The Right to Explanation, Explained*; see also CATHY O’NEILL, *WEAPONS OF MATH DESTRUCTION* 205 (2016) (proposing a Hippocratic Oath for data scientists).

high-risk claim or party types, which receive low computationally determined viability scores. Government lawyers, too, may function as cause lawyers, pushing legal arguments, in accordance with administration position, that diverge from courts' past practices. Government agencies should study trends in private attorneys' use of litigation outcome prediction tools in the areas in which they regulate, and should make their own case selection decisions to fill gaps in representation.⁸⁶

7.7 CONCLUSION

This chapter has explored the consequences of computationally driven litigation outcome prediction tools for the civil justice system, with a focus on increasing access to justice. It has mapped the current state of the outcome prediction field in academic work and commercial applications, as well as in pro bono and low bono practice settings. It has also raised concerns about unintended consequences for litigants and for our legal system as a whole.

I conclude that there is plenty of reason for “techno-optimism,” to use Tanina Rostain’s term, about the potential for computationally driven litigation outcome prediction tools to close the civil justice gap.⁸⁷ However, reaching that optimistic future, while also guarding against potential harms, will require substantially more money and data, continued methodological improvement, careful organizational implementation, and strategic deployment of government resources.

⁸⁶ For an analogous use of government resources to fill private enforcement gaps, see David Weil, *Improving Workplace Conditions through Strategic Enforcement*, RUSSELL SAGE FOUND. (2010), <https://www.russellsage.org/research/report/strategic-enforcement>.

⁸⁷ Tanina Rostain, *Techno-Optimism and Access to the Legal System*, 148 DAEDALUS 93 (2019).