QUANTITATIVE LEGAL PREDICTION—OR—HOW I LEARNED TO STOP WORRYING AND START PREPARING FOR THE DATA-DRIVEN FUTURE OF THE LEGAL SERVICES INDUSTRY

Daniel Martin Katz[*]

INTRODUCTION

Welcome to law's information revolution[1]—revolution already in progress.[2] While the 2008 financial crisis can be seen as the precipitating event, developments in legal information technology are actually a root cause of many of the long-term changes in the legal services market. When considering the downturn in the legal employment market, one should understand there are two distinct trends at play—one is cyclical and the other is structural.[3] The cyclical downturn in the market for legal services is related to broader economic conditions.[4] Some portion of the downturn in demand specifically associated with the broader business cycle will likely abate once broader economic conditions improve. Driven by technology, the structural portion of the downturn appears to be permanent, such that many of those legal jobs displaced both before and by the great recession will not return.[5]

[*] Assistant Professor of Law, Michigan State University. Ph.D. University of Michigan (2011), M.P.P. University of Michigan (2005), J.D. University of Michigan (2005). I would like to thank everyone who has helped in the development of this paper, but I would particularly like to thank the late Larry Ribstein for the extensive thoughts he offered on this and other related projects.

[1] Bruce H. Kobayashi & Larry E. Ribstein, *Law's Information Revolution*, 53 ARIZ. L. REV. 1169 (2011); *see also* Larry E. Ribstein, *The Death of Big Law*, 2010 WIS. L. REV. 749.

[2] *See, e.g.*, Jeff Gray, *Welcome to Robot, Android & Automaton*, GLOBE & MAIL (Can.), June 15, 2011, at B9; Nolan M. Goldberg & Micah W. Miller, *The Practice of Law in the Age of 'Big Data*,*'* NAT'L L.J. (Apr. 11, 2011), http://www.law.com/jsp/nlj/PubArticleNLJ.jsp?id=1202489457214; Tam Harbert, *Big Data Meets Big Law*, LAW TECH. NEWS (Dec. 27, 2012), www.law.com/jsp/lawtechnologynews/PubArticleLTN.jsp?id=1202555605051; Farhad Manjoo, *Will Robots Steal Your Job?*, SLATE (Sept. 29, 2011, 2:42 AM), http://www.slate.com/articles/technology/robot_invasion/2011/09/will_robots_steal_your_job_5.html; *see also* NEIL RICKMAN & JAMES M. ANDERSON, INNOVATIONS IN THE PROVISION OF LEGAL SERVICES IN THE UNITED STATES: AN OVERVIEW FOR POLICYMAKERS (2011), *available at* http://www.rand.org/content/dam/rand/pubs/occasional_papers/2011/RAND_OP354.pdf.

[3] *See* William D. Henderson & Rachel M. Zahorsky, *Paradigm Shift*, A.B.A. J., July 2011, at 40, 42, 47.

[4] *Id.* at 40.

[5] *Id.* at 40–41.

Aided by design and new alternative delivery models, legal information technology is the centerpiece of the "new normal."[6] Such innovative technologies include platforms designed to help drive down legal costs for potential clients at all price points—from a simple consumer using LegalZoom.com[7] to the sophisticated general counsel[8] applying informatics techniques to lower his or her company's legal bill.[9]

For better or for worse, when it comes to building software, a nontrivial subset of tasks undertaken by lawyers is subject to automation. In this vein, law is similar to other white-collar industries.[10] The bundle of skills associated with the practice of law falls on a continuum where a number of basic tasks have already been displaced by computation, automation, and "soft" artificial intelligence.[11] Faced with cost pressures, clients and law firms are leveraging legal information technology to either automate or semi-automate tasks previously performed by teams of lawyers.[12] Namely, a series of first-

[6] *See* Paul Lippe, *Welcome to 'the New Normal,'* A.B.A. J. (Oct. 13, 2010, 5:31 PM), http://www.abajournal.com/legalrebels/new_normal/.

[7] LEGALZOOM, http://www.legalzoom.com/ (last visited May 10, 2013) (advertising low-cost, legal document creation).

[8] *See* David B. Wilkins, *Team of Rivals? Toward a New Model of the Corporate Attorney–Client Relationship*, 78 FORDHAM L. REV. 2067, 2085–87 (2010) (noting that corporations have begun to trim the number of firms they use by creating preferred provider networks). *See generally* Larry E. Ribstein, *Delawyering the Corporation*, 2012 WIS. L. REV. 305 (describing how the expanded role of corporate counsel and the use of new legal technologies has affected the legal market). It is unclear whether this is a permanent feature of the market. It really depends upon the submarket in which the general counsel is working. As described *infra* in Part II.A, many law divisions and their general counsels (GC) are moving toward predictive analytics in order to drive down rates. Consolidation is obviously a helpful part of that conversation as GCs that spend more and are effective negotiators should be able to drive down their legal costs.

[9] *See infra* Part II.A. Leveraging more than fifteen billion dollars in legal spending data, the TyMetrix Division of the legal informatics conglomerate Wolters Kluwer has published the *2012 Real Rate Report*, which advises corporate counsels and other sophisticated clients of the actual rate (not the rack rate) charged by law firms in a number of major metropolitan areas. *See Products*, TYMETRIX, http://tymetrix.com/products/legal-analytics/2/2012-real-rate-report/ (last visited May 10, 2013).

[10] *See generally* Timothy F. Bresnahan et al., *Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence*, 117 Q.J. ECON. 339 (2002); Antonio Regalado, *When Machines Do Your Job*, MIT TECH. REV. (July 11, 2012), http://www.technologyreview.com/news/428429/when-machines-do-your-job/ (demonstrating the more general principle in the domain of skilled labor and previewing what is likely to be seen in white-collar domains such as legal services). For a useful analogy involving another white-collar industry, consider the case of finance. As outlined *infra* in Part II.C, over the past generation finance has transitioned from an industry dominated by "mental models" to one driven by quantitative prediction.

[11] *Cf.* Bresnahan et al., *supra* note 10, at 344.

[12] *See* John Markoff, *Armies of Expensive Lawyers, Replaced by Cheaper Software*, N.Y. TIMES, Mar. 5, 2011, at A1; Joe Palazzolo, *Why Hire a Lawyer? Computers Are Cheaper*, WALL ST. J., June 18, 2012, at B1.

generation innovations, such as e-discovery[13] and automated document assembly,[14] already has imposed significant consequences on the legal services market.[15] Like many industries before it, the march of automation, process engineering, informatics, and supply chain management will continue to operate and transform our industry.[16] Informatics, computing, and technology are going to change both what it means to practice law and to "think like a lawyer." When it comes to the application of the leading ideas in computation,

[13] *See, e.g.*, William P. Barnette, *Ghost in the Machine:* Zubulake *Revisited and Other Emerging E-Discovery Issues Under the Amended Federal Rules*, 18 RICH. J.L. & TECH. 4 (2012), http://jolt.richmond.edu/v18i3/article11.pdf (describing the "increasing prevalence and cost of e-discovery" and resulting disputes); Richard L. Marcus, *E-Discovery & Beyond: Toward* Brave New World *or* 1984*?*, 25 REV. LITIG. 633 (2006) (describing the potential impact of digital technology on litigation and the issues raised dealing with e-discovery); Carey Sirota Meyer & Kari L. Wraspir, *E-Discovery: Preparing Clients for (and Protecting Them Against) Discovery in the Electronic Information Age*, 26 WM. MITCHELL L. REV. 939 (2000) (describing how to address e-discovery issues); Rebecca N. Shwayri, *Preserving the Needle in the Electronic Haystack: Proposed Federal Rule Amendments and Their Impact on E-Discovery*, 38 J. LEGIS. 118 (2012) (analyzing the impact of e-discovery on ESI preservation obligations and the framework for preservation); Salvatore J. Bauccio, Comment, *E-Discovery: Why and How E-Mail Is Changing the Way Trials Are Won and Lost*, 45 DUQ. L. REV. 269 (2007); *see also supra* note 8. E-discovery is an extremely active area of modern practice and the tools applied in this domain have shifted some of the profits associated with document review to software and third-party vendors. Law firms are fighting to retain the remaining work by developing extensive litigation support departments devoted to executing various task in the e-discovery work flow. There is some evidence that this approach is working. *See* Monica Bay, *Survey Shows Surge in E-Discovery Work at Law Firms and Corporation*s, LAW TECH. NEWS (July 6, 2012), http://www.law.com/jsp/lawtechnologynews/PubArticleLTN.jsp?id=1202561944663&Survey_Shows_Surge_in_EDiscovery_Work_at_Law_Firms_and_Corporations; Gina Passarella, *Law Firms as E-Discovery Vendors? Could Be*, LEGAL INTELLIGENCER (Sept. 27, 2012), http://www.law.com/jsp/pa/PubArticlePA.jsp?id=1202572729762. The move to second-generation e-discovery tools, such as predictive coding, threatens to once again shift the labor-versus-software distribution in favor of the machines. *See* Evan Koblentz, *Judge Carter OKs Peck's Predictive Coding Decision in 'Da Silva Moore*,*'* LAW TECH. NEWS (Apr. 26, 2012), www.law.com/jsp/lawtechnologynews/PubArticleLTN.jsp?id=1202550377104.

[14] *See* Marc Lauritsen, *Fall in Line with Document Assembly: Applications to Change the Way You Practice*, LAW OFF. COMPUTING, Feb./Mar. 2006, at 71; Darryl R. Mountain, *Disrupting Conventional Law Firm Business Models Using Document Assembly*, 15 INT'L J.L. & INFO. TECH. 170 (2007); Elizabeth J. Goldstein, *Kiiac's Contract Drafting Software: Ready for the Rapids?*, LAW TECH. NEWS (May 18, 2012), http://www.law.com/jsp/lawtechnologynews/PubArticleLTN.jsp?id=1202555105751&Kiiacs_Contract_Drafting_Software_Ready_for_the_Rapids; Richard S. Granat, *Document Assembly over the Internet*, AM. B. ASS'N (Dec. 2011), http://www.americanbar.org/content/dam/aba/publications/law_practice_today/document-assembly-over-the-internet.authcheckdam.pdf; Stephanie Francis Ward, *Kingsley Martin's Analysis Software Spots Contract Flaws*, A.B.A. J. (Sept. 18, 2012, 9:00 AM), http://www.abajournal.com/legalrebels/article/kingsley_martin_a_better_benchmark/.

[15] *See* Henderson & Zahorsky, *supra* note 3, at 42–44, 46.

[16] *See* RICHARD SUSSKIND, THE END OF LAWYERS? RETHINKING THE NATURE OF LEGAL SERVICES (2010) (predicting changes in the legal market, including the commoditization of such sources and the implementation of IT solutions, and describing the consequences for the legal industry); *see also* THOMAS D. MORGAN, THE VANISHING AMERICAN LAWYER (2010); Johnathan Jenkins, Note, *What Can Information Technology Do for Law?*, 21 HARV. J.L. & TECH. 589, 597, 604 (2008).

informatics, and other allied disciplines, the market for legal services lags behind many other industries.[17] In other words, yesterday's fast is today's slow, and this is only the beginning.

Aided by growing access to large bodies of semi-structured legal information, the most disruptive of all possible displacing technologies—quantitative legal prediction (QLP)—now stands on the horizon. Although different variants of QLP exist, the march toward quantitative legal prediction will define much of the coming innovation in the legal services industry. And it will occur whether you like it or not.

Do I have a case? What is our likely exposure? How much is this going to cost? What will happen if we leave this particular provision out of this contract? How can we best staff this particular legal matter? These are core questions asked by sophisticated clients such as general counsels, as well as consumers at the retail level. Whether generated by a mental model or a sophisticated algorithm, prediction is a core component of the guidance that many lawyers offer. Indeed, it is by generating informed answers to these types of questions that many lawyers earn their respective wages.

Every single day lawyers and law firms are providing predictions to their clients regarding the likely impact of choices in business planning and transactional structures, as well as their prospects in litigation and the costs associated with its pursuit. How are these predictions being generated? Precisely what data or model is being leveraged? Could a subset of these predictions be improved by various forms of outcome data drawn from a large number of "similar" instances? Simply put, the answer is yes. Quantitative legal prediction already plays a significant role in certain practice areas and this role is likely to increase as greater access to appropriate legal data becomes available. This Article is dedicated to highlighting the coming age of quantitative legal prediction with the hopes that entrepreneurial lawyers, law

[17] *See* Manjoo, *supra* note 2 ("The legal industry is one of the few remaining outposts of the corporate world whose operations are dictated mainly by human experience."). In particular, this Article is devoted to highlighting one such technology—predictive analytics. In many other industries other than law, "big data" and predictive analytics have already obtained a significant foothold. *See generally* JAMES MANYIKA ET AL., MCKINSEY GLOBAL INST., BIG DATA: THE NEXT FRONTIER FOR INNOVATION, COMPETITION, AND PRODUCTIVITY 1 (2011), *available at* http://www.mckinsey.com/mgi/publications/big_data/index.asp; Steve Lohr, *Learning the Power of Teamwork in a Netflix Race for $1 Million*, N.Y. TIMES, July 28, 2009, at B1; *The Data Deluge*, ECONOMIST, Feb. 27, 2010, at 11; Clive Thompson, *What Is I.B.M.'s Watson?*, N.Y. TIMES, June 20, 2010, (Magazine), at 30; Steven Levy, *The AI Revolution Is On*, WIRED (Dec. 27, 2010, 12:00 PM), http://www.wired.com/magazine/2010/12/ff_ai_essay_airevolution/.

students, and law schools will take heed and prepare to thrive in the new ordering.[18]

As it is important to begin with a discussion of the broader environment, Part I identifies the underlying trends that are driving the future of technology, and in turn, legal information technology—"Big Data," Moore's Law, and the "Artificial Intelligence Revolution." Transitioning from the general to the applied case of the market for legal services, Part II describes the emerging age of data-driven law practice with specific attention to quantitative legal prediction. QLP is invading a number of components of the legal service industry, including prediction of cost, outcomes, and potential financial exposure in various legal disputes. Despite all of its potential, Part III describes some of the limits inherent in developing prediction models, including the underlying system volatility and the proper modeling of all relevant dynamics. Part IV concludes with a brief perspective on legal education and the future of the legal services industry.

I. MOORE'S LAW, KRYDER'S LAW, THE AI REVOLUTION, AND EVER-EXPANDING POSSIBILITY FRONTIER

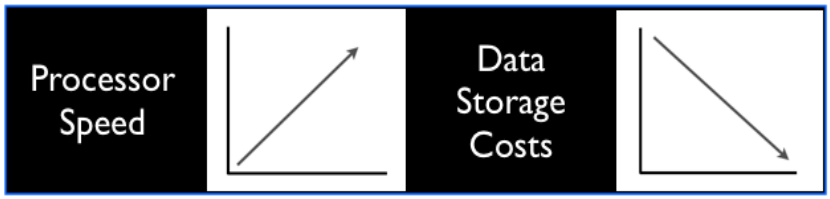
This is the era of "Big Data" and soft artificial intelligence.[19] Increases in computing power and decreases in data storage costs—taken together with significant improvements in machine learning and artificial intelligence—threaten to disrupt white-collar industries in much the manner that process engineering and automation reset the labor-versus-capital tradeoff in blue-

[18] There exist a variety of very strong critiques of the modern law school. Perhaps, the most visceral of these critiques comes from the so-called scamblog movement. Among the more focused of these critiques are the arguments relating to the diminished return on investment (ROI) associated with a law degree. For documentation of the scamblog movement, see Daniel D. Barnhizer, *Cultural Narratives of the Legal Profession: Law School, Scamblogs, Hopelessness, and the Rule of Law*, 2012 MICH. ST. L. REV. 663. Among other things, the goal of this Article is to highlight one area in which law schools could help increase the ROI attached to a law degree. Namely, preparing their students for the era of data-driven law practice. In addition to this external critique offered by "scambloggers" are critiques by insiders such as Professor Tamanaha. *See, e.g.*, BRIAN Z. TAMANAHA, FAILING LAW SCHOOLS (2012). Building on Professor Tamanaha's work—particularly the call for product differentiation in the market for legal education—this Article highlights a way forward for entrepreneurial law schools that embrace the *data-driven era of law practice*.

[19] *See supra* note 13 and accompanying text; *see also Community Cleverness Required*, NATURE, Sept. 4, 2008, at 1 (introducing an entire issue of *Nature* that "examines what big data sets mean for contemporary science"); *The Data Deluge*, *supra* note 17 (noting how the "data deluge is already starting to transform business, government, science and everyday life"); Conrad Quilty-Harper, *10 Ways Data Is Changing How We Live*, TELEGRAPH (Aug. 25, 2010, 2:56 PM), http://www.telegraph.co.uk/technology/7963311/10-ways-data-is- changing-how-we-live.html.

collar industries.[20] Before considering the specific contours of the legal services market, it is worth exploring several related but differentiable trends that are operating to change the future of work in many industries, including the legal services industry.

Figure 1: Moore's Law & Kryder's Law



A. Moore's Law

For more than forty years, the transistor count (speed) of the world's leading central processing unit (CPU) has doubled every twelve to eighteen months.[21] This simple fact has helped usher in a significant amount of technology innovation.[22] Named for Intel Corporation founder Gordon Moore, Moore's Law was first outlined in his well-cited 1965 article.[23] While Moore originally predicted this trend would last for close to a decade,[24] the exact timeline for the respective doubling has been consistently recalibrated, with time windows such as eighteen to thirty-six months typically predicted.[25] Figure 2 highlights the rapid growth in CPU transistor count, which is the typical metric used to benchmark CPU speed. Of course, doubling in the early years featured increases from 5,000 hertz to 10,000 hertz, while recent patterns

[20] *See, e.g.*, STANLEY ARONOWITZ & WILLIAM DIFAZIO, THE JOBLESS FUTURE (2d ed. 2010); GEORGES FRIEDMANN, THE ANATOMY OF WORK (Transaction Publishers 1992) (1961); JEREMY RIFKIN, THE END OF WORK (1995).

[21] *See, e.g.*, Chris A. Mack, *Fifty Years of Moore's Law*, 24 IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING 202, 202–03 (2011).

[22] *See* Samuel Arbesman, *The Hidden Rules That Shape Human Progress*, BBC (Oct. 18, 2012), http://www.bbc.com/future/story/20121018-hidden-rules-of-human-progress.

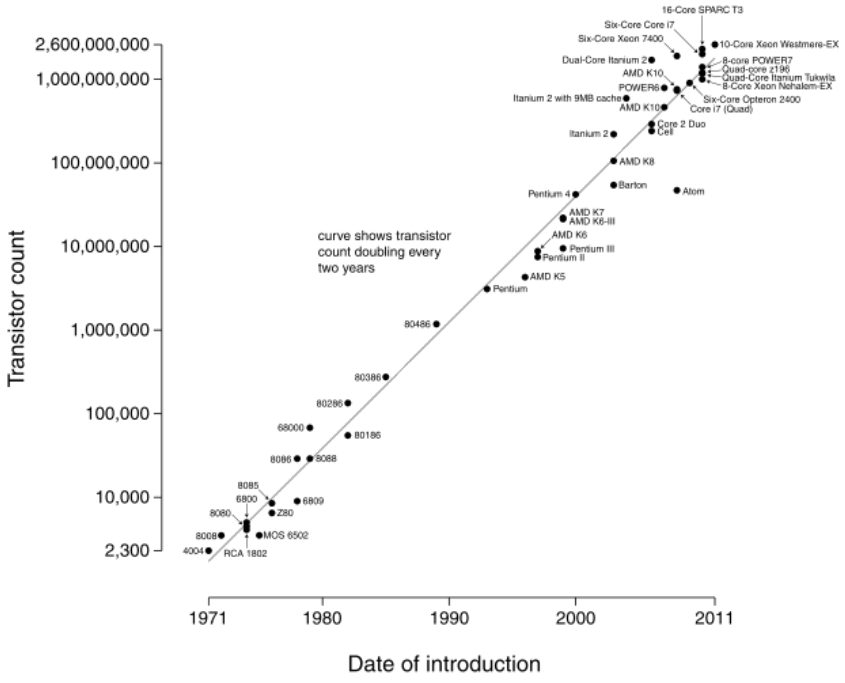
[23] *See* Gordon E. Moore, *Cramming More Components onto Integrated Circuits*, ELECTRONICS, Apr. 19, 1965, at 114, *reprinted in* 86 PROC. IEEE 82 (1998).

[24] *Id.*

[25] *See, e.g.*, UNDERSTANDING MOORE'S LAW: FOUR DECADES OF INNOVATION (David C. Brock ed., 2006); Mark Lundstrom, *Moore's Law Forever?*, 299 SCIENCE 210, 210 (2003); Robert R. Schaller, *Moore's Law: Past, Present, and Future*, IEEE SPECTRUM, June 1997, at 53, 54–55; Scott E. Thompson & Srivatsan Parthasarathy, *Moore's Law: The Future of Si Microelectronics*, MATERIALS TODAY, June 2006, at 20, 21.

of doubling have ushered in increases in transistor counts on the order of billions.[26] Although there is reason to believe it will ultimately abate,[27] it appears this process of doubling will continue for the foreseeable future.[28]

Figure 2: Moore's Law—Microprocessor Transistor Counts, 1971–2011[29]



[26] *See* UNDERSTANDING MOORE'S LAW, *supra* note 25.

[27] *See, e.g.*, Lundstrom, *supra* note 25, at 211; Karl Rupp & Siegfried Selberherr, *The Economic Limit to Moore's Law*, 24 IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING 1 (2011).

[28] *See* Sharon Gaudin, *Physicist Says Moore's Law Is 'Collapsing,'* COMPUTER WORLD (May 2, 2012, 6:00 AM), http://www.computerworld.com/s/article/9226758/Physicist_says_Moore_s_Law_is_collapsing. *But see* Adam Sneed, *A Brief History of Warnings About the Demise of Moore's Law*, SLATE (May 3, 2012, 4:56 PM), http://www.slate.com/blogs/future_tense/2012/05/03/michio_kako_and_a_brief_history_of_warnings_about_the_end_of_moore_s_law_.html (noting that there is a long history of claims discounting the continued viability of Moore's Law, all of which have proven to be incorrect).

[29] *Moore's Law*, WIKIPEDIA, http://en.wikipedia.org/wiki/Moore%27s_law (last updated Feb. 21, 2013, 8:16 PM).

B. Kryder's Law and Big Data

CPU speed is not responsible alone for rapidly expanding the possibility frontier. Equally important has been the rapid and consistent decline in data storage cost. Kryder's Law (the storage analog to Moore's Law) holds that the decrease in data storage costs follows a pattern similar to, if not exceeding, the pace of the corresponding increase in transistor count.[30] This decrease in storage cost is a key component in the rise of Big Data. Indeed, many commentators have identified this as the age of Big Data and those prepared to deal with this data deluge will drive productivity, innovation, and the future of the economy.[31]

So how big is "BIG"? The target is subjective and ever moving, but the conventional understanding refers to "datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze."[32] Given that this sort of definition is somewhat illusory, it is useful to offer a benchmarked perspective on the ever-expanding Big Data frontier. Indeed, a backward-looking approach is probably the best way to understand just how big today's "BIG" actually is.

A very common unit of measurement familiar to many is the gigabyte. The typical thumb drive distributed at a tradeshow or academic conference commonly holds several "gigs" of data. A gigabyte can store 10^9 (1,000,000,000) bytes of information.[33] In terms of information content, one gigabyte is the equivalent of roughly seven minutes of high definition (HD) video or about twenty yards of books on a typical shelf.[34] The price of a gigabyte has declined very rapidly over the past several decades. In 1981, a gigabyte cost about $300,000. In 1997, it cost around $100, and by 2011 it cost about $0.10.

---

[30] *See* Chip Walter, *Kryder's Law*, SCI. AM., Aug. 2005, at 32.

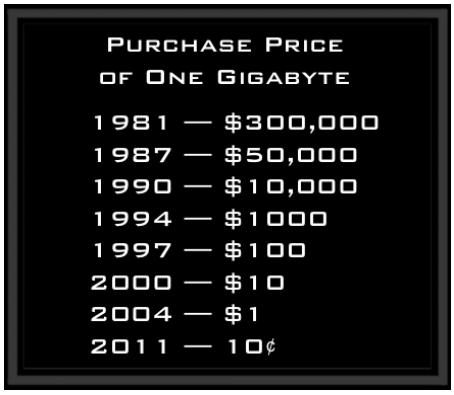
[31] *See, e.g.*, Joseph Walker, *Meet the New Boss: Big Data*, WALL ST. J., Sept. 20, 2012, at B1; Lisa Arthur, *The Surprising Way eBay Used Big Data Analytics to Save Millions*, FORBES (Aug. 23, 2012, 9:11 AM), http://www.forbes.com/sites/lisaarthur/2012/08/23/the-surprising-way-ebay-used-big-data-analytics-to-save-millions/; Tam Harbert, *Big Data, Big Jobs?*, COMPUTER WORLD (Sept. 20, 2012, 6:00 AM), http://www.computerworld.com/s/article/9231445/Big_data_big_jobs.

[32] *See* MANYIKA ET AL., *supra* note 17, at 1 ("This definition [of Big Data] is intentionally subjective and incorporates a moving definition of how big a dataset needs to be in order to be considered big data—i.e., we don't define big data in terms of being larger than a certain number of terabytes . . . . We assume that, as technology advances over time, the size of datasets that qualify as big data will also increase.").

[33] *All Too Much*, ECONOMIST, Feb. 27, 2010, at 3.

[34] *How Much Is a Petabyte?*, MOZY BLOG (July 2, 2009), http://mozy.com/blog/misc/how-much-is-a-petabyte/.

Figure 3: Decreasing Data Storage Costs[35]



What was previously available only to those operating at the level of enterprise computing has trickled down to the average consumer. Indeed, moving up the scale from the gigabyte is the terabyte (10^12 or 1,000,000,000,000 bytes) and the petabyte (10^15 or 1,000,000,000,000,000 bytes).[36] The terabyte is now commonly available at the consumer level and the petabyte should be available at the retail level in the coming years.[37]

Simply put, a petabyte is a lot of data.[38] By way of example, one petabyte is equal to more than thirteen consecutive years of HD video and fifty petabytes is roughly equal to the information content of the "entire written works of mankind from the beginning of recorded history in all languages."[39] A major data storage company has predicted that within the next five years it will be possible for the retail consumer to purchase a petabyte for approximately $750.[40] Thus, in principle, an individual or organization will be

[35] *See* Matthew Komorowski, *A History of Storage Cost*, MKOMO.COM, http://www.mkomo.com/cost-per-gigabyte (last visited May 10, 2013).

[36] *All Too Much*, *supra* note 33, at 3.

[37] The precise timeline is an open question. *See* David S. H. Rosenthal et al., *The Economics of Long-Term Digital Storage*, *in* THE MEMORY OF THE WORLD IN THE DIGITAL AGE: DIGITIZATION AND PRESERVATION 513 (Luciana Duranti & Elizabeth Shaffer eds., 2012), *available at* http://www.unesco.org/webworld/download/mow/mow_vancouver_proceedings_en.pdf; *see also supra* note 34.

[38] *See, e.g.*, *The Petabyte Age: Because More Isn't Just More—More Is Different*, WIRED (June 23, 2008), http://www.wired.com/science/discoveries/magazine/16-07/pb_intro.

[39] *How Much Is a Petabyte?*, *supra* note 34.

[40] *1.2 Petabytes of Storage*, P2PNET (Feb. 15, 2006, 9:39 AM), http://www.p2pnet.net/story/7929 (quoting Michael Thomas, owner of Colossal Storage, as stating, "I'd say we can expect a finished product to

able to store all of the written works of humanity for approximately $37,500. This is what is meant when folks talk about Big Data, and even more sophisticated developments are occurring at the level of enterprise computing.

C. The AI Revolution Is Here but It Is Nothing like We Expected[41]

Of course, data storage and processor speed alone are insufficient to generate the sort of aggregated insights that drive productivity and innovation. Facilitated by both Moore's Law and Kryder's Law, the final leg of this new age of productivity is the Artificial Intelligence (AI) Revolution.[42] The AI Revolution is indeed ongoing, but it is not the fanciful world drawn up by futurists in the 1950s and 1960s. The AI dreams of that bygone era were centered on "mimicking the logic-based reasoning that human brains were thought to use."[43] That was a borderline fruitless effort and is commonly referred to as the "AI winter."[44] Summer has come to artificial intelligence, but this has come by focusing on the discrete tasks that current computers are actually well suited to perform.[45]

Today's AI is "soft AI" because it attempts to mimic human intelligence in outcomes, but not in its underlying processes.[46] It turns out that we still have only a very limited understanding of the human brain, and thus the direct artificial intelligence attempts to model its internal processes have borne little fruit.[47] By contrast, for a certain class of problems, the outcome-based approach has been quite successful. These approaches generally place a black box around the internal dynamics used by human reasoners and instead model and predict the choices made by actors.[48] Using large segments of observational data, today's soft AI is built upon modeling what people actually do, thereby allowing a machine to probabilistically emulate their behavior

be on the market in about four to five years" and noting "the cost would probably be in the range of $750 each").

41 *See* Levy, *supra* note 17.

42 *Id.*

43 *Id.*

44 *Id.*

45 *See id.*

46 *See* Robert Emmett Mueller, *The Leonardo Paradox: Imagining the Ultimately Creative Computer*, 23 LEONARDO 427, 427 (1990) (highlighting the distinction between "hard" and "soft" artificial intelligence).

47 Levy, *supra* note 17.

48 Examples of the black-box approach are numerous. For a broad overview of application of so-called black-box models in machine learning, see generally CHRISTOPHER M. BISHOP, PATTERN RECOGNITION AND MACHINE LEARNING (2006); Klaus-Robert Müller et al., *An Introduction to Kernel-Based Learning Algorithms*, 12 IEEE TRANSACTIONS ON NEURAL NETWORKS 181 (2001).

under analogous conditions.[49] This "inverse approach"[50] is the core of modern machine learning, and it has led to a number of breakthrough technologies previously thought to be either impossible or only possible in the far-distant future.[51]

D. The Second Half of the Chessboard?

Does the combination of Moore's Law, Big Data, and the "soft AI" revolution represent a fundamental transformation, rather than some sort of predictable, incremental change? It is pretty common for those with a vested interest in the status quo to argue that what they do is outside the possibility frontier. When an individual argues, "You cannot replace me with a machine," it is useful to begin by evaluating his or her basis for that belief. Do they have the requisite technical understanding to evaluate what is and is not possible? Typically, claims of this type represent more of a hope than a grounded analysis. This is the age of robotics, AI, and the "race against the machine."[52] Be wary of backward-looking statements such as, "That was already tried and did not work." The ground is rapidly shifting. Peril and possibility, as well as disruption, are fundamental features of our times.

Mistaking exponential change for linear change is a very common mistake. Indeed, our desire to linearize a nonlinear function is a well-studied cognitive bias.[53] A classic prism through which the nonlinear march of technological progress is sometimes described is the chessboard problem.[54] There are many versions of this story, but it generally surrounds payment by a ruler to the local

---

[49] Levy, *supra* note 17.

[50] For a description of the conceptual distinction between a forward and inverse approach, see *infra* Part III.A.

[51] *See infra* Part I.D.1.

[52] *See* ERIK BRYNJOLFSSON & ANDREW MCAFEE, RACE AGAINST THE MACHINE (2011); *see also* Levy, *supra* note 17.

[53] *See, e.g.*, Maya Bar-Hillel, *On the Subjective Probability of Compound Events*, 9 ORGANIZATIONAL BEHAV. & HUM. PERFORMANCE 396 (1973); Simon Kemp, *Perception of Changes in the Cost of Living*, 5 J. ECON. PSYCHOL. 313 (1984); Andrew J. Mackinnon & Alexander J. Wearing, *Feedback and the Forecasting of Exponential Change*, 76 ACTA PSYCHOLOGICA 177, 177–78 (1991); Peter A. O'Donnell et al., *An Experimental Study of the Impact of a Computer-Based Decision Aid on the Forecast of Exponential Data*, *in* PACIS 1997 PROCEEDINGS 279 (1997); Willem A. Wagenaar & Han Timmers, *The Pond-and-Duckweed Problem; Three Experiments on the Misperception of Exponential Growth*, 43 ACTA PSYCHOLOGICA 239 (1979); *see also* Richard Webby & Marcus O'Connor, *Judgemental and Statistical Time Series Forecasting: A Review of the Literature*, 12 INT'L J. FORECASTING 91 (1996).

[54] *See* BRYNJOLFSSON & MCAFEE, *supra* note 52, at 19 (revisiting Raymond Kurzweil's description of the technology and the chessboard problem).

individual who invented the chessboard.[55] Whether it be a Persian king or Indian leader, the basic thrust is as follows: The ruler was so pleased with the game of chess that he allowed the inventor of the game to name a prize for the invention.[56] The inventor, who was both brilliant and wise, asked the king to provide him with one grain of rice for the first square on the chessboard, two grains for the second square, and four grains for the third square with continued doubling until payment was received for all 64 squares.[57] The ruler quickly accepted the inventor's offer and was even offended that the inventor was asking for such a low price.[58] The story ends with the inventor becoming the new leader because the promise yielded a pile of rice that was larger than the size of the tallest mountain.[59]

The ruler's fatal mistake was equating linear growth with exponential growth. In the early portions of the chessboard, the returns associated with each doubling are quite close. For example, assume that the ruler were to give 100 grains for each square. This function—*100x*—would yield the follow series: 100, 200, 300, 400, 500, 600, 700, 800, and so on. This would look quite similar to values such as 1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024 (i.e., 2^0, 2^1, 2^2, 2^3, and so on). Eventually, the exponential function overtakes and rapidly passes the values returned by the linear function—6400 versus 2^{64} grains on the final square alone.

[55] For the mathematics of the question, see THEONI PAPPAS, THE JOY OF MATHEMATICS 17 (rev. ed. 1989); *see also* Eric W. Weisstein, *Wheat and Chessboard Problem*, WOLFRAM MATHWORLD, http://mathworld.wolfram.com/WheatandChessboardProblem.html (last visited May 10, 2013).

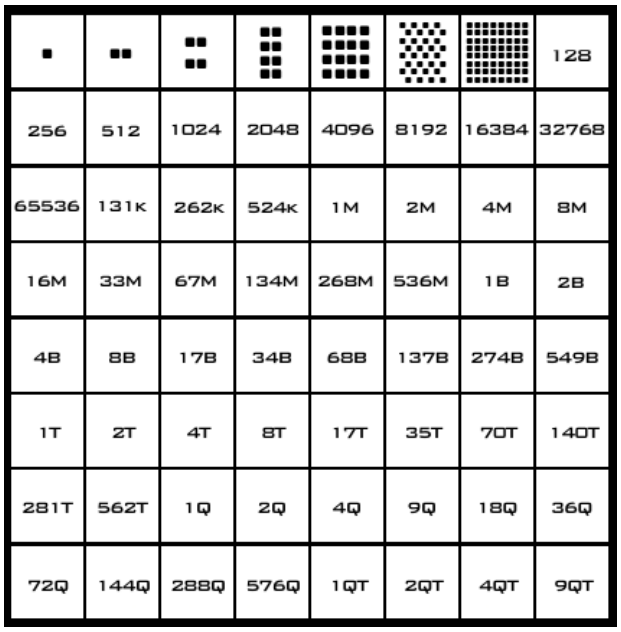
[56] *Wheat and Chessboard Problem*, WIKIPEDIA, http://en.wikipedia.org/wiki/Wheat_and_chessboard_problem (last modified Apr. 5, 2013).

[57] *Id.*

[58] *Id.*

[59] *Id.*

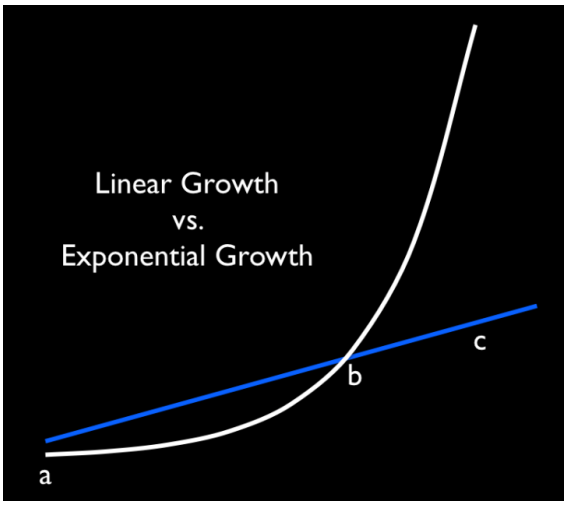
Figure 4: The Grains on the Chessboard



For a visual depiction of the question, consider Figure 5 below. Again, in the lower ranges of values along the X-axis, the linear and exponential growth functions return very similar values. In fact, if a researcher were to observe values between points *A* and *B* it would be reasonable to conclude that both functions were quite similar. Only after observing values between points *B* and *C* would it be clear that the underlying functions were quite distinct. This simple example points to the more general bias faced by human reasoners. Namely, it is difficult for decision makers to distinguish linear and nonlinear growth processes,[60] and undoubtedly this bias is particularly acute when the most recently observed range of values suggests the function behaves as if it were linear.

[60] *See supra* note 53.

Figure 5: Linear Growth Versus Exponential Growth



Arguably we are transitioning to the second half of the chessboard where new technological possibilities are consistently presenting themselves. Each doubling of an already massive number is extremely significant because "[e]xponential increases initially look a lot like standard linear ones, but they're not. As time goes by—as we move into the second half of the chessboard—exponential growth confounds our intuition and expectations."[61] In the world of technology, the synergy of Moore's Law, Big Data, and the AI Revolution is doing precisely this. With each doubling of processor speed, halving of data storage costs, and major advances in machine learning, the possibility frontier is opening up and doing so at a drastically nonlinear rate.

1. "You Cannot Replace What I Do with a Computer"—Aspirational Spelling, Driverless Cars, and IBM's Watson

Opportunities are created with each step forward for those who do not fall prey to the notion that elements of their respective jobs cannot be subjected to some form of automation, process engineering, data analytics, etc. Thus, before discussing the coming breakthrough technologies in data-driven law practice, it is useful to enhance one's understanding of the current state of affairs with specific reference to three concrete instances where the mixture of processor

[61] *See* BRYNJOLFSSON & MCAFEE, *supra* note 54, at 19; *see also* Levy, *supra* note 17.

speed, data storage, and soft AI have opened up the possibility frontier. Even if one knows little about this topic, it should be clear that in an age of aspirational spelling, driverless cars, and IBM's Watson, the practice of law is likely to change.

a. Welcome to the World of Aspirational Spelling

Human reasoners (well, many of them other than me)[62] learn the rules of spelling at an early age. They also learn how to apply the relevant exceptions to those rules and in many instances are able to effortlessly write paragraph upon paragraph with very limited errors.[63] From an algorithmic perspective, the difficult question for AI scholars is how to mimic that behavior. How can a researcher reproduce the simple outcome that so many humans are easily able accomplish? Historically, spelling and spell checking were "hard" problems and the "best" available solutions to the problems were only moderately successful.[64] Many researchers and technology companies, including Microsoft, invested millions of dollars and countless hours trying to develop a robust, flexible, and scalable spell-checking algorithm.[65] Big Data broke the logjam and pioneered what is the current approach to spell checking.[66]

Google succeeded where others had previously failed by leveraging massive "click data" and more than three billion daily queries to harvest out probabilistically likely matches to commonly misspelled words.[67] The key is to develop an approach that generates a series of often-correct answers to the specific problem. As individuals click through to specific answers, that data is harvested so that, in the aggregate, the program quickly begins to approximate

[62] I am not one of them but, as I will describe in a moment, I am on the right side of history because this is the age of *aspirational spelling*. If you can reasonably aspire to spell a word, then you can spell a word with the help of Google and its millions of users.

[63] *See, e.g.*, Kristine F. Anderson, *The Development of Spelling Ability and Linguistic Strategies*, 39 READING TCHR. 140, 140–42 (1985); Carol Sue Englert et al., *Spelling Unfamiliar Words by an Analogy Strategy*, 19 J. SPECIAL EDUC. 291 (1985); Sandra Wilde, *Learning to Spell and Punctuate: A Study of Eight- and Nine-Year-Old Children*, 2 LANGUAGE & EDUC. 35 (1988).

[64] *See Clicking for Gold*, ECONOMIST, Feb. 27, 2010, at 9.

[65] *Id.* ("Microsoft says it spent several million dollars over 20 years to develop a robust spell-checker for its word-processing program. But Google got its raw material free: its program is based on all the misspellings that users type into a search window and then "correct" by clicking on the right result. With almost 3 billion queries a day, those results soon mount up.").

[66] *See id.*

[67] *Id.*; Peter Norvig, *How to Write a Spelling Corrector*, NORVIG.COM, http://www.norvig.com/spell-correct.html (last visited May 10, 2013).

the correct answer in a wide number of instances.[68] Specifically, the model + "click data" is stored in a massive relational database or network of linked correct and incorrect answers, all of which are consistently being updated as new queries are being typed into the Google main page.[69]

Figure 6: Google's Spell Checking

This is a simple example highlighting the artificial intelligence of today: machines mimicking the same outcomes that are typically produced by humans—even if their specific internal processes might differ. Is Spelling 1.0 thinking? Okay, probably not. However, this is the age of aspirational spelling where, in many instances, if you can aspire to spell something, then Google can help take you the rest of the way.

b. Welcome to the Age of Driverless Cars

In the spring of 2004, the Defense Advanced Research Projects Agency (DARPA) held its first Grand Challenge in the Mojave Desert near the Nevada–California border.[70] The rules were simple: Build a self-driving car that could traverse the 150-mile DARPA course. The first car to pass the finish line would receive a $1 million prize.[71] More than 100 teams registered for the competition.[72] Despite all of the diverse approaches and technical expertise that was assembled and directed at the problem, the 2004 DARPA Grand Challenge was not terribly successful. The longest lasting car only managed to complete seven miles of the course after getting hung up on a rock.[73]

---

[68] For a general description of these and other related uses of Big Data, see Alistair Croll, *The Feedback Economy*, STRATA (Jan. 4, 2012), http://strata.oreilly.com/2012/01/the-feedback-economy.html.

[69] *See Clicking for Gold*, *supra* note 64. *See generally* Croll, *supra* note 68.

[70] *Urban Challenge*, DARPA, http://archive.darpa.mil/grandchallenge/ (last visited May 10, 2013).

[71] *See The DARPA Grand Challenge: A Historic Demonstration of Autonomous Robotic Vehicles*, DARPA, http://archive.darpa.mil/grandchallenge04/sponsor_toolkit/brochure.pdf (last visited May 10, 2013); Marsha Walton, *Robots Fail to Complete Grand Challenge*, CNN (May 6, 2004, 10:44 AM), http://articles.cnn.com/2004-03-14/tech/darpa.race_1_darpa-grand-challenge-desert-tortoise-robots.

[72] Walton, *supra* note 71.

[73] *See id.*

Of course, this is not the end of the story. In the years that followed, several additional competitions were held and significant technical progress was made. Aiding these efforts was the continued march of processor speed increases, data storage decreases, and significant advances in the field of soft artificial intelligence. Fast forward just eight years later and the Google self-driving car is now licensed in the State of Nevada[74] and has completed more than 300,000 miles without causing an accident.[75] Beyond likely being one of the most transformational and disruptive technologies of our time, the self-driving car is emblematic of a deeper move into the second half of the chessboard. The impossible becomes possible, and it does so mighty quickly.

Figure 7: Nevada's Autonomous Car License Plate[76]

c. IBM's Watson Says "Hello World"

"From the TJ Watson Research Center in Yorktown Heights, New York—This is Jeopardy!—The IBM Challenge."[77] On February 14, 2011, famed announcer Johnny Gilbert stepped to the microphone and unveiled the greatest example to date of performance computing that threatens the core of typical white-collar work.[78] The IBM Challenge pitted IBM's Watson versus Brad Rutter and Ken Jennings, the two most successful Jeopardy champions in

[74] John C. Dvorak, *Google's Revolutionary Self-Driving Car*, PCMAG.COM (May 9, 2012), http://www.pcmag.com/article2/0,2817,2404199,00.asp.

[75] Rebecca J. Rosen, *Google's Self-Driving Cars: 300,000 Miles Logged, Not a Single Accident Under Computer Control*, ATLANTIC (Aug. 9, 2012, 12:29 PM), http://www.theatlantic.com/technology/archive/2012/08/googles-self-driving-cars-300-000-miles-logged-not-a-single-accident-under-computer-control/260926/.

[76] *Autonomous Vehicles*, NEV. DEPARTMENT MOTOR VEHICLES, http://www.dmvnv.com/autonomous.htm (last visited May 10, 2013).

[77] *Jeopardy! The IBM Challenge* (CBS television broadcast Feb. 14, 2011), *available at* https://www.youtube.com/watch?v=seNkjYyG3gI.

[78] John Markoff, *Computer Wins on 'Jeopardy!': Trivial, It's Not*, N.Y. TIMES, Feb. 17, 2011, at A1.

history.[79] After the multiday challenge, there was a clear winner—Machines 1, Humans 0.[80] Watson made it look easy.[81] On the edge of facing defeat, Jennings, the 74-time consecutive Jeopardy champion wrote on his video screen: "I, for one, welcome our new computer overlords."[82]

It is hard to understate just how difficult of a problem it is for a machine to compete in a game such as Jeopardy.[83] Topics are wide ranging and include detailed questions in domains such as history, literature, politics, arts and entertainment, and science.[84] Contestants often confront clues that "involve analyzing subtle meaning, irony, riddles, and other complexities in which humans excel and computers traditionally do not."[85] Finally, answers typically must be given very quickly—often in roughly 3 seconds.[86]

Watson accomplishes its task without access to the Internet, and instead uses large bodies of structured and semi-structured data as it interprets text and refines its answers.[87] Watson applies a mixture of technologies including natural language processing (NLP), information retrieval (IR), knowledge representation and reasoning, and machine learning (ML).[88] The complete hardware features 2880 cores, 16 terabytes of RAM,[89] and is the size of 10 refrigerators.[90]

When a new clue is offered, Watson begins by parsing the question into its parts of speech, thereby better understanding the role of each word within the respective clue.[91] This allows Watson to try to determine the call of the

[79] *Id.*

[80] *Id.*

[81] *Id.*

[82] *Id.*

[83] For an additional description of IBM's Watson, as well as the future role of IT in law, see John O. McGinnis & Steven Wasick, *Law: An Information Technology* 30–32 (Nw. Univ. Sch. of Law Pub. Law & Legal Theory Series, Paper No. 12-22, 2012).

[84] *FAQs*, IBM, http://www.research.ibm.com/deepqa/faq.shtml#1 (last visited Feb. 22, 2013).

[85] *Id.*

[86] *See* David Ferrucci et al., *Building Watson: An Overview of the DeepQA Project*, AI MAG., Fall 2010, at 59, 69–70.

[87] *FAQs*, *supra* note 84.

[88] Ferrucci et al., *supra* note 86, at 62.

[89] Tami Deedrick, *It's Technical, Dear Watson*, IBM SYS. MAG. (Feb. 2011), http://www.ibmsystemsmag.com/ibmi/trends/whatsnew/It%E2%80%99s-Technical,-Dear-Watson/.

[90] Eyder Peralta, *Are You Smarter Than a Computer the Size of 10 Refrigerators?*, NPR (Jan. 13, 2011, 1:19 PM), http://www.npr.org/blogs/thetwo-way/2011/01/13/132902908/are-you-smarter-than-a-computer-the-size-of-10-refrigerators. While IBM's Watson is large today, it will undoubtedly grow smaller and smaller in the years to come. The iPhone 14, now with Watson—yeah, that is where this is all heading.

[91] *See* Ferrucci et al., *supra* note 86, at 69–70.

question.[92] Analogous to *breadth first search*, at the earliest stages the computer casts a very wide search for possible information relevant to answering the current clue.[93] Next, it analyzes and scores the resulting information using a proprietary scoring algorithm.[94] This algorithm develops a statistical confidence level for each potential answer.[95] Based upon constantly adapting factors, including the confidence score of the most likely answer, the money held by each of the other players, and the remaining money left on the board, Watson determines whether it will or will not attempt to push its buzzer.[96] All of this, and more, happens in less than 3 seconds!

Figure 8: The IBM Jeopardy Challenge[97]

Similar to other forms of soft AI, it is important to note that Watson does not necessarily mimic the internal processes used by human reasoners. Instead, it mimics the outcomes generated by humans while following a method that is somewhat similar, but not precisely akin, to human reasoning. This is an important conceptual distinction that in part differentiates "hard" AI from "soft" AI. So, while Watson does not always get the answer correct and even makes obvious mistakes,[98] it is a major step forward and points to a very different future for domains where human expertise has historically dominated.

92 *See id.*
93 *Id.* at 71.
94 *Id.* at 72, 74.
95 *Id.* at 74.
96 *Id.* at 75.
97 *IBM's Computer Wins 'Jeopardy!' but . . . Toronto?*, CTV NEWS (Feb. 15, 2011, 11:15 PM), http://www.ctvnews.ca/ibm-s-computer-wins-jeopardy-but-toronto-1.608022.
98 Stephen Baker, *How Could IBM's Watson Think That Toronto Is a U.S. City?*, HUFFINGTON POST (February 16, 2011, 9:08 AM), http://www.huffingtonpost.com/stephen-baker/how-could-ibms-watson-thi_b_823867.html.

In other words, Watson is far more than a demonstration project. It is a working computer system that is actively being applied to a variety of professional domains—most notably the field of medicine (i.e., data-driven medicine)—where individual doctors are called upon to analyze large amounts of information and rapidly execute the best possible judgment.[99]

II. DATA-DRIVEN LAW PRACTICE AND THE AGE OF QUANTITATIVE LEGAL PREDICTION

Do I have a case? What is our likely exposure? How much is this going to cost? Are these documents relevant? What will happen if we leave this particular provision out of this contract? How can we best staff this particular legal matter? These are core questions asked by sophisticated clients such as general counsels as well as consumers at the retail level.

Many lawyers earn their respective wages by generating informed responses to these and other related types of questions. For many years, the answers to these questions have been the exclusive province of human assessment. While sometimes used in a pejorative manner, it is worth noting that such "mental models" can be well specified. In other words, experience can, under certain conditions, dramatically improve one's ability. A seasoned lawyer can draw upon both extensive legal training as well as personal experience developed over years of law practice. At the same time, such individuals are expensive and even experts cannot escape their respective limitations. This is the entry point for quantitative legal prediction.

QLP-based technologies are designed to remedy or supplement the shortcomings of human reasoners. For example, human reasoners are limited in the scope of their observations. They only posses the observational data they have observed. While an experienced lawyer might be familiar with hundreds, if not thousands, of prior events, he or she is unlikely to have observed tens of thousands, hundreds of thousands, or millions of prior events. Thus, when answering the question, "Do I have a case?" an individual's particular understanding of likelihood might be driven by personal observations that are anecdotal, censored, or otherwise not indicative of the true distribution of outcomes. This is particularly problematic for rare events.[100] The best way to

[99] Lucas Mearian, *IBM's Watson Expands Cancer Care Resume*, COMPUTER WORLD (Mar. 23, 2012, 3:28 PM), http://www.computerworld.com/s/article/9225515/IBM_s_Watson_expands_cancer_care_resume.

[100] This is a problem faced by both human reasoners and model builders. *See* Xavier Gabaix, *Power Laws in Economics and Finance*, 1 ANN. REV. ECON. 255 (2009); Paul Goodwin & George Wright, *The Limits of*

remedy these and other related issues is to observe a large-scale and truly representative selection of the relevant event data.

In addition to data censoring issues, Big Data-based prediction engines also help overcome other limitations. When it comes to processing and deriving insights from large-scale data or document sets, humans have important cognitive limitations.[101] Even if one has access to all of the relevant information, without the aid of technology in many cases, it is essentially impossible to completely process all relevant data or its potentially relevant dimensions. It is just too much. Human reasoners have well-documented cognitive biases, such as the availability heuristic, optimism bias, anchoring, confirmation bias, illusion of validity, and the frequency illusion.[102] While the use of a quantitative prediction solution does not necessarily eliminate all of these potential limits, the transparency associated with developing predictive models can ultimately help engineer around some of these important and well-known human deficits.

In sum, for the appropriate tasks, the age of quantitative legal prediction is about a mixture of humans and machines working together to outperform either working in isolation. The equation is simple: Humans + Machines > Humans or Machines.

A. Predicting the Expected Bill

How much is this going to cost? From both the sophisticated client as well as the average consumer, this is a major question raised prior to, or early in, legal representation. Particularly for the retail and small-scale business clients,

Forecasting Methods in Anticipating Rare Events, 77 TECHNOLOGICAL FORECASTING & SOC. CHANGE 355 (2010); Spyros Makridakis et al., *Forecasting and Uncertainty in the Economic and Business World*, 25 INT'L J. FORECASTING 794 (2009); Didier Sornette, *Dragon-Kings, Black Swans and the Prediction of Crises*, 2 INT'L J. TERRASPACE SCI. & ENGINEERING 1 (2009); Nassim Nicholas Taleb, *Black Swans and the Domains of Statistics*, 61 AM. STATISTICIAN 198 (2007) (book review); *see also* NASSIM NICHOLAS TALEB, THE BLACK SWAN: THE IMPACT OF THE HIGHLY IMPROBABLE (2010).

[101] *See, e.g.*, Thomas Hills & Ralph Hertwig, *Why Aren't We Smarter Already: Evolutionary Trade-Offs and Cognitive Enhancements*, 20 CURRENT DIRECTIONS PSYCHOL. SCI. 373 (2011); Dwight W. Read, *Working Memory: A Cognitive Limit to Non-Human Primate Recursive Thinking Prior to Hominid Evolution*, EVOLUTIONARY PSYCHOL. 676 (2008), http://www.epjournal.net/wp-content/uploads/ep06676714.pdf; *see also* Wolfgang Gaissmaier et al., *An Ecological Perspective to Cognitive Limits: Modeling Environment-Mind Interactions with ACT-R*, JUDGMENT & DECISION MAKING 278 (2008), http://journal.sjdm.org/bn7.pdf.

[102] *See, e.g.*, Daniel Kahneman & Amos Tversky, *Subjective Probability: A Judgment of Representativeness*, 3 COGNITIVE PSYCHOL. 430, 431 (1972); Amos Tversky & Daniel Kahneman, *Availability: A Heuristic for Judging Frequency and Probability*, 5 COGNITIVE PSYCHOL. 207 (1973); Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 SCIENCE 1124 (1974).

the agency problems surrounding billing are a significant source of complaints. Clients worry that their lawyer or law firm is padding the bill—whether by charging too much per hour or adding largely unnecessary hours. In an effort to defend their fees or more generally avoid the commoditization of their work, lawyers commonly highlight the unique properties of the current dispute, transaction, or matter. So the mantra goes, "These things are hard to predict—you know every case is different." While each case may be different, and although its entire structure cannot be fully captured by measurements, metrics, etc., these are no longer the days of "[f]or professional services rendered."[103] There is an acute and growing understanding within the market regarding the arbitrage opportunity that exists in intelligently assisting clients in reducing their legal spending. Several analytics companies are actively working to both aggregate large-scale datasets and leverage approaches from the world of procurement to identify value propositions throughout the legal service marketplace.[104]

No one is employing these tools better than sophisticated general counsels in their purchasing of legal services for their respective law divisions. "Bob, it looks like we have a potential employment discrimination case coming out of our Phoenix regional office. I do not want to get soaked on the bill here. Let's find out how much a law firm partner with this specialization, in this geographic market, with say fifteen years of experience might cost per hour." Platforms, dashboards, and other management platforms designed to solve many of the information deficits around these and other related questions are actively being developed by a variety of entrepreneurial entities. The goal is simple: figure out how to intelligently reduce both their outside and inside legal spending.

Consider the rapidly growing legal analytics company, TyMetrix (a division of Wolters Kluwer).[105] TyMetrix builds information technology (IT) systems that are designed to "improve the performance of internal operations, and [provide] data solutions that give legal professionals an information

[103] *See* James B. Stewart, *Dewey's Fall Underscores Law Firms' New Reality*, N.Y. TIMES, May 5, 2012, at B1.

[104] *See generally* BUYING LEGAL: PROCUREMENT INSIGHTS AND PRACTICE (Silvia Hodges ed., 2012) (discussing the sourcing of legal services).

[105] TYMETRIX, tymetrix.com (last visited May 10, 2013). TyMetrix is not the only company working in this emerging space. Two other major companies are DataCert and Sky Analytics. DATACERT, http://www.datacert.com (last visited May 10, 2013); SKY ANALYTICS, http://www.skyanalytics.com/ (last visited May 10, 2013).

advantage in any scenario."[106] Included among its product offerings is a legal analytics platform that delivers industry-wide legal spending and performance data that can be used by clients to determine an acceptable rate to pay for a given legal service.[107] To develop this immense data apparatus, TyMetrix leveraged its existing role as provider of backend billing and payment software to various law departments.

Understanding large-scale data aggregated across multiple clients was the key to garnering some deep insights, TyMetrix convinced its respective clients to pool and aggregate anonymous billing information for purposes of better understanding the contours of the respective legal marketplace.[108] Using this and other associated metadata, TyMetrix has published the *Real Rate Report*, a report highlighting trends and insights from billions of dollars in legal spending.[109] Among other things, the *Real Rate Report* features the actual dollar amounts spent by various purchasers of legal services (typically corporate law departments).[110]

Law firms and other legal service providers often offer "rack rate[s],"[111] a term developed in the travel industry to describe the often inflated prices that a person pays at a hotel if he or she deals directly with the hotel under high demand conditions.[112] The *Real Rate Report* is particularly useful because it highlights the actual rates paid by purchasers.[113] In much the manner that online travel sites (e.g., Orbitz, Travelocity, and Kayak) revolutionized the travel industry, this aggregated information can help high-end purchasers of legal services overcome various information deficits.[114] The information within the broader TyMetrix platform is extensive and includes more than $42 billion in legal spending, 398 million hours of legal services, 105 million activities

106 *About TyMetrix*, TYMETRIX, http://tymetrix.com/about-tymetrix/ (last visited May 10, 2013).

107 *Products*, TYMETRIX, http://tymetrix.com/products/legal-analytics/ (last visited May 10, 2013).

108 *See* Press Release, CT TyMetrix and the Corporate Executive Board Provide the Industry's First True Look at Legal Billing Rates and Trends (Sept. 7, 2010), *available at* http://tymetrix.com/press-releases/16/2010/showArticle/.

109 *See id.*

110 *Id.*

111 Debra Cassens Weiss, *Why Law Firms Are like Hotels: 'Rack Rates' Are Negotiable, Real Rates Vary by Client*, A.B.A. J. (May 26, 2010, 8:08 AM), http://www.abajournal.com/news/article/client_beware_law_firm_rack_rates_are_negotiable_and_real_rates_vary_even_f/.

112 Roger Collis, *Hotels: Never Pay the Rack Rates*, N.Y. TIMES (Mar. 20, 1992), http://www.nytimes.com/1992/03/20/style/20iht-freq_0.html.

113 *See* Press Release, *supra* note 108.

114 Bobbie Johnson, *The Great Online Travel Revolution*, GUARDIAN (Dec. 15, 2009, 9:07 PM), http://www.guardian.co.uk/travel/2009/dec/15/travel-websites-noughties-decade.

captured, 17,000 law firms and vendors, and 286,000 individual billers and time keepers.[115] Benchmarking, analyzing, and projecting future legal spending costs while also contesting existing legal bills is a significant portion of what the modern general counsel must do as he or she operates as the maestro of the company's global legal supply chain.[116]

With respect to the costs of legal services, it is hard to understate the amount of disruption this class of technology potentially introduced.[117] Is this lawyer really worth a $125 wage premium? Can we shift this matter over to a cheaper firm? Can we send this matter to the Raleigh office instead of the New York office? Once the purchasers of legal services start asking these types of questions, there is no retreat to the good old days of "[f]or professional services rendered."[118]

B. Staffing the Matter—Measuring Attorney Quality and Performance

Every client wants to pay less for its respective legal services.[119] Yet, year after year, law divisions, wealthy individual clients, and retail consumers continue to expend significant sums to vindicate their rights and protect their respective interests. As described above, some of the surplus collected by lawyers is attributable to the information deficit surrounding lawyer and law firm prices. In addition, the information environment surrounding the market

---

[115] *TyMetrix Legal Analytics*, TYMETRIX, http://tymetrix.com/products/legal-analytics/13/legalview (last visited Apr. 11, 2013).

[116] *See* MARI SAKO, GENERAL COUNSEL WITH POWER? (2011), *available at* http://www.sbs.ox.ac.uk/centres/professionalservices/Documents/Sako%20GC%20with%20Power%20Aug%202011.pdf; Milton C. Regan, Jr. & Palmer T. Heenan, *Supply Chains and Porous Boundaries: The Disaggregation of Legal Services*, 78 FORDHAM L. REV. 2137, 2167 (2010) ("Continuing and perhaps increasing use of networks by legal departments means that corporate counsel may begin to function more as general contractors who coordinate activities among a multitude of suppliers that make contributions at various points in the legal services value chain. If so, project management skills will become more important for such lawyers, as will the ability to structure governance arrangements that align incentives as much as possible among network members. Departments may also turn more to nonlawyers with such skills, much as many have come to rely on corporate procurement officers in negotiating the terms of law firm engagements."); *see also General Counsel Eyeing Legal Services "Production Line," Oxford Research Finds*, LEGALFUTURES (Sept. 7, 2011), http://www.legalfutures.co.uk/legal-services-act/market-monitor/general-counsel-eyeing-legal-services-production-line-oxford-research-finds (discussing Mari Sako's work).

[117] This is the general wisdom provided that it is not a "bet the company" case. If it is a "bet the company" case, then cost is generally less of a consideration. However, the vast majority of disputes are not likely to lead to the demise of the company.

[118] *See* Stewart, *supra* note 103 (highlighting the prior days when law firms simply submitted aggregated bills with the simple statement, "[f]or professional services rendered").

[119] In some rare instances, the selection of a lawyer is driven by noneconomic considerations, but in general it is the case that consumers would be happy to pay less for an otherwise equivalent legal solution.

for lawyers features significant noise attached to assessing both attorney quality and attorney performance. While both clients and law firms have an interest in assessing various aspects of lawyer quality and performance, these are among the most challenging measurement questions. Although it is often declared that a given attorney is "really good," typically the model underlying this assessment is not fully specified. In other words, what gives rise to the idea that a particular lawyer is great, average, or below average?

At the purchasing level, reputational bonding historically allowed law firms to develop brands that helped partially overcome the information deficit in the broader market.[120] It is difficult to directly assess quality, but clients understood that certain law firm brands were believed to be high quality.[121] Even in instances where the client was a general counsel, price was not typically a matter up for consideration.[122] If you wanted Cravath to be your lawyer, then as a client you needed to pay top dollar and not question the bill.[123] Quality was maintained and enforced through "mentoring, screening, and monitoring."[124] This particular economic structure was somewhat unstable, as individual lawyers who held positions as overseers had strong incentives to defect on mentoring, screening, and monitoring, and in the extreme case high performers had an incentive to take their book of business and strike out on their own.[125] As general counsels and other clients increased their sophistication and needed to figure out how to reduce their legal costs, those who formerly did not question the bill or the status of their firm(s) as the preferred vendor began to search for potential alternatives.[126] But the question of cost to value still lingers, and this tradeoff has stymied the accelerated move across the spectrum of bespoke service to commoditized legal services and legal information products.[127]

120 *See* Ribstein, *supra* note 1, at 753.

121 *See id.* at 754.

122 *Id.*

123 Stewart, *supra* note 103.

124 *See* Ribstein, *supra* note 1, at 754. The temptation to slack on these functions is strong because they typically require strong managerial cultures and long-time horizons. *See id.* at 754–55 ("In order for large law firms to perform their reputational bonding function they must motivate their lawyers to provide the mentoring, screening, and monitoring that supports the firm's reputation. The problem is that lawyers constantly must allocate time and effort between building the firm's reputation and building their own clienteles. If the ties binding lawyers to firms unravel, lawyers' temptation to build their personal human capital and client relationships may outweigh their incentive to invest in building the firm.").

125 *Id.* at 754–55.

126 *See generally* BUYING LEGAL: PROCUREMENT INSIGHTS AND PRACTICE, *supra* note 104.

127 *See* SUSSKIND, *supra* note 16.

It is difficult for clients to assess the quality of their lawyers. The question falls into two different, yet related, questions: How good is my attorney as a general matter? How well has my attorney performed on this specific case? Assessments for these questions are among the most challenging matters in our industry. Clients typically use proxies for attorney quality and performance because direct measurement is so challenging. One of those proxies is firm brand, and at the individual lawyer level such proxies include law school attended, clerkship, years of experience, cases handled, notoriety, etc. In transitioning from the mental model to a quantitative approach, obviously some of these parameters are easier to operationalize than others. However, one could imagine a range of plausible implementations and measurements that could be undertaken. Even more difficult is to develop measures to test or validate any particular model one might develop. It is an open question in need of a solution, and one should expect to see various entrepreneurial entities attempting to enter this space.

While the client-facing aspect of the question is understandable but challenging, the quality and assessment questions are equally present for law divisions and law firms. Providers themselves have a strong incentive to assess attorney performance in a manner not limited to their end service or product. These entities devoted to delivering legal services must assess whether their current employees are worthy of retention or promotion. In addition, for entry-level employees such as law firm associates, the hiring question is considered in an environment where increasingly sophisticated clients have a very limited appetite for paying for first- and second-year associates.[128] Thus, whether a firm should still make a significant up-front investment in a particular entry-level employee is a very challenging one. The calculus becomes even more strained when, with some probability, young associates leave their firms before the firms' investments in training are recouped.[129] Entry-level lawyers are having great difficulty getting a start in the traditional legal industry because a hiring mistake by a law firm can be a particularly costly one.[130] The selection

[128] *See* David Segal, *What They Don't Teach Law Students: Lawyering*, N.Y. TIMES, Nov. 20, 2011, at A1.

[129] Sometimes this is okay as direct revenues because revenues are, of course, not the only manner for a firm to recapture its investment. For example, former associates can become future clients if they become in-house counsel. The investment can be recaptured if that in-house counsel is persuaded to drive business in the firm's direction.

[130] *See* Joe Palazzolo, *First-Year Associates: Are They Worth It?*, WALL ST. J.L. BLOG (Oct. 17, 2011, 9:59 AM), http://blogs.wsj.com/law/2011/10/17/first-year-associates-are-they-worth-it/. It is particularly costly because many general counsels have imposed limits and bans on junior associates working on their respective matters. *Id.* ("More than 20% of the 366 in-house legal departments that responded are refusing to pay for the

of would-be associates has historically been far from an exact exercise. Indeed, it is a prediction problem in need of a data-driven solution.

Consistent with the *Moneyball*[131] ethos and the entrepreneurial spirit of the newly emerging legal services and product market, Lawyer Metrics is a company devoted to developing data-driven and scientifically informed forecasting models that predict the future success of individual lawyers (particularly at or near the entry level) in law firms and other related legal enterprises.[132] Its approach is designed to de-bias[133] both the hiring decision and the subsequent employee evaluation process. In other words, the value proposition offered by Lawyer Metrics and other related companies is linked to law firm efficiency and the "huge gains to be made by focusing on traits or attributes that are actually correlated with performance."[134] The company relies upon a battery of well-designed assessment tools:

> [These tools explore the] association[s] between performance and several dozen success traits that can be observed on a lawyer's resume or transcript. These range from traditional success criteria such as grades, law review, clerkships, and law school rank to nontraditional criteria that many firms overlook or give less weight to—blue- or pink-collar work experience, advanced degrees, publications, participation in team sports, etc. Using this wide range of biographical data, [its] Moneyball analyses reveal that law firms are often systematically overvaluing some attributes, ignoring others that really matter, and generally making bad tradeoffs in both entry level and lateral lawyer "drafts."[135]

work of first- or second-year attorneys, in at least some matters. Almost half of the companies, which have annual revenues ranging from $25 million or less to more than $4 billion, said they put those policies in place during the past two years, and the trend appears to be growing."). Thus, these entry-level associates are taking a fairly large salary and at the same time are not able to staff many of a given firms' matters.

131 *See generally* MICHAEL LEWIS, MONEYBALL: THE ART OF WINNING AN UNFAIR GAME (2004).

132 *See* LAWYER METRICS, http://www.lawyermetrics.com/home.html (last visited May 10, 2013).

133 *See* Steve Gibson et al., *Moneyball for Law Firms*, AMLAW DAILY (Oct. 10, 2011, 4:00 PM), http://amlawdaily.typepad.com/amlawdaily/2011/10/moneyball-for-law-firms.html ("Bias among brilliant equity partners? Yes, it happens. A good example is attitudes toward law school pedigree. The data suggests that, in several firms, a subset of partners who attended elite law schools often give higher performance ratings to associates who also attended elite law schools—even when nonelite associates are statistically identical on every other measure. In contrast, when looking at the same group of associates, partners who did not attend elite law schools observe no performance gap.").

134 *See id.*

135 *See id.*

By identifying this obvious weakness in the labor market[136] and fashioning a solution, Lawyer Metrics has developed what should be a very profitable niche in the newly emerging legal data analytics space. Notwithstanding the substantial progress that has been made to date, much more can still be accomplished, as evaluating lawyer quality and performance is key to the sort of commoditization discussed by scholars such as Richard Susskind.[137] For example, using easily available inputs, how can I evaluate the actual work product generated by lawyers? Is the quality of my lawyer output improving or regressing? What is the match between the complexity of my work and the necessary level of sophistication required of my lawyer? Obviously, these are just a few questions that one could pose.

C. Predicting Case Outcomes

Do I have a case? How many zeros worth of exposure are we likely facing here? In addition to the question of cost, the prediction of case outcomes is among the top questions of interest to a potential client. Legal prediction is a long-standing idea that can be traced back to some of our foremost legal thinkers. More a concept than a technical reality, prediction was a centerpiece of Justice Oliver Wendell Holmes's conception of jurisprudence.[138] In addition, the related question of legal uncertainty is one that has been considered by many scholars applying a wide-ranging set of approaches including law and economics as well as legal philosophy. Despite all of the work describing the problems surrounding prediction and uncertainty, until very recently there was an overall dearth of active technical research in the space. The rise of Big Data and soft artificial intelligence, however, has invigorated the formerly dormant field of legal prediction.

1. Predicting Judicial Decisions—#PredictSCOTUS

Reading the tea leaves and predicting its decisions is a bit of a sport for the sophisticated observers of the United States Supreme Court. Every year, law reviews, magazine and newspaper articles, television and radio time, conference panels, blog posts, and tweets are devoted to questions such as: How will Justice *X* side on this particular matter? In these and other related

---

[136] It is only obvious now. As the saying goes, most innovation lives at the intersection of "seems like a bad idea" and "good idea."

[137] *See, e.g.*, SUSSKIND, *supra* note 16.

[138] *See* Oliver Wendell Holmes, Jr., *The Path of the Law*, 10 HARV. L. REV. 457 (1897).

forums, individual commentators offer all sorts of theories about what the Court will do and why it will choose to do so.[139]

Suffice to say, as a matter of scientific forecasting, the quality of many of these theories is very unclear. Without some sort of a validation scheme, it is fairly difficult to determine whether a forecast is well-specified or little more than informed speculation. This is all to say that most theories that have been offered in traditional legal scholarship are not tested in a manner that could demonstrate their basic validity as either an explanatory or a predictive model (emphasis on the latter).

Now it is worth noting that during the years when doctrinal approaches reigned supreme, there were scientific approaches being applied to these questions. Tracing back to the early work of the political scientist Harold Spaeth as well as others, there existed a long-standing tradition of empirically analyzing the decisions of the Court.[140] Using regression analysis and other related techniques,[141] the existing social science work provided significant insight into the case and political factors that helped drive the Court's decision making.[142] Despite all of this quality scholarship, until fairly recently there was very little in the way of a forward-facing predictive science in either quantitative social science or empirical legal studies.

As described and popularized in *Super Crunchers* by Ian Ayres,[143] the 2002 Supreme Court Forecasting Project represented an important break with

[139] Of course, prediction is hardly the only goal of the enterprise. For example, commentators are interested in outlining policy concerns, flagging failed attempts to harmonize doctrines, raising emerging legal issues in society, and highlighting ongoing disputes between lower courts.

[140] *See, e.g.*, JEFFREY A. SEGAL & HAROLD J. SPAETH, THE SUPREME COURT AND THE ATTITUDINAL MODEL (1993).

[141] Of course, much of this work applied first-generation social science statistical methods such as OLS (ordinary least squares) and later categorical dependent models (probit and logit). This includes efforts at prediction. *See, e.g.*, Jeffrey A. Segal, *Predicting Supreme Court Cases Probabilistically: The Search and Seizure Cases, 1962–1981*, 78 AM. POL. SCI. REV. 891 (1984).

[142] *See, e.g.*, LEE EPSTEIN & JACK KNIGHT, THE CHOICES JUSTICES MAKE (1998); Forrest Maltzman & Paul J. Wahlbeck, *May It Please the Chief? Opinion Assignments in the Rehnquist Court*, 40 AM. J. POL. SCI. 421, 425–26 (1996); Jan Palmer, *An Econometric Analysis of the U.S. Supreme Court's Certiorari Decisions*, 39 PUB. CHOICE 387 (1982); Jeffrey A. Segal & Albert D. Cover, *Ideological Values and the Votes of U.S. Supreme Court Justices*, 83 AM. POL. SCI. REV. 557 (1989); Donald R. Songer & Stefanie A. Lindquist, *Not the Whole Story: The Impact of Justices' Values on Supreme Court Decision Making*, 40 AM. J. POL. SCI. 1049 (1996); James F. Spriggs II et al., *Bargaining on the U.S. Supreme Court: Justices' Responses to Majority Opinion Drafts*, 61 J. POL. 485 (1999); Paul J. Wahlbeck et al., *Marshalling the Court*: *Bargaining and Accommodation on the United States Supreme Court*, 42 AM. J. POL. SCI. 294 (1998).

[143] IAN AYRES, SUPER CRUNCHERS (2007).

traditional social science research on the Supreme Court: "Rather than focus retrospectively, and proceed to analyze, critique, quantify, regress, debunk, reconcile, classify, or applaud some set of the Court's past decisions, we instead applied two different methods to predict the outcome of every case argued in the Term."[144] Building somewhat off their prior technical work,[145] political scientists Andrew Martin and Kevin Quinn, together with legal scholars Theodore Ruger and Pauline Kim, set up a tournament for the 2002–2003 Supreme Court Term.[146] Like any Term of the Court, the issues presented were wide ranging and included controversial topics such as *Miranda* rights, affirmative action, state sovereign immunity, the First Amendment, sex offender registration, and three strikes laws.[147]

The tournament pitted a classification tree (Method #1) against the predictions of elite lawyers and law professors (Method #2) with the straightforward task of determining the votes of individual United States Supreme Court Justices (affirm or reverse) in upcoming cases.[148] Through the tournament, the goal was to observe both the relative and absolute performance of individuals and the machines.[149] For many, the results were surprising: "[T]he machine did significantly better at predicting outcomes than did the experts. While the *experts correctly forecast outcomes in 59.1% of cases, the machine got a full 75% right*."[150]

The pattern is familiar—machines outperforming humans on a task that typically requires expert judgment (and this was about ten years ago).[151] This is

144 Theodore W. Ruger et al., Essay, *The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting Supreme Court Decisionmaking*, 104 COLUM. L. REV. 1150, 1151 (2004).

145 *See* Andrew D. Martin & Kevin M. Quinn, *Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999*, 10 POL. ANALYSIS 134 (2002).

146 Ruger et al., *supra* note 144, at 1160.

147 *Id.* at 1151.

148 *Id.* at 1160.

149 *Id.*

150 *Id.* at 1152 (emphasis added).

151 More recent work has attempted to improve on both the performance and generalizability of the approach undertaken during the Supreme Court Forecasting Project of 2002–2003. Among these efforts, a recent approach treating Justice votes as blocks within a complex network seems particularly promising. *See* Roger Guimerà & Marta Sales-Pardo, *Justice Blocks and Predictability of U.S. Supreme Court Votes*, PLOS ONE (Nov. 2011), http://www.plosone.org/article/info:doi/10.1371/journal.pone.0027188. One alternative to the model-based approach is crowd-sourced prediction. By far the leading platform is FantasySCOTUS created by Josh Blackman. *See* FANTASYSCOTUS, http://www.fantasyscotus.net/ (last visited May 10, 2013); *see also* Josh Blackman et al., *FantasySCOTUS: Crowdsourcing a Prediction Market for the Supreme Court*, 10 NW. J. TECH. & INTELL. PROP. 125 (2012). One very sound and open question raised by Ian Ayres is whether a purely model-driven approach will outperform crowd-sourced prediction. For a brief comparison of

arguably a real social science triumph for the predictive enterprise. It points to the real potential for such tools in other related endeavors that are most closely aligned with demand in the extant legal services marketplace. Supreme Court prediction is interesting and exciting, but the market is less interested in predicting the roughly eighty Supreme Court cases per year and much more interested in predicting *outcomes* in more pedestrian cases and other circumstances. Namely, in virtually every domain of law, settlements and dismissals are far more likely than actual decisions by judges or juries.[152] And Supreme Court cases are exceedingly rare events. Lawyers and their clients often bargain in the shadow of the law, and thus their interest surrounds a different question: When will this case settle and for how much?

2. Predicting Case Outcomes—Predicting Patent Disputes

Fast forward just a few years and we observe a rise of various data analytics companies working in the prediction space. In the case outcome prediction space, LexMachina is probably the most mature company. An offshoot replication from the work of the Stanford IP Litigation Clearinghouse (IPLC), LexMachina is a private analytics company that was spun off in 2009.[153] Founded by law professor Mark Lemley, together with cofounders Joshua Walker and George Gregory, "the IPLC mapped every electronically available *patent litigation event* and outcome to bring openness and

the approaches, see Ian Ayres, *Prediction Markets vs. Super Crunching: Which Can Better Predict How Justice Kennedy Will Vote?*, FREAKONOMICS (Dec. 23, 2009, 3:03 PM), http://freakonomics.blogs.nytimes.com/2009/12/23/prediction-markets-vs-super-crunching-which-can-better-predict-how-justice-kennedy-will-vote/. There is almost certainly meaningful information provided by each approach, so the proper but nontrivial question is how to properly blend the respective data streams to outperform the results offered by using just one approach.

[152] *See* Theodore Eisenberg & Charlotte Lanvers, *What Is the Settlement Rate and Why Should We Care?*, 6 J. EMPIRICAL LEGAL STUD. 111, 112 (2009) ("Settlement dominates outcomes of civil litigation in the United States yet surprisingly little systematic knowledge exists about settlement rates. Casual conventional wisdom often has it that about 95 percent of cases settle."). While Eisenberg and Lanvers successfully challenge the prevailing wisdom of a 95% settlement rate, the basic proposition that many cases settle still remains intact. *See* Jason Scott Johnston & Joel Waldfogel, *Does Repeat Play Elicit Cooperation? Evidence from Federal Civil Litigation*, 31 J. LEGAL STUD. 39, 40 (2002) ("[S]ettlement rates for some type[s] of cases—such as torts—exceed[] 90 percent."); Frank E.A. Sander, *The Obsession with Settlement Rates*, 11 NEGOTIATION J. 329, 331 (1995) ("[Ninety-five] percent of all cases filed in court are likely to settle eventually . . . [.]"); W. Kip Viscusi, *Product and Occupational Liability*, J. ECON. PERSP., Summer 1991, at 71, 84 ("95 percent of [fully pursued product liability claims] lead to a positive out-of-court settlement."). Eisenberg and Lanvers's numbers fluctuate by jurisdiction, but virtually all available evidence indicates that the rate of settlement in most practice areas is quite high. *See* Eisenberg & Lanvers, *supra*.

[153] *About*, LEX MACHINA, http://lexmachina.com/about/ (last visited May 10, 2013).

transparency to IP law."[154] Major technology companies such as Apple, Cisco, Genentech, Intel, Microsoft, and Oracle funded LexMachina's development of a massive and extensive dataset with more than 130,000 cases featuring in excess of 6,000,000 docket entries and direct access to more than 4 million documents.[155] Taken together, it represents the most extensive data platform for a given topical domain. LexMachina's board of advisors includes leading law professors, some of the most serious industry leaders in the Bay Area,[156] as well as intellectual heavyweights such as Andrew Ng, who teaches the massive online course Machine Learning for Coursera.[157]

Building useful technology is of course not the entire question. For any would-be technology startup, the question is not only whether the product can be built, but also whether the technology will be adopted by the relevant consumer market. Many amazing companies whose ideas were solid failed on this secondary question. It is often a question of timing and the appetite of the relevant market. One of the most famous examples of a timing failure is the legendary innovator and venture capitalist, Marc Andreessen. In 1999, Andreessen founded Loudcloud one of the very first (if not the first) cloud business services.[158] The idea seemed solid:

> [Y]ou should be able to buy all this software by the drink, instead of having to shell out for the bottle up front. By capitalizing on economies of scale, Loudcloud could provide higher levels of service than you could get in-house, and a startup could get its product to market almost instantaneously. It could spend its time and energy building the actual product instead of trying to figure out how to host it and keep it live.[159]

It was actually reasonably successful right up until the burst of the technology bubble. After the NASDAQ crashed, the company narrowly escaped but was able to persist until 2007 when it was purchased by HP.[160] "In retrospect, [the company was] five or six years too early."[161]

---

154 *Id.* (emphasis added).

155 *Id.*; *Product*, LEX MACHINA, http://lexmachina.com/product/ (last visited May 10, 2013).

156 *Team*, LEX MACHINA, https://lexmachina.com/about/team/#board (last visited May 10, 2013).

157 *See Machine Learning*, COURSERA, https://www.coursera.org/course/ml (last visited May 10, 2013).

158 Chris Anderson, *The Man Who Makes the Future:* Wired *Icon Marc Andreessen*, WIRED (Apr. 24, 2012, 7:35 PM), http://www.wired.com/business/2012/04/ff_andreessen/all/.

159 *See id.*

160 *Id.*

161 *Id.*

On the question of timing and market demand, patent litigation is arguably one of the more fertile grounds for case prediction. The stakes can be huge and many of the relevant customers have tremendous resources at their disposal. Indeed, predicting the success of a patent in a thicket of competing claims is important not only for the respective inventor, but in many cases the probability of patent failure could represent the discount rate applied to a company's valuation. Thus, it is a key tool for venture capital firms and investment banks whose valuations of patents could likely benefit from the more complete data source as well as the more rigorous methods that can be applied with the massive scope of data that LexMachina possesses.

3. Predicting Case Outcomes—Securities Fraud Class Actions

While analytics for patent litigation is the most well-developed domain for case outcome prediction, several other areas have shown significant early promise.[162] For example, a recent paper by Blakeley McShane, Oliver Watson, Tom Baker, and Sean Griffith, published in the *Journal of Empirical Legal Studies*, articulates the first predictive model of securities fraud class action lawsuits.[163] It predicts both the likelihood of settlement and the expected settlement amount.[164] The model is fully predictive from the initial stages of litigation, as it uses only variables that are known at the day of filing.[165] Additionally, the model is able to flag the high exposure cases that are simultaneously fairly unlikely to settle, but will settle for a large amount if settlement occurs.[166]

Unlike most typical *empirical legal studies* papers, the authors first develop a model and then validate their model in two conceptually distinct manners: First, they tested their model on an out-of-sample prediction of the relevant

---

[162] One other area that has shown significant progress is the study of veil piercing and the conditions under which veil-piercing arguments will be entertained by courts. While Christina L. Boyd and David A. Hoffman do not specifically develop a prediction model, their combined research efforts represent a significant move in this direction. *See* Christina L. Boyd & David A. Hoffman, *Disputing Limited Liability*, 104 NW. U. L. REV. 853, 856 (2010). For related work, see Christina L. Boyd & David A. Hoffman, *Litigating Toward Settlement*, 30 J.L. ECON. & ORG. (forthcoming 2014), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1649643; Christina L. Boyd et al., *Building a Taxonomy of Litigation: Clusters of Causes of Action in Federal Complaints*, 10 J. EMPIRICAL LEGAL STUD. (forthcoming 2013), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2045733.

[163] Blakeley B. McShane et al., *Predicting Securities Fraud Settlements and Amounts: A Hierarchical Bayesian Model of Federal Securities Class Action Lawsuits*, 9 J. EMPIRICAL LEGAL STUD. 482 (2012).

[164] *See id.* at 484.

[165] *See id.*

[166] *See id.*

dataset, and second, they engaged in a form of forward prediction using cases from the end of their dataset.[167] Specifically, the authors note:

> [W]e held out a *random 25 percent of our observations*, thus leaving 899 cases as in-sample and 299 as out-of-sample. All 899 observations were used to fit the settlement/dismissal model, whereas the 592 cases that settled were used to fit the settlement amount model. . . .
>
> . . . .
>
> . . . .
>
> [W]e also performed a more difficult out-of-sample evaluation. In particular, we held out the 286 cases that were filed in either 2003 or 2004 (i.e., the last two years of our data; these cases account for 24 percent of the data). . . . Out-of-sample results under this more difficult hold-out schema remained strong. In particular, the diagnostic plots and fit statistics for this hold-out schema differed minimally . . . .[168]

Both of these validation steps outlined by the authors are critical to demonstrating that their model is robust and does not "overfit" the respective data.[169] As the field moves forward into greater use of prediction models, it is critical for these validation efforts to be undertaken and demanded prior to their actual deployment in any real world application.[170]

4. Predicting Relevant Documents—Electronic Discovery & Predictive Coding

Expense and uncertainty often surrounds the decision to litigate. That decision is guided by considerations of the total cost of litigation and the likelihood of ultimate success on the merits. The total cost of litigation is driven by a number of factors: lawyers, expert witnesses, investigators,

---

167 *Id.* at 505–06.

168 *Id.* at 504–06 (emphasis added).

169 Overfitting is a serious problem in data mining and machine learning. It originates from a model that is too complex or one that mistakes noise for signal in particular application. *See Overfitting*, WIKIPEDIA, http://en.wikipedia.org/wiki/Overfitting (last updated Feb. 26, 2013, 4:25 PM); *see also* Andrew Y. Ng, *Preventing "Overfitting" of Cross-Validation Data*, *in* PROCEEDINGS OF THE FOURTEENTH INTERNATIONAL CONFERENCE ON MACHINE LEARNING 245 (1997).

170 Beyond this and other efforts described herein, much of the case-prediction space remains open for innovation, entrepreneurship, and technology.

employee time and distraction, and to an ever-increasing extent the costs of discovery.[171]

Today's discovery is electronic discovery (e-discovery). In the current business environment, paper is now the exception. Even a decade ago, 93% of all information was electronic and the percentage is almost certainly even higher today.[172] The decrease in data storage cost and increase in processor speed has brought with it a massive proliferation of electronically stored information (ESI), including information on work computers, personal computers, e-mail, removable media (i.e., flash drives and portable hard drives), corporate intranets, mobile devices, file servers, backup systems, computerized voicemail, etc.[173] By one estimate, in large corporations and other equally large institutions, an average of 45%–50% of civil litigation respondents' costs are attributable to discovery.[174]

E-discovery is so expensive in part because the proliferation of ESI has made the review process so expansive. Under Federal Rule of Civil Procedure 26, "[p]arties may obtain discovery regarding any nonprivileged matter that is relevant to any party's claim or defense."[175] Relevance is judged by the process one undertakes insomuch as "[r]elevant information need not be admissible at the trial *if the discovery appears reasonably calculated* to lead to the discovery of admissible evidence."[176] Litigants and their lawyers are obligated to produce all relevant ESI unless the producing party obtains a limiting order[177] or the information is not reasonably accessible.[178] Otherwise, the party must produce

171 *See generally* NICHOLAS M. PACE & LAURA ZAKARAS, RAND INST. FOR CIVIL JUSTICE, WHERE THE MONEY GOES: UNDERSTANDING LITIGANT EXPENDITURES FOR PRODUCING ELECTRONIC DISCOVERY (2012), *available at* http://www.rand.org/content/dam/rand/pubs/monographs/2012/RAND_MG1208.pdf.

172 *See* MICHAEL R. ARKFELD, PROLIFERATION OF "ELECTRONICALLY STORED INFORMATION" (ESI) AND REIMBURSABLE PRIVATE CLOUD COMPUTING COSTS 4 (2011), *available at* http://www.lexisnexis.com/documents/pdf/20110721073226_large.pdf.

173 *See id.*

174 *See* NAVIGANT CONSULTING, THE STATE OF DISCOVERY ABUSE IN CIVIL LITIGATION: A SURVEY OF CHIEF LEGAL OFFICERS 8 (2008) (surveying Fortune 1000 chief legal officers to show that, on average in 2007, 45%–50% of corporations' civil litigation costs related to discovery activities). A "significant share" of those costs are attributable to electronically stored information. *Id.*

175 FED. R. CIV. P. 26(b)(1); *see also* Lee H. Rosenthal, *A Few Thoughts on Electronic Discovery After December 1, 2006*, 116 YALE L.J. POCKET PART 167, 171 n.4 (2006), http://www.yalelawjournal.org/the-yale-law-journal-pocket-part/procedure/a-few-thoughts-on-electronic-discovery-after-december-1,-2006/.

176 FED. R. CIV. P. 26(b)(1) (emphasis added).

177 FED. R. CIV. P. 26(b)(2)(B).

178 *Id.*; *see also* Rosenthal, *supra* note 175, at 171 ("Rule 26(b)(2) applies a two-tier structure to this distinctive and recurring problem of electronic discovery. The first tier is party-managed discovery; the second tier is available only on court order and under court supervision. A party must provide discovery of the first

the information, and in order to do so they must wade through the sea of ESI.[179]

How does a party find those relevant nonprivileged documents or records? In the "golden days" of document review, the days prior to the proliferation of electronically stored information, law firms would execute manual review of paper documents using teams of young associates.[180] This was a major profit center for law firm partners (particularly for those in "Big Law"), and it served as an entry point for young associates in the profession.[181] The economics were quite simple: While slowly learning the intricacies of practice, young associates could help offset the cost of their salaries and benefits by executing a variety of otherwise mundane tasks such as document review.[182]

Nothing is more responsible for undercutting that particular economic ordering than the ubiquitous use of e-mail. Each day roughly 144 billion e-mails are sent from the roughly 2 billion email accounts worldwide.[183] The vast majority of employees in most professional environments have an e-mail account, and taken together those accounts contain massive volumes of information. Depending upon their precise data retention policy, it is not uncommon for a large organization to have millions of e-mails stored on their respective servers. The prospect of executing an exhaustive manual review of this amount of ESI is entirely implausible, thereby necessitating lawyers and their sophisticated clients to seek alternative, technologically infused approaches to review and produce otherwise relevant information.

Now it is certainly the case that law firms—and their clients—have not been uniformly innovative in response to the new world of e-discovery. Indeed, rather than innovate and capture this work, law firms have witnessed the rise of companies such as IBM, Symantec, EMC, Recommind, Clustify, Clearwell Systems, Autonomy, FTI Technology, kCura, and many others. One

tier—relevant, reasonably accessible, electronically stored information—without a court order. A party need not review or provide discovery of electronically stored information that it identifies as 'not reasonably accessible.' Information contained on such sources is in the second tier, subject to discovery if the requesting party can show good cause for a court to order production." (footnote omitted)).

179 *See* FED. R. CIV. P. 26(b)(1).

180 Markoff, *supra* note 12, at A1.

181 *Id.*

182 *Id.*

183 *See, e.g.*, THE RADICATI GRP., INC., EMAIL MARKET, 2012–2016, at 2, 5 (Sara Radicati ed., 2012), *available at* http://www.radicati.com/wp/wp-content/uploads/2012/08/Email-Market-2012-2016-Executive-Summary.pdf.

recent estimate pegs total e-discovery revenues at $1.5 billion for 2013 with significant growth expected throughout the balance of this decade.[184]

Taken on the whole, e-discovery represents perhaps the most mature incursion of technology into the practice of law. Like many other subsectors of the industry, even this otherwise technically infused domain is about to be transformed through quantitative prediction-based technology. Namely, while the first generation of e-discovery was focused upon platforms for collection, processing, search, and review, the costs associated with e-discovery did not come down. Rather, the cost of production has actually increased.[185] This is in part because the cost of review still represents more than 70% of the total cost of e-discovery because review is still primarily driven by human labor.[186]

We now stand on the cusp of the next generation of e-discovery centered around "predictive coding"[187] technology, which should reduce costs to clients[188] and in turn increase profits to high-performing law firms and legal product companies engaged in the enterprise.

Predictive coding, or more generally, "technology aided review," seeks to reduce the extent of human involvement in the e-discovery process. Predictive coding "is a function, not a specific technology; so the technical methods, process, and workflow behind different vendors' underlying search and text mining may vary."[189]

---

[184] *See* Evan Koblentz, *E-Discovery Market Predicted to Reach $1.5B in 2013*, LAW TECH. NEWS (May 23, 2011), http://www.law.com/jsp/lawtechnologynews/PubArticleLTN.jsp?germane=1202555971820&id=1202494788349.

[185] *See* PACE & ZAKARAS, *supra* note 171, at xiii.

[186] *Id.* at xiv.

[187] *See, e.g.*, Kramer Levin Naftalis & Frankel LLP, *Courts Continue to Address Technology-Assisted Review and "Predictive Coding,"* ELECTRONIC DISCOVERY UPDATE, Fall 2012, at 1, *available at* http://www.kramerlevin.com/files/Publication/f067d2fb-2658-4f82-87ec-1940afd32bf2/Presentation/PublicationAttachment/f4b2ea5b-0a2f-4e55-9f5c-19f0c33389a5/2012_0901_Electronic%20Discovery%20Update_Newsletter.pdf; David Hill, *Big Data's Evolving Role in E-Discovery: What Is Predictive Coding?*, NETWORK COMPUTING (Aug. 17, 2012), http://www.networkcomputing.com/e-discovery/big-datas-evolving-role-in-e-discovery-w/240005739.

[188] It will certainly drive down the costs per terabyte. The open and unclear question is whether the reduction in costs per unit will exceed the increases in ESI that are expected in the coming years.

[189] KATEY WOOD & BRIAN BABINEAU, PREDICTIVE CODING: THE NEXT PHASE OF ELECTRONIC DISCOVERY PROCESS AUTOMATION 5 (2011), *available at* http://www.recommind.com/sites/default/files/ESG_WP_Recommind_Predictive_Coding_2011.pdf. Although it has the potential to be highly transformative, predictive coding in e-discovery represents nothing more than a slightly repackaged applied case of classification methods for text and other metadata that have existed in other academic and industrial sectors for quite some time. This is in part why the patent issued to the software company Recommind is highly questionable (at best). There are a number of potential avenues for challenging its patent, including

One important distinction that transfers from the machine learning realm to the predictive coding e-discovery world is the distinction between supervised and unsupervised methods. All approaches to predictive coding in the e-discovery space rely upon either semi-supervised or supervised learning approaches. Such approaches are inductive and typically involve the seeding of the algorithm with training (or labeled) data from which the machine infers the "true" function for assigning a document to a particular group (i.e., relevant versus not relevant). This inference is achieved using some sort of a cost function where the goal is to minimize that cost function while at the same time not overfitting the relevant data. The search for this function is iterative because the space of possible model configurations is searched and tested until satisfactory results are obtained consistently.

There exist a variety of approaches to achieve these ends, and the specific distinctions are best left for a more appropriate technical outlet. However, for a general sketch, just consider the sort of data that is associated with a basic record in electronic discovery. In general, this record features a mixture of its text (which usually reduces to its keywords and perhaps other semantic information) and its associated metadata (author, date, etc.). The task in predictive coding is to apply one (or more) of the set of supervised learning algorithms to classify each new record relative to the "gold standard data" that has been preidentified or preclassified by an expert reviewer. Such applicable methods include latent semantic analysis, naïve bayes classifiers, support vector machines, genetic algorithms, neural networks, etc.

Given all of the available methods, one might wonder how to select the appropriate approach between them. Specifically, it would appear that the ultimate "meta-method" would be a higher order algorithm that could preselect among some of the candidate approaches mentioned above. As the adage typically goes in the search-and-optimization community, there is no free

novelty and nonobviousness. *See* U.S. Patent No. 7,933,859 B1 (filed May 25, 2010); Christopher Danzig, *Predictive Coding Patented, E-Discovery World Gets Jealous*, ABOVE L. (June 9, 2011, 3:29 PM), http://abovethelaw.com/2011/06/predictive-coding-patented-e-discovery-world-gets-jealous/; Evan Koblentz, *Recommind Intends to Flex Predictive Coding Muscles*, LAW TECH. NEWS (June 8, 2011), http://www.law.com/jsp/lawtechnologynews/PubArticleLTN.jsp?id=1202496430795&slreturn=1&hbxlogin=1; Press Release, Recommind, Inc., Recommind Patents Predictive Coding (June 8, 2011), *available at* http://www.recommind.com/releases/recommind-patents-predictive-coding. For a basic thrust of the argument, see Devin Krugly, *Recommind's "Predictive Coding" Patent: More PR than IP*, EDISCOVERY INSIGHT (June 13, 2011), http://ediscoveryinsight.com/2011/06/recommind%E2%80%99s-%E2%80%9Cpredictivecoding%E2%80%9D-patent-more-pr-than-ip.

lunch.[190] It is not possible to develop this "meta-method" because each of the existing solution concepts has strengths and weaknesses that vary depending on the underlying problem. In other words, it is not possible to clearly identify, in advance, the optimal solution concept for a given problem. For any nontrivial problem, the notion of global optima is thus impossible to pre-evaluate. This does not leave the researcher or vendor without any recourse, but serves as an important cautionary limit.

In short, while the existing methods differ and a significant number of technical questions still remain unanswered, document review as well as e-discovery as we currently know it is about to be substantially reset. This demise is driven by yet another form of quantitative legal prediction: predictive coding.

---

[190] The "no free lunch" (NFL) theorem has received extensive treatment in the search-and-optimization community. Indeed, it is a key question for those involved in supervised learning. The NFL theorem demonstrates the futility of efforts to search for bias-free learning. For example, Wolpert and Macready explain:

> A number of 'no free lunch' (NFL) theorems are presented which establish that for any algorithm, any elevated performance over one class of problems is offset by performance over another class. These theorems result in a geometric interpretation of what it means for an algorithm to be well suited to an optimization problem.

David H. Wolpert & William G. Macready, *No Free Lunch Theorems for Optimization*, 1 IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION 67, 67 (1997); *see also* Cullen Schaffer, *A Conservation Law for Generalization Performance*, *in* PROCEEDINGS OF THE ELEVENTH INTERNATIONAL CONFERENCE ON MACHINE LEARNING 259 (1994). Building off earlier work, Droste, Jansen, and Wegener further outline:

> [A]n Almost No Free Lunch (ANFL) theorem shows that for each function which can be optimized efficiently by a search heuristic there can be constructed many related functions where the same heuristic is bad. As a consequence, search heuristics use some idea how to look for good points and can be successful only for functions "giving the right hints."

Stefan Droste et al., *Optimization with Randomized Search Heuristics—The (A)NFL Theorem, Realistic Scenarios, and Difficult Functions*, 287 THEORETICAL COMPUTER SCI. 131, 131 (2002). The impossibility of bias-free learning can actually be traced back to the canonical work of Hume who noted:

> [T]here is nothing in any object, consider'd in itself, which can afford us a reason for drawing a conclusion beyond it; and, [t]hat even after the observation of the frequent or constant conjunction of objects, we have no reason to draw any inference concerning any object beyond those of which we have had experience . . . .

DAVID HUME, A TREATISE OF HUMAN NATURE 139 (L.A. Selby-Bigge ed., Clarendon Press 1967) (1888) (emphasis omitted).

5. *Quantitative Finance Meets Quantitative Legal Prediction—Lessons for the Age of Data-Driven Law Practice*

The modern general counsel is called upon to be not only a legal supply-chain manager, but also a legal-risk portfolio manager. Both of these respective exercises can be substantially aided through the use of data, metrics, and models. Whether sourcing a particular legal matter, determining the outcome of a given piece of litigation, or forecasting the long-run implications of a given contract provision, the core questions involve matters of prediction. Given that we are likely heading into an age of data-driven law practice, an open question remains: Are current law students, lawyers, and general counsels well prepared to engage in this sort of new ordering? It is pretty clear that as a general matter the answer is no.

For a glimpse of the future, finance offers instructive lessons for the legal industry. Not long ago, the vast majority of trading activity was guided by individual brokers selecting stocks in direct consultation with individual clients.[191] Such human reasoners would typically leverage a mental model the reasoner developed through experience in the field. While the human element has not been completely removed from finance, the rise of the quants displaced many of the status quo practices.[192] The emphasis has shifted from human to machine judgment. Thus, on any given day, a majority of trades executed on the New York Stock Exchange are generated algorithmically.[193]

As it is a domain that involves sophisticated reasoning, finance offers important instructive lessons for lawyers and legal educators. *Finance has not gone away*, but the role of prediction within finance has undergone a radical transformation. Whether it is a human or a machine executing a prediction, the relevant standard is not perfection, but rather improvement over competing approaches. In finance, in a large number of instances, the machines have outperformed.[194] As such, the qualitative skills that were formerly privileged in

[191] For more on flash trading, see Roger Lowenstein, Op-Ed, *A Speed Limit for the Stock Market*, N.Y. TIMES, Oct. 2, 2012, at A31; Michael Mackenzie, *High-Frequency Trading Under Scrutiny*, FIN. TIMES (July 28, 2009, 6:44 PM), http://www.ft.com/intl/cms/s/0/d5fa0660-7b95-11de-9772-00144feabdc0.html.

[192] *High-Frequency Trading Prospers at Expense of Everyone*, BLOOMBERG (Dec. 25, 2012, 6:30 AM), http://www.bloomberg.com/news/2012-12-25/high-frequency-trading-prospers-at-expense-of-everyone.html.

[193] *See* Mackenzie, *supra* note 191 (noting that "high-frequency trading accounts for as much as 73 per cent of US daily equity volume").

[194] *High-Frequency Trading Prospers at Expense of Everyone*, *supra* note 192. That said, there are important instances where machines have performed far worse than humans.

finance are simply of diminished value after the advent of soft AI.[195] In some cases, new academic tracks, such as financial engineering, blossomed and helped place students in positions that were previously reserved for students with traditional MBA training.[196]

Of course, in light of the 2007–2009 financial crisis, it is easy to point to the shortcomings of finance and argue against these sorts of developments in the legal services market.[197] Whether the questions surround the financing of lawsuits or engaging in the types of predictions described above, it does not matter what you think ought to happen; it only matters what the relevant market will embrace. The market will (or already has) embraced this sort of technology and there is likely much more coming down the pipeline. Whether this happens within the physical borders of the United States or is done abroad is an open question for regulators, but that this will occur somewhere is inevitable.

III. THE SCIENCE AND LIMITS OF PREDICTION—A PRIMER FOR THE AGE OF QUANTITATIVE LEGAL PREDICTION

Quantitative legal prediction is of course an applied case of the broader science of prediction. In describing this important and growing segment of the legal services industry, it is worth highlighting some of the properties associated with the more general science because the leading concepts from this domain are highly relevant to the future of the legal industry. These include the theoretical orientation (inverse versus forward solutions), the various methods (feature selection and extraction, classification, clustering, similarity methods, etc.), and important limits associated with prediction models.

A. The Theoretical Orientation: Inverse Versus Forward Problems and Machine Learning Versus Causal Inference

In comparing the sort of "mental models" developed by human reasoners against competing algorithms, the question is simple: Can your model predict

[195] Obviously judgment is still a valuable skill for those interested in value investing and those who hold long positions. The point is that the short-to-intermediate arbitrage has gone the way of the machines.

[196] For a history, see Xiaozhuo Yang, *Financial Engineering Education Risk Management*, CHINESE ASS'N PROFS. SCI. & TECH. (Dec. 2005), www.capst.org/events/FinancialEngineeringOverview.pdf.

[197] As outlined *infra* in Part III.C, one should approach questions of prediction with the appropriate level of humility.

better than the leading existing approach? Whether the question is well posed or whether the causality is well understood is not particularly critical.[198] In other words, with relatively stable temporal dynamics, it is not always necessary to have a deep theory in order to generate a well-functioning prediction engine. This is a tremendously important point because it represents a significant departure from the traditional hypothesis-testing–falsification framework typically undertaken in many scientific inquiries.

There is an important conceptual difference between many of the approaches used in machine learning and those used in causal inference. When attempting to identify and measure the specific relationship between a series of potential causal variables and the *left-hand-side variable* of interest, the gold standard in medical, physical, and social sciences is often considered to be the randomized control trial (RCT).[199] If one is interested in cleanly identifying the impact of a new drug, a new chemical combination, or a new law or social policy, randomization in assignment and a reasonably large *N* is typically considered a good way to measure the causal relationship.[200]

While RCTs are the preferred approach, they are simply unavailable in a wide variety of applications. Indeed, much of the "credibility" revolution in social science has surrounded the development and application of statistical tools designed to approximate RCT-style conditions. These include instrumental variables,[201] regression discontinuity,[202] propensity matching,[203]

[198] Please do not misunderstand—understanding the nature of the causal relationship is not harmful. Indeed, it is helpful. The important point is that when attempting to build a prediction engine that improves over existing status quo approaches, disentangling the precise causal relationship is just not always necessary. Many popular methods in AI and machine learning, such as neural networks and genetic algorithms, are "black-box methods" that still serve the goals of the particular task without deep concerns about proper assignments of causality.

[199] *See, e.g.*, A K Akobeng, *Understanding Randomised Controlled Trials*, 90 ARCHIVES DISEASE CHILDHOOD 840 (2005). *But see* Nancy Cartwright, *Are RCTs the Gold Standard?*, 2 BIOSOCIETIES 11, 11–20 (2007); Ted J. Kaptchuk, Commentary, *The Double-Blind, Randomized, Placebo-Controlled Trial: Gold Standard or Golden Calf?*, 54 J. CLINICAL EPIDEMIOLOGY 541 (2001). It is worth noting that even under such ideal conditions potential confounds can disrupt the analysis. Most notably, unforeseen failures in randomization and efforts to generalize results beyond the tested group can limit the reach of a particular model. Heckman, for example, points "to the difficulty of generalizing from experimental to real-world settings, argu[ing] that randomization is not any sort of 'gold standard' of causal inference, but this is a minority position." *See* Andrew Gelman, Essay, *Causality and Statistical Learning*, 117 AM. J. SOC. 955, 956 (2011).

[200] Akobeng, *supra* note 199.

[201] *See, e.g.*, Joshua D. Angrist & Alan B. Krueger, *Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments*, J. ECON. PERSP., Fall 2001, at 69; Joshua D. Angrist et al., *Identification of Causal Effects Using Instrumental Variables*, 91 J. AM. STAT. ASS'N 444 (1996); James Heckman, *Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making*

and many others. Given its increasingly interdisciplinary nature, legal scholarship has also embraced the causal inference revolution.[204] Indeed, it is causal inference, and *not* the science of prediction, that is the current mainstream in social science and empirical legal studies.[205] In general, this is a welcome development because for certain legal–scientific questions causal inference approaches are the most sound manner in which to proceed. However, when it comes to prediction, the tools of causal inference are not necessarily all that useful. If anything, the obsession with causal inference has

Program Evaluations, 32 J. HUM. RESOURCES 441 (1997); James J. Heckman & Richard Robb, Jr., *Alternative Methods for Evaluating the Impact of Interventions: An Overview*, 30 J. ECONOMETRICS 239 (1985); James J. Heckman, *Econometric Causality*, 76 INT'L STAT. REV. 1, 3 (2008).

[202] *See, e.g.*, Jinyong Hahn et al., *Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design*, 69 ECONOMETRICA 201 (2001); Guido W. Imbens & Thomas Lemieux, *Regression Discontinuity Designs: A Guide to Practice*, 142 J. ECONOMETRICS 615 (2008); Brian A. Jacob & Lars Lefgren, *Remedial Education and Student Achievement: A Regression-Discontinuity Analysis*, 86 REV. ECON. & STAT. 226 (2004); David S. Lee, *Randomized Experiments from Non-Random Selection in U.S. House Elections*, 142 J. ECONOMETRICS 675 (2008); Miguel Urquiola & Eric Verhoogen, *Class-Size Caps, Sorting, and the Regression-Discontinuity Design*, 99 AM. ECON. REV. 179 (2009).

[203] *See, e.g.*, Jeffrey B. Bingenheimer et al., *Firearm Violence Exposure and Serious Violent Behavior*, 308 SCIENCE 1323 (2005); Jinyong Hahn, *On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects*, 66 ECONOMETRICA 315 (1998); David J. Harding, *Counterfactual Models of Neighborhood Effects: The Effect of Neighborhood Poverty on Dropping Out and Teenage Pregnancy*, 109 AM. J. SOC. 676 (2003); Suzanne O'Keefe, *Job Creation in California's Enterprise Zones: A Comparison Using a Propensity Score Matching Model*, 55 J. URB. ECON. 131 (2004); Paul R. Rosenbaum & Donald B. Rubin, *Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score*, 39 AM. STATISTICIAN 33 (1985); Paul R. Rosenbaum & Donald B. Rubin, *Reducing Bias in Observational Studies Using Subclassification on the Propensity Score*, 79 J. AM. STAT. ASS'N 516 (1984); William R. Shadish et al., *Can Nonrandomized Experiments Yield Accurate Answers? A Randomized Experiment Comparing Random and Nonrandom Assignments*, 103 J. AM. STAT. ASS'N 1334 (2008).

[204] There exist a vast number of recent articles in a wide number of substantive legal domains. For just a small slice, see for example, Christina L. Boyd et al., *Untangling the Causal Effects of Sex on Judging*, 54 AM. J. POL. SCI. 389 (2010); D. James Greiner, *Causal Inference in Civil Rights Litigation*, 122 HARV. L. REV. 533 (2008); Jonathan Klick & Thomas Stratmann, *The Effect of Abortion Legalization on Sexual Behavior: Evidence from Sexually Transmitted Diseases*, 32 J. LEGAL STUD. 407 (2003); Leandra Lederman & Warren B. Hrung, *Do Attorneys Do Their Clients Justice? An Empirical Study of Lawyers' Effects on Tax Court Litigation Outcomes*, 41 WAKE FOREST L. REV. 1235 (2006); Yair Listokin, *Does More Crime Mean More Prisoners? An Instrumental Variables Approach*, 46 J.L. & ECON. 181 (2003). For a more general treatment of the question, see Lee Epstein & Gary King, *The Rules of Inference*, 69 U. CHI. L. REV. 1 (2002); Daniel E. Ho & Donald B. Rubin, *Credible Causal Inference for Empirical Legal Studies*, 7 ANN. REV. L. & SOC. SCI. 17 (2011).

[205] In a recent review paper, Andrew Gelman highlighted one important distinction and described this distinction as "[f]orward causal inference" and "[r]everse causal inference." With respect to this distinction, it is reverse causal inference that is the heart of mainstream modern econometrics. *See* Gelman, *supra* note 199, at 955–56. On some spectrum between forward causal inference and reverse causal inference lies *forward prediction* where the goal is simply to develop the "best" predictive model up to time t and then try to predict the next interval (i.e., $t+1$).

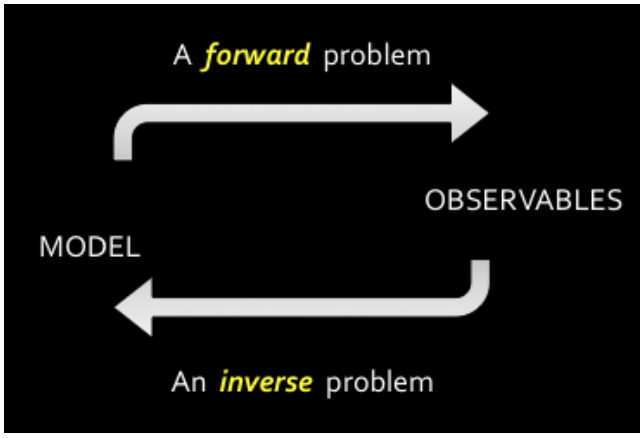
distracted many leading legal academics from what will transform the market for legal services—quantitative legal prediction.

Unlike causal inference, the science of prediction is driven by disciplines such as computer science, physics, and applied mathematics. Many of the most successful approaches in the science of prediction apply inverse (fully or partially inductive) style solution concepts that black box causality, and thus are quite different from the sort of experimental or quasi-experimental approaches undertaken by causal inference scholars.[206] While the labels differ across disciplines, the distinction between causal and predictive models is conceptually analogous to the distinction between a forward and an inverse problem. The study of inverse problems is a very active area of modern applied mathematics. The insights derived by such work are quite useful in the development of models in the age of quantitative legal prediction. Simply put, the inverse approach is the heart of machine learning. There is a wide variety of approaches but in general here is a common approach: Given this time series of data up to time *t*, which parameters and what weighting of those relevant parameters are most useful in predicting the next time step? Simply put, one uses the observables to build the model rather than using the model to assign causal weight to those observables.[207]

206 *See* Gelman, *supra* note 199.

207 *See* KEVIN P. MURPHY, MACHINE LEARNING: A PROBABILISTIC PERSPECTIVE (2012); *see also* ETHEM ALPAYDIN, INTRODUCTION TO MACHINE LEARNING (2d ed. 2010); MEHRYAR MOHRI, AFSHIN ROSTAMIZADEH & AMEET TALWALKAR, FOUNDATIONS OF MACHINE LEARNING (2012).

Figure 9: A Forward Versus Inverse Problem



While a hypothesis, in the traditional sense, is not employed[208] in developing a robust prediction engine, it is important to note that a serious scientific validation of approach is still employed. Validation for this class of inductive models is achieved using either "out of sample prediction" or "forward prediction."[209] Indeed, while forward testing is the ultimate test of a prediction model, there are still open questions associated with changing temporal dynamics. In other words, defining how much of the past is useful for predicting the future is a very challenging matter. As the dynamics of the system being modeled become more volatile, so too do the predictions of that system's behavior. The most refined algorithmic approaches must use some sort of learning rule in an effort to search the landscape of possible model configurations. If the dynamics are too volatile, even the most refined approach will struggle. The general goal, however, is to optimally update the model automatically as time ticks forward.[210]

[208] There is a hypothesis that is being tested, but it is conceptually distinct from the traditional approach seen in social science and empirical legal studies. For example, when developing a *classifier* in supervised learning, this form of data mining represents a form of hypothesis test. The goal of classification is to learn or inductively recover the boundary separating the instances of one class from the instances of all other classes. This involves testing multiple hypotheses regarding that exact boundary. While traditional regression analysis involves developing an analytical solution for the optimum of a lower order polynomial, machine learning typically requires a significant incursion into optimization theory. Specifically, evaluating all potential hypothesized boundary configurations is typically a problem that cannot be solved analytically. *See generally* ALPAYDIN, *supra* note 207.

[209] *See* MURPHY, *supra* note 207; *see also* ALPAYDIN, *supra* note 207; MOHRI ET AL., *supra* note 207.

[210] *See* MURPHY, *supra* note 207; *see also* ALPAYDIN, *supra* note 207; MOHRI ET AL., *supra* note 207.

B. A Perspective on the Applicable Methods: The Science of Similarity

People, music, and movies are objects that feature a lot of potential dimensions. Take music for example. A given song features a number of high-level elements such as composition, rhythm, ostinato, roots, tonality, instrumentation, stylings, recording techniques, influences, ensembles, individual instruments, lyrical content, vocals, and elements.[211] With all of these higher order properties and hundreds of lower level properties potentially associated with each song, it would appear to be impossible to algorithmically match songs with this many theoretical dimensions. Of course, this is quite possible, as the Music Genome Project and its popular associated technology, Pandora, are enjoyed by millions of end users.

Collaborative filtering technologies, such as those used by Amazon in recommending book purchases, are all "inverse" or inductive solutions to the respective problem.[212] Amazon does not have a deep theory of books. It simply wants to predict which books you are likely to purchase conditioned upon observing your purchases up to the present time *t*. The same is true of Facebook in recommending friends, Netflix in recommending movies, Pandora in recommending songs, and many others. Underneath the hood, these commercial products rely on some sort of concept of similarity that is implemented and refined using large bodies of data. Indeed, virtually all existing solutions embrace some concepts from the broader "science of similarity."

In broad strokes, when individuals engage in legal reasoning they engage in a high-level, high-dimensional search of the space of possible reference cases. In that search, similarity and dissimilarity are the drivers. Heuristics are used to define the stopping conditions. The science of legal search (legal information retrieval) is driven in substantial part by a notion of similarity. Humans do not—and cannot—exhaust the space and this is just one reason why humans + machines > humans or machines. Legal search intermediary companies such as Google, Lexis, and Westlaw aid lawyers by allowing them to make better sense of the sea of potentially relevant legal information. The problem with today's legal search is that the body of results is typically substantial, and thus the human (lawyer) must still engage in substantial filtering of the results.

---

[211] *See Music Genome Project*, WIKIPEDIA, http://en.wikipedia.org/wiki/Music_Genome_Project (last updated Feb. 27, 2013, 12:30 PM).

[212] *See* Greg Linden et al., *Amazon.com Recommendations*, IEEE INTERNET COMPUTING, Jan./Feb. 2003, at 76.

Much of the weight is put on the human reasoner to determine which cases are potentially useful or harmful to his particular position.

Whether using the case for argumentation purposes or for the purposes of prediction, it is really important to obtain similar cases to the underlying base case. This is actually fairly difficult because most cases share some level of similarity with other cases. The key research-and-development challenge is to develop a refined, but also scalable, method for defining similarity. Similarity measures are sometimes called distance measures, where each object is projected into an *n*-dimensional space and their respective distances are driven by some sort of a scoring function. Those similarity measures are composite scores of a variety of inputs. In the context of legal documents, those inputs include text (keywords, semantic information, etc.) and metadata (author, date, votes, citations, etc.). Leveraging this information in the appropriate mix is the "black magic" of algorithm development and is an important thrust of active technical research in the field of legal informatics.[213]

C. Analogical Reasoning: An Impossible Dream? It Starts with the Golden Nugget of Feedback Economy → Click Data

Analogical reasoning is at the core of how lawyers reason and how lawyers argue. The casebook method developed by Christopher Columbus Langdell is designed to tune the understanding of law students in an effort to help perfect their ability to reason by analogy.[214] Through immersion, students are bombarded with analogy after analogy in case after case. Much like the sort of inductive models discussed herein, successful students are able to harvest, retrieve, and induce the relevant method of common law reasoning and develop the sort of clever arguments that others find persuasive. They can execute this task despite not fully understanding the underlying model of persuasion. Although an individual human reasoner cannot precisely identify the rules, it is quite clear that some phrases and analogies captivate the

[213] *See, e.g.*, Kevin D. Ashley & Stefanie Brüninghaus, *Automatically Classifying Case Texts and Predicting Outcomes*, 17 ARTIFICIAL INTELLIGENCE & L. 125 (2009); Michael J. Bommarito II et al., *Distance Measures for Dynamic Citation Networks*, 389 PHYSICA A 4201 (2010); Jack G. Conrad, *E-Discovery Revisited: The Need for Artificial Intelligence Beyond Information Retrieval*, 18 ARTIFICIAL INTELLIGENCE & L. 321 (2010).

[214] *Casebook Method*, WIKIPEDIA, http://en.wikipedia.org/wiki/Casebook_method (last updated Feb. 20, 2013, 5:13 PM).

imagination.[215] Masterful persuasion and masterful legal argumentation have historically fallen in the "I know it when I see it" camp.

The interesting question is to what extent art can become a science, whereby a partial incursion into the domain of analogical reasoning can be undertaken. In other words, a key question for researchers in both artificial intelligence and law is precisely what leads individuals to, in fact, "know it when they see it."[216] What precisely is being triggered?

While useful work has been done in this basic thrust, we are still far away from a machine that can engage in "hard" analogical reasoning. The more immediate question is whether it is possible to develop some sort of second-best or "soft" analogical reasoning technology designed to aid human reasoners in their efforts to develop more persuasive arguments. On this front, there is significant hope, and it is likely that advances will be driven by an iterative mix of data + model + more data and so on. It will start with a more intelligent legal search and then will move up the intellectual value chain.

---

[215] We are still quite a distance from a fully developed approach, but some early work in this vein has been undertaken. *See, e.g.*, Katie Greenwood et al., Towards a Computational Account of Persuasion in Law (2003) (unpublished manuscript), *available at* http://cgi.csc.liv.ac.uk/~katie/icail03.pdf. For the more general exploration of analogy in law, see among others, MELVIN ARON EISENBERG, THE NATURE OF THE COMMON LAW (1988); Lawrence C. Becker, *Analogy in Legal Reasoning*, 83 ETHICS 248 (1973); Scott Brewer, *Exemplary Reasoning: Semantics, Pragmatics, and the Rational Force of Legal Argument by Analogy*, 109 HARV. L. REV. 923 (1996); Emily Sherwin, *A Defense of Analogical Reasoning in Law*, 66 U. CHI. L. REV. 1179 (1999); Cass R. Sunstein, Commentary, *On Analogical Reasoning*, 106 HARV. L. REV. 741 (1993); Richard A. Posner, *Reasoning by Analogy*, 91 CORNELL L. REV. 761 (2006) (reviewing LLOYD L. WEINREB, LEGAL REASON: THE USE OF ANALOGY IN LEGAL ARGUMENT (2005)); Joshua C. Teitelbaum, Analogical Legal Reasoning: Theory and Evidence (Sept. 1, 2012) (unpublished manuscript), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2145478; *see also* Daniel Martin Katz et al., *Legal N-Grams? A Simple Approach to Track the Evolution of Legal Language*, *in* LEGAL KNOWLEDGE AND INFORMATION SYSTEMS 167 (Katie M. Atkinson ed., 2011).

[216] There is a rich tradition of research in artificial intelligence and law including a peer-reviewed journal published by Springer and an international association with hundreds of members. *See, e.g.*, KEVIN D. ASHLEY, MODELING LEGAL ARGUMENT: REASONING WITH CASES AND HYPOTHETICALS (1990); ANNE VON DER LIETH GARDNER, AN ARTIFICIAL INTELLIGENCE APPROACH TO LEGAL REASONING (1987); Vincent Aleven, *Using Background Knowledge in Case-Based Legal Reasoning: A Computational Model and an Intelligent Learning Environment*, 150 ARTIFICIAL INTELLIGENCE 183 (2003); Layman E. Allen & C. Rudy Engholm, *Normalized Legal Drafting and the Query Method*, 29 J. LEGAL EDUC. 380 (1978); Katie Atkinson et al., *Computational Representation of Practical Argument*, 152 SYNTHESE 157 (2006); Kevin W. Saunders, *A Logic for the Analysis of Collateral Estoppel*, 12 RUTGERS COMPUTER & TECH. L.J. 99 (1986); Adam Wyner, *An Ontology in OWL for Legal Case-Based Reasoning*, 16 ARTIFICIAL INTELLIGENCE & L. 361 (2008); Edwina L. Rissland, Comment, *Artificial Intelligence and Law: Stepping Stones to a Model of Legal Reasoning*, 99 YALE L.J. 1957 (1990); Kevin D. Ashley, Ontological Requirements for Analogical, Teleological, and Hypothetical Legal Reasoning (2009) (unpublished manuscript), *available at* http://dl.acm.org/citation.cfm?id=1568236&bnc=1.

Consider three increasingly sophisticated forms of tools designed to guide a lawyer's legal reasoning: (1) People who cite *Case X* also cite *Case Y*; (2) Lawyers who argue *Principle X* also typically argue *Principle Y*; and (3) Given the mixture of argument and content in your brief, have you considered this argument and content, which is largely parallel (analogous) to your argument and content? On the spectrum of AI complexity, each of these models is increasingly sophisticated and would be hard to fully express and predict in a model alone. However, like many of the breakthroughs in commercialized artificial intelligence and in our feedback economy, it was not the model, but rather the click data, or the feedback effect, that ultimately led to a transformative product. The key to getting that feedback (click data) is to express a plausible-enough model to maintain end user participation during a "burn in" period.[217] This issue was faced in the move between the Music Genome Project (model) and Pandora (the commercialized product).

With the development of an elaborate and adaptive mapping of the case space, a number of potential technologies become possible. For example, to a greater degree of precision one should be able to predict the cases that lawyers should read in crafting their arguments (intelligent legal search), better predict relevant documents (e-discovery), and perhaps even suggest (predict) cases with analogous content or argumentation structure.

All of these search-based technologies could improve and streamline the core task of legal argumentation undertaken by many lawyers. In addition, such information retrieval is also needed for the highest end prediction engine. In order to develop a case prediction engine, one needs to be able to "pre-predict" the set of cases that are sufficiently similar to the base case to be indexed for purposes of executing the actual prediction of case outcomes. In order to deliver optimal results, the retrieval or "pre-predicted" set of comparison cases needs to include cases that share an analogical structure to the reference case. Since analogy is so powerful in law, its development cannot be ignored when it comes to the question of selecting the comparison group of cases. Modeling this sort of analogical reasoning is nontrivial, but it is not impossible. Thus, it should be one of the most pressing goals of research in the legal informatics and artificial-intelligence-and-law community.

[217] The "burn in" period is so critical because the click data can be used to back fill and refine the model. Over time, the model can forecast the proper mix of model and crowd-sourced prediction that best delivers the "answer" to the end user.

D. The Limits of Prediction: LaPlace's Mistake, Weather Versus Tides, and Law as a Complex Adaptive System

1. LaPlace's Demon & Prediction in a Complex Environment

Despite all technical possibilities, prediction is a difficult enterprise, and as such, one should confront the question with humility. Determinism is *not* the order of the day. While we already have entered the age of quantitative legal prediction, it is important to understand the limits of prediction as even some of the greatest minds in human history have fallen into the trap of overconfidence. Consider the work of the great French mathematician and astronomer, Pierre-Simon Laplace. In the most vigorous claim of deterministic thinking, Laplace argued:

> [If] at any given moment [one] knew all of the forces that animate nature and the mutual positions of the beings that compose it, if this intellect were vast enough to submit the data to analysis, could condense into a single formula the movement of the greatest bodies of the universe and that of the lightest atom; for such an intellect nothing could be uncertain and the future just like the past would be present before its eyes.[218]

This is the "Laplace Demon"—a claim that in the strong form argues the past completely determines the future.

As a matter of physics, such deterministic thinking has been discredited. For example, Werner Heisenberg (with the Heisenberg Uncertainty Principle), as well as many others, has demonstrated that for virtually all systems that feature more than just trivial dynamics there exists a maximum level of precision for which components can be measured and in turn predicted.[219] The maximum level of precision that is possible is a function of the quality and scope of measurement, the complexity of the interacting dynamics, and other related factors.[220] It is in this respect that the science of complex systems and the study of legal complexity are among the most practical of questions.

---

[218] *See Laplace's Demon*, CHAOS & FRACTALS, http://www.stsci.edu/~lbradley/seminar/laplace.html (last visited May 10, 2013).

[219] *See generally* Von W. Heisenberg, *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik*, 43 ZEITSCHRIFT FÜR PHYSIK A HADRONS AND NUCLEI 172 (1927) (Ger.) [hereinafter *Heisenberg's Uncertainty*]. For a useful description, see *Uncertainty Principle*, WIKIPEDIA, http://en.wikipedia.org/wiki/Uncertainty_principle (last modified Apr. 3, 2013).

[220] *See generally Heisenberg's Uncertainty*, *supra* note 219; *Uncertainty Principle*, *supra* note 219.

2. *How Complex Are the Underlying Dynamics? The Economics of Bubbles & Weather Versus Tides*

In making his grand pronouncement, Laplace failed to recognize the difference between two different sorts of systems—"simple" and "complex." Duncan Watts describes simple systems as "those for which a model can capture all or most of the variation in what we observe. The oscillations of pendulums and the orbits of satellites are therefore 'simple' in this sense, even though [it is] not necessarily a simple matter to be able to model and predict them."[221] By contrast, "complex systems" are those composed of a significant number of interconnected parts that as a whole tend to interact in a nonlinear manner.[222]

Economic and political systems, biological systems, and physical systems all feature such properties. This can frustrate attempts at predicting the outputs generated from such systems. Consider the world economy. Influenced by physicists, economists developed a number of models of both the overall economy as well as the performance of capital markets. Such canonical models include general equilibrium theory[223] and the efficient capital markets hypothesis.[224] As a first-order description of various market dynamics, these

---

[221] *See* Duncan Watts, *The Dream of Prediction: Why You Should Be Skeptical*, YAHOO! 2011 YEAR REV. (Dec. 27, 2011, 9:00 PM), http://2011.yearinreview.yahoo.com/2011/blog/8569/predictions-why-you-should-be-skeptical/.

[222] *Id.*

[223] *See, e.g.*, Kenneth J. Arrow & Gerard Debreu, *Existence of an Equilibrium for a Competitive Economy*, 22 ECONOMETRICA 265 (1954); Lionel W. McKenzie, *On the Existence of General Equilibrium for a Competitive Marke*t, 27 ECONOMETRICA 54 (1959).

[224] *See, e.g.*, Eugene F. Fama & Kenneth R. French, *The Capital Asset Pricing Model: Theory and Evidence*, J. ECON. PERSP., Summer 2004, at 25; Eugene F. Fama & Kenneth R. French, *The Cross-Section of Expected Stock Returns*, 47 J. FIN. 427 (1992); Eugene F. Fama & Kenneth R French, *Dividend Yields and Expected Stock Returns*, 22 J. FIN. ECON. 3 (1988); Eugene F. Fama, *Efficient Capital Markets: A Review of Theory and Empirical Work*, 25 J. FIN. 383 (1970). Leveraging ideas from behavioral science as well as complex systems, scholars have recently attempted to remedy the obvious weakness of the efficient capital market hypothesis. *See* Andrew W. Lo, *Reconciling Efficient Markets with Behavioral Finance: The Adaptive Markets Hypothesis*, 7 J. INVESTMENT CONSULTING, no. 2, 2005, at 21; Andrew W. Lo, *The Adaptive Markets Hypothesis: Market Efficiency from an Evolutionary Perspective*, J. PORTFOLIO MGMT., no. 5, 2004, at 15. One particularly stinging critique of modern finance comes from Fama's thesis advisor and one of the leading mathematicians of the twentieth century—Benoit Mandelbrot. The crux of the debate surrounds both the use of Brownian motion and the assumption of statistical independence present in much ECM literature. The key point from Mandelbrot is that price changes behave very differently from the simple geometric Brownian motion. Thus, he argued that the use of the standard ARCH (*Autoregressive Conditional Heteroskedastic Model*) was improper. Instead, his *fractal brownian motion* approach should be used. *See generally* BENOIT B. MANDELBROT & RICHARD L. HUDSON, THE (MIS)BEHAVIOR OF MARKETS: A FRACTAL VIEW OF RISK, RUIN, AND REWARD (2004); BENOIT B. MANDELBROT, FRACTALS AND SCALING IN FINANCE: DISCONTINUITY,

models provide a very solid depiction. That said, excessive reliance upon these models is the same class of mistake as that made by Laplace's Demon. Economic and political systems are not deterministic, and equilibrium actually does not exist.[225] Rather, equilibrium is a convenient description of countervailing dynamics that over some moving window have achieved stasis.[226] The shortcomings of equilibrium perspective are on clear display in the case of economic bubbles. Although they are not possible within a neoclassical framework, bubbles are fundamental features of markets.[227] Yet,

CONCENTRATION, RISK (1997); Benoit B. Mandelbrot & John W. Van Ness, *Fractional Brownian Motions, Fractional Noises and Applications*, 10 SIAM REV. 422 (1968).

[225] Equilibrium is a useful placeholder but ultimately a stylized description of the real dynamics of economic systems. For economic systems that feature anything other than trivial dynamics (i.e., something other than a world of two firms and two goods), nonequilibrium properties of those systems are what is actually interesting. One of the first economists to make this point was Nicholas Kaldor, famous among other reasons for his contribution to welfare economics. *See* N. Kaldor, *The Irrelevance of Equilibrium Economics*, 82 ECON. J. 1237 (1972). There are a variety of existing threads of nonequilbrium economics including econophysics and ecological economics. Such approaches are beginning to gain traction even in mainstream economics circles. *See, e.g.*, W. BRIAN ARTHUR, INCREASING RETURNS AND PATH DEPENDENCE IN THE ECONOMY (1994); ERIC D. BEINHOCKER, THE ORIGIN OF WEALTH (2006); JEAN PHILIPPE BOUCHAUD & MARC POTTERS, THEORY OF FINANCIAL RISK AND DERIVATIVE PRICING (2003); THE ECONOMY AS AN EVOLVING COMPLEX SYSTEM II (W. Brian Arthur et al. eds., 1997); W. Brian Arthur, *Complexity and the Economy*, 284 SCIENCE 107 (1999); A. Drăgulescu & V.M. Yakovenko, *Statistical Mechanics of Money*, 17 EUR. PHYSICAL J. B 723 (2000) (Ger.); Herbert Gintis, *The Dynamics of General Equilibrium*, 117 ECON. J. 1280 (2007); Herbert Gintis, *The Emergence of a Price System from Decentralized Bilateral Exchange*, 6 B.E. J. THEORETICAL ECON., no. 1, 2006, at 1; César A. Hidalgo & Ricardo Hausmann, *The Building Blocks of Economic Complexity*, 106 PROC. NAT'L ACAD. SCI. U.S. 10,570 (2009); John McCombie & Mark Roberts, *On Competing Views of the Importance of Increasing Returns, Cumulative Causation and Path-Dependence*, *in* THE FOUNDATIONS OF NON-EQUILIBRIUM ECONOMICS 12 (Sebastian Berger ed., 2009); J. Barkley Rosser Jr., *On the Complexities of Complex Economic Dynamics*, J. ECON. PERSP., Fall 1999, at 169; Wayne M. Saslow, *An Economic Analogy to Thermodynamics*, 67 AM. J. PHYSICS 1239 (1999); Tânia Sousa & Tiago Domingos, *Equilibrium Econophysics: A Unified Formalism for Neoclassical Economics and Equilibrium Thermodynamics*, 371 PHYSICA A 492 (2006); Tânia Sousa & Tiago Domingos, *Is Neoclassical Microeconomics Formally Valid? An Approach Based on an Analogy with Equilibrium Thermodynamics*, 58 ECOLOGICAL ECON. 160 (2006); K. Vela Velupillai, *Non-Linear Dynamics, Complexity and Randomness: Algorithmic Foundations*, 25 J. ECON. SURVS. 547 (2011); Martin L. Weitzman, *Economic Profitability Versus Ecological Entropy*, 115 Q.J. ECON. 237 (2000); Michael H. R. Stanley et al., *Scaling Behaviour in the Growth of Companies*, 379 NATURE 804 (1996); *see also* M. MITCHELL WALDROP, COMPLEXITY: THE EMERGING SCIENCE AT THE EDGE OF ORDER AND CHAOS (1992).

[226] *See supra* note 224 and accompanying text.

[227] To be clear, bubbles do not exist in the neoclassical model. *See* David Laibson, Professor, Harvard Univ., Asset Bubbles and Economic Dynamics (May 2010), *available at* http://www.econ.cam.ac.uk/cremic/news/stoneDL.html (noting that it is the neoclassical view that bubbles do not exist). It turns out that this is an area of economic theory in serious need of revision. While divided into different intellectual camps, a number of leading scholars have begun to bridge this gap in the literature. *See, e.g.*, CHARLES P. KINDLEBERGER & ROBERT Z. ALIBER, MANIAS, PANICS AND CRASHES: A HISTORY OF FINANCIAL CRISES (6th ed. 2011); CARMEN M. REINHART & KENNETH S. ROGOFF, THIS TIME IS DIFFERENT: EIGHT CENTURIES OF FINANCIAL FOLLY (2009); ROBERT J. SHILLER, IRRATIONAL EXUBERANCE (2d ed. 2005); DIDIER SORNETTE, WHY STOCK

they are difficult to predict as they arise through nonlinear interactions between components in a particular economic ordering.[228]

The emerging theory of financial bubbles highlights the weakness of deterministic models and deterministic thinking. What is needed is a higher order understanding of the relationship between a system's complexity and its predictability. Consider two different systems—weather systems and tide systems. Both fall on the fairly complex end of the spectrum, but from a prediction standpoint they could not be more different. Tides are generated by fairly complex dynamics, including tidal constituents such as the Earth's rotation, the topography of the ocean, and the position of the Moon and the Sun relative to Earth. Mathematicians such as Laplace, Kelvin, and Poincaré formulated a system of partial differential equations relating to properties such as the ocean's horizontal flow to its surface height. These equations, as well as a variety of subsequent refinements, have helped produce the types of quantitatively derived predictions that are published in books such as tide tables. Thus, in the case of tide systems, although they are complex they are often highly predictable.

Weather systems, by contrast, are not particularly predictable, although their precise predictability varies based on the underlying stochastic dynamics that are present. For example, consider temperature prediction in a midwestern state such as Michigan. In February, the expected high temperature hangs around the freezing point and features very little variation. However, as the calendar turns toward the threshold of a new season, the temperature can vary significantly. Even within a 48-hour period, the April temperature can change from a high of 30 degrees to 75 degrees to 30 degrees once again. Even under fairly ideal conditions, weather is a hard prediction problem and our best success is obtained within small time windows around the given event. In general, for weather prediction outside of a seven-to-fourteen-day window, the best level of prediction that is typically available is the almanac.[229]

MARKETS CRASH: CRITICAL EVENTS IN COMPLEX FINANCIAL SYSTEMS (2003); Kyle Chauvin et al., *Asset Bubbles and the Cost of Economic Fluctuations*, 43 J. MONEY CREDIT & BANKING (SUPPLEMENT) 233 (2011); Peter M. Garber, *Tulipmania*, 97 J. POL. ECON. 535 (1989); Sornette, *supra* note 100; Didier Sornette et al., *The 2006–2008 Oil Bubble: Evidence of Speculation, and Prediction*, 388 PHYSICA A 1571 (2009); Jean Tirole, *Asset Bubbles and Overlapping Generations*, 53 ECONOMETRICA 1071 (1985); Didier Sornette & Ryan Woodard, Financial Bubbles, Real Estate Bubbles, Derivative Bubbles, and the Financial and Economic Crisis (May 2, 2009) (unpublished manuscript), *available at* http://arxiv.org/abs/0905.0220.

[228] *See supra* notes 224–25 and accompanying text.

[229] It is important to note that numerical methods have brought significant improvement to the science of weather prediction. That said, given the complexity of the underlying dynamics there are real limits to

3. *The Limits of Prediction: Law as a Complex Adaptive System*

Legal systems are *complex adaptive systems* with elaborate levels of complexity[230] and extensive feedback loops between their respective institutions and agents as well as outside institutions and agents.[231] The precise level of complexity, of course, differs across sub-domains, but in general quantitative legal prediction is akin to weather prediction, not tide prediction. As such, an almanac-style level of prediction might be all that can be accomplished for law. Obviously, the almanac is hardly the quality of prediction that is offered in a tide table, and one would be ill advised to walk around with an almanac and months in advance boast with confidence about the precise temperature on a given day. That said, the almanac was still

predicting weather systems because they are dramatically nonlinear. *See* Edward N. Lorenz, *Deterministic Nonperiodic Flow*, 20 J. ATMOSPHERIC SCI. 130 (1963) (introducing, among other things, chaos theory which would later evolve into the science of complex systems). One key concept is the signal-to-noise ratio, which in general is at a fairly low ebb outside the ten-to-fourteen-day prediction window. *Id.*

[230] Much like the study of complexity in weather systems is instructive for its prediction, so to is the theoretical and empirical study of legal complexity. *See, e.g.*, RICHARD A. EPSTEIN, SIMPLE RULES FOR A COMPLEX WORLD (1995); Michael J. Bommarito II & Daniel M. Katz, *A Mathematical Approach to the Study of the United States Code*, 389 PHYSICA A 4195 (2010); Danièle Bourcier & Pierre Mazzega, *Toward Measures of Complexity in Legal Systems*, *in* PROCEEDINGS OF THE ELEVENTH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 211 (2007); Louis Kaplow, *A Model of the Optimal Complexity of Legal Rules*, 11 J.L. ECON. & ORG. 150 (1995); Susan B. Long & Judyth A. Swingen, *An Approach to the Measurement of Tax Law Complexity*, J. AM. TAX'N ASS'N, Spring 1987, at 22; Diarmuid Rossa Phelan, *The Effect of Complexity of Law on Litigation Strategy*, *in* LEGAL STRATEGIES: HOW CORPORATIONS USE LAW TO IMPROVE PERFORMANCE 335 (Antoine Masson & Mary J. Shariff eds., 2010); Peter H. Schuck, *Legal Complexity: Some Causes, Consequences, and Cures*, 42 DUKE L.J. 1 (1992); Joel Slemrod, *The Etiology of Tax Complexity: Evidence from U.S. State Income Tax Systems*, 33 PUB. FIN. REV. 279 (2005); Gordon Tullock, *On the Desirable Degree of Detail in the Law*, 2 EUR. J.L. & ECON. 199 (1995); Michelle J. White, *Legal Complexity and Lawyers' Benefit from Litigation*, 12 INT'L REV. L. & ECON. 381 (1992); R. George Wright, *The Illusion of Simplicity: An Explanation of Why the Law Can't Just Be Less Complex*, 27 FLA. ST. U. L. REV. 715 (2000); Byron Holz, Note, *Chaos Worth Having: Irreducible Complexity and Pragmatic Jurisprudence*, 8 MINN. J.L. SCI. & TECH. 303 (2006); *see also* Daniel Martin Katz & Michael Bommarito II, Measuring the Complexity of the Law: The U.S. Code (May 2013) (unpublished manuscript) (on file with author).

[231] *See, e.g.*, Daniel A. Farber, *Earthquakes and Tremors in Statutory Interpretation: An Empirical Study of the Dynamics of Interpretation*, 89 MINN. L. REV. 848 (2005); Gregory Todd Jones, *Dynamical Jurisprudence: Law as a Complex System*, 24 GA. ST. U. L. REV. 873 (2008); Daniel M. Katz & Derek K. Stafford, *Hustle and Flow: A Social Network Analysis of the American Federal Judiciary*, 71 OHIO ST. L.J. 457 (2010); David G. Post & Michael B. Eisen, *How Long Is the Coastline of the Law? Thoughts on the Fractal Nature of Legal Systems*, 29 J. LEGAL STUD. 545 (2000); J.B. Ruhl, *Law's Complexity: A Primer*, 24 GA. ST. U. L. REV. 885 (2008); J.B. Ruhl, *Regulation by Adaptive Management—Is It Possible?*, 7 MINN. J.L. SCI. & TECH. 21 (2005); J.B. Ruhl, *The Fitness of Law: Using Complexity Theory to Describe the Evolution of Law and Society and Its Practical Meaning for Democracy*, 49 VAND. L. REV. 1407 (1996); Bernard Trujillo, *Patterns in a Complex System: An Empirical Study of Valuation in Business Bankruptcy Cases*, 53 UCLA L. REV. 357 (2005).

extremely useful as the information contained helped farmers in their efforts to increase agriculture production. The standard is not perfection but rather benchmarking against alternative comparative models. Simply put, if one person has an almanac and the other does not, in the long run, the one with the almanac is likely to outperform.

IV. INNOVATION IN A MATURE INDUSTRY: PREPARING TO THRIVE (SURVIVE) IN THE AGE OF QUANTITATIVE LEGAL PREDICTION

A. "You Cannot Replace What I Do with a Computer"—The Legal Services Edition

Transitioning from the general to the applied case, in case it is not clear already, lawyers can be (and already have been) replaced by variants of the sort of technologies and approaches discussed *supra* in Parts I, II, and III. More generally stated, a nontrivial proportion of the tasks that white-collar professionals (including lawyers) undertake has been subjected to automation, process engineering, and displacement. The distribution of units of work will continue to move in one direction. For white-collar professions such as law, medicine, or finance, the medium-term future centers on a mixture of humans and machines working together to more efficiently deliver the services than either could alone. However, with respect to the existing market for legal services, the total number of humans needed to service the current demand for legal services[232] is simply going to decline. Without tapping previously untapped markets (and there is good reason to believe they can be tapped), law is an otherwise mature industry whose total labor market participation will likely never exceed its prior peak.[233]

[232] One way to change the existing demand is through a principled deregulation of the legal services market, and the development of a robust retail legal sector akin to "H&R Block Law." This is actively underway in the United Kingdom via the 2007 Legal Services Act. *See* Jane Croft, *Law Firms Look for Tie-Ups to Profit After Deregulation*, FIN. TIMES (London), Feb. 13, 2012, at 4; Neil Rose, *Wait for ABSs Is Over: Tesco Law Is Here*, GUARDIAN (Apr. 2, 2012, 7:13 PM), http://www.guardian.co.uk/law/2012/apr/02/abs-tesco-law-here. There exist a variety of paths to developing a robust retail legal sector. Perhaps the most important of these is nonlawyer ownership. How might this be achieved? The most likely avenues are interjurisdictional competition (Delaware-style liberalization) or through litigation (in the vein of *Bates v. State Bar of Arizona*, 433 U.S. 350 (1977)). For the litigation blueprint, see Renee Newman Knake, *Democratizing the Delivery of Legal Services*, 73 OHIO ST. L.J. 1 (2012).

[233] It is quite possible that those markets will be tapped. However, to do so requires a different type of lawyer—an entrepreneurial lawyer focused on the intersection of law, technology, and innovation in the delivery of legal services. *See* William Henderson, Commentary, *Why Are We Afraid of the Future of Law?*, NAT'L JURIST, Sept. 2012, at 8; *see also* Renee Newman Knake, *Cultivating Learners Who Will Invent the*

Consider the following simple hypothetical example. Imagine there were 1,000 units of legal work in the world and it currently required 100 humans to service those units. Assume just 300 units of work (not a terribly implausible number) were those where machines could (with assistance) mimic the outputs developed by humans. Further, assume that fifteen dually skilled or hybrid workers with a mastery of technology and law were required in order to develop such equivalent human + machine products. As designated in Figure 10 below, ten such individuals might be individuals with dual law and technology capacities.

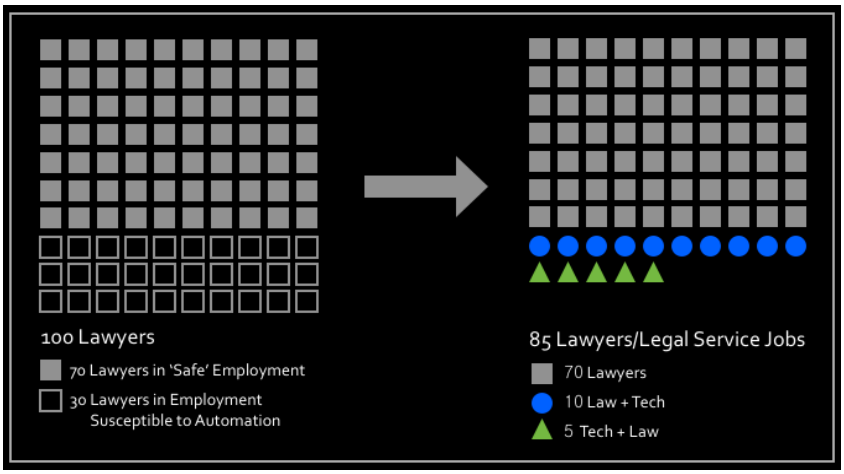
The very dynamics that create peril for some create possibility for others. As the traditional market for professional services continues to experience significant disruption and permanent contraction, there will be corresponding employment opportunities for those with very particular forms of dual capacities.[234] It turns out that going forward not every undergraduate major will be equally valid prelaw training. Namely, as displayed in Figure 10 below, the residual of this abstract labor market might feature seventy traditional jobs and fifteen new human + machine jobs. This would constitute a 30% decline in the traditional legal employment market. Further, this would represent a 15% decline in the size of the total legal services and legal product market and substantial returns for the entrepreneur who develops new, innovative delivery models to solve various legal problems using the appropriate mix of law, technology, and design.[235]

Future of Law Practice: Some Thoughts on Educating Entrepreneurial and Innovative Lawyers, 38 OHIO N.U. L. REV. 847 (2012). The status quo imposes significant access-to-justice consequences. *See* Gillian K. Hadfield, *The Price of Law: How the Market for Lawyers Distorts the Justice System*, 98 MICH. L. REV. 953 (2000); *see also* Gillian K. Hadfield, *Higher Demand, Lower Supply? A Comparative Assessment of the Legal Resource Landscape for Ordinary Americans*, 37 FORDHAM URB. L.J. 129 (2010).

[234] Perhaps this will be a very special form of legal consultant—one with expertise in legal information technology. *See* Tanina Rostain, *The Emergence of "Law Consultants,"* 75 FORDHAM L. REV. 1397 (2006). In addition, there are great opportunities for various legal entrepreneurs (whether or not they are lawyers). *See* William D. Henderson, From Big Law to Lean Law (Nov. 9, 2012) (unpublished manuscript) (on file with author).

[235] One very important component of legal information engineering is design. Design and aesthetic is the key to developing the sort of solutions that help solve the real problems of end users.

Figure 10: The "Death" of Big Law and the Rise of Law + Tech, Tech + Law



B. The Great Transition in the Market for Legal Services (and Legal Education?)

We are undergoing a great transition in the market for legal services and how to respond to it is arguably the most important question facing law schools, law students, law firms, and practicing lawyers. The legal service industry has experienced very little net job growth over the past fifteen years and significant contraction since the great recession starting in 2008.[236] Law schools are currently graduating roughly two students for every projected job opening,[237] and this trend is predicted to continue into the foreseeable future. A variety of factors are of course responsible for this overarching trend, including the aftermath of the 2008 financial crisis. However, going forward it is legal information technology, including but not limited to quantitative legal prediction, that will help define the future of the legal services industry. It is in this space where arbitrage opportunities abound for entrepreneurially minded law schools, law students, and practicing lawyers.

At its core, a professional school is designed to train students for success in professional careers in the relevant employment market. That starts with

[236] *See* William D. Henderson, *A Blueprint for Change*, 40 PEPP. L. REV. 461 (2013); *see also* BRUCE MACEWEN, GROWTH IS DEAD: NOW WHAT? (2013); Toby Brown, *Is the Legal Market Flat?*, 3 GEEKS & L. BLOG (July 10, 2012, 4:23 PM), http://www.geeklawblog.com/2012/07/is-legal-market-flat.html.

[237] *See* Henderson, *supra* note 236, at 476; *see also* TAMANAHA, *supra* note 18.

providing the sort of theory and skills training that can help students secure employment and become successful professionals.[238] Obviously, the training should be better tailored to the economic realities of the new legal labor market and many of those realities are being driven either directly or indirectly by technology.[239] In order to ensure that there is more there, some of the hothouse walls will have to come down.[240] The future belongs to those institutions and individuals who act as though their livelihoods depend upon it—because in many cases they do.[241]

[238] *See* Larry E. Ribstein, *Practicing Theory: Legal Education for the Twenty-First Century*, 96 IOWA L. REV. 1649 (2011); *see also* Daniel Martin Katz, *Thoughts on the State of American Legal Education—The* New York Times *Editorial Edition*, COMPUTATIONAL LEGAL STUD. (Nov. 28, 2011), http://computationallegalstudies.com/2011/11/28/thoughts-on-the-state-of-american-legal-education-the-new-york-times-editorial-edition/. Now it is important to note that better training and a greater return on investment are not likely to create more overall law jobs. If anything, the future of law is going to have fewer (and very different) lawyers. The ROI is not really within the control of any particular institution. What institutions control is the curriculum, and it is fair to say that the curriculum offered at most institutions is in need of a serious reboot. Some institutions are already embracing the future. *See, e.g.*, Jordan Furlong, *Law School Revolution*, LAW21 (June 25, 2012), http://www.law21.ca/2012/06/law-school-revolution/; Joanna Goodman, *Unconference! Beat Poetry and Quantitative Analysis—We Are All Futurists Now!*, LEGAL IT PROFS. (July 5, 2012), http://www.legalitprofessionals.com/index.php/col/joanna-goodman/columns/4438-unconference-beat-poetry-and-quantitative-analysis-we-are-all-futurists-now; Neil Rose, *The Next Big Thing*, LEGAL FUTURES (July 3, 2012), http://www.legalfutures.co.uk/blog/the-next-big-thing.

[239] *See* Ribstein, *supra* note 238. A classic adage is "someone *outside* your industry is working hard to disrupt it." That is certainly the case in law and this article outlines exactly how the advances in commercialized prediction technology are working to disrupt the market.

[240] *See id.* Professor Ribstein famously described the American legal academy as a hothouse—a place where some strange plants had grown because legal educators were almost entirely untethered from the legal marketplace. "Protected from the harsh winds of the markets, legal educators were free to develop a hothouse plant that bore little resemblance to anything that grew in the natural soil of law practice." *Id.* at 1655. "The hothouse walls are falling, leaving law schools to cope with markets." *Id.* at 1652.

[241] Ribstein, *supra* note 238. As I have argued elsewhere:

> Law school needs to transition from its liberal arts predisposition to a polytechnic research and teaching operation. From both a scholarship and training perspective, it is time to get serious about science, computation, data analytics and technology. [There is an] arbitrage opportunity in the market for legal education . . . for an institution(s) [to] move toward an "MIT School of Law."

Katz, *supra* note 238. "Here is the iron rule of the law school reform business—platitudes abound and specific proposals are few and far between." *Id.* So here is my proposal—the MIT School of Law. Daniel Martin Katz, *The MIT School of Law: A Perspective on Legal Education in the 21st Century*, SLIDESHARE (Oct. 14, 2011), http://www.slideshare.net/Danielkatz/the-mit-school-of-law-presentation-version-102-101411.