Cornell Journal of Law and Public Policy

Volume 27
Issue 2 *Winter 2017*

Article 3

The Law and Big Data

Caryn Devins
Western District of North Carolina

Teppo Felin
University of Oxford

Stuart Kauffman
Institute for Systems Biology, Seattle

Roger Koppl
Syracuse University

THE LAW AND BIG DATA

Caryn Devins, Teppo Felin,** Stuart Kauffman****
*& Roger Koppl*****

In this Article we critically examine the use of Big Data in the legal system. Big Data is driving a trend towards behavioral optimization and "personalized law," in which legal decisions and rules are optimized for best outcomes and where law is tailored to individual consumers based on analysis of past data. Big Data, however, has serious limitations and dangers when applied in the legal context. Advocates of Big Data make theoretically problematic assumptions about the objectivity of data and scientific observation. Law is always theory-laden. Although Big Data strives to be objective, law and data have multiple possible meanings and uses and thus require theory and interpretation in order to be applied. Further, the meanings and uses of law and data are indefinite and continually evolving in ways that cannot be captured or predicted by Big Data.

Due to these limitations, the use of Big Data will likely generate unintended consequences in the legal system. Large-scale use of Big Data will create distortions that adversely influence legal decision-making, causing irrational herding behaviors in the law. The centralized nature of the collection and application of Big Data also poses serious threats to legal evolution and democratic accountability. Furthermore, its focus on behavioral optimization necessarily restricts and even eliminates the local variation and heterogeneity that makes the legal system adaptive. In all, though Big Data has legitimate uses, this Article cautions against using Big Data to replace independent legal judgment.

INTRODUCTION 358
I. BIG DATA, LITTLE PERSONS: THE RISE OF BEHAVIORAL OPTIMIZATION AND "PERSONALIZED" LAW 362
A. *The "Big Data" Paradigm* 363
B. *Behavioral Optimization and "Personalized Law"* ... 365

* Research and Writing Attorney, Federal Public Defender for the Western District of North Carolina. This Article is submitted in an independent capacity and does not represent the views of the institutions with which the authors are affiliated.

** Professor of Strategy, Saïd Business School, University of Oxford.

*** Institute for Systems Biology, Seattle.

**** Professor of Finance, Whitman School of Management of Syracuse University. Roger Koppl gratefully acknowledges the financial support of the Earl V. Snyder Innovation Management Center of Syracuse University's Whitman School of Management.

II. BIG DATA'S LACK OF OBJECTIVITY AND THE NECESSITY OF THEORY 371
A. *Why Data, Observation and Measurement Need Theory* 372
1. Data and Observation is Theory-Laden 373
2. Law and Evidence: Seeing What You Believe ... 375
B. *The Indefinite Meanings of Law and Data: Law as Metaphor, Data as Compression* 379
1. Law as Metaphor and Language Game 379
2. *Data as Compression* 383
3. Law in a Normative Universe 386
III. LAW AND BIG DATA'S ILLUSORY PREDICTIVE POWER 388
A. *The "Frame" Problem: Algorithms in a World of No Entailing "Laws"* 389
B. *Due to the Frame Problem, Predictive Analytics Fails in Law* 393
1. Risk Assessment Tools in Sentencing 394
2. Risk Assessment Models in Financial Regulation 399
IV. LAW GONE VIRAL—BIG DATA'S UNINTENDED CONSEQUENCES 401
A. *Big Data and the Knowledge Problem: The Stifling of Legal Evolution* 403
B. *Network Effects, Herding Effects, Big Data's Centralization of Power* 408
CONCLUSION 413

INTRODUCTION

Big Data is considered the greatest innovation or the greatest peril of our times, depending on whom you ask. The legal system is certainly not immune from its effects. Legal tradition prizes consistency, stability, and uniformity in legal rules. Big Data promises to provide a scientific and evidence-based approach to law.[1] Simultaneously, Big Data signals the rise of behavioral optimization and "personalized law," as large-scale data analysis and predictive technologies are used to prescribe behavior and generate legal directives and recommendations precisely tailored to the client or regulated entity. In a Big Data world, laws are supposedly calibrated to policy objectives and optimal human behavior, based on a

[1] *See generally* Richard H. Fallon, Jr., *"The Rule of Law" as a Concept in Constitutional Discourse*, 97 COLUM. L. REV. 1 (1997); Amir N. Licht et al., *Culture Rules: The Foundations of the Rule of Law and Other Norms of Governance*, 35 J. COMP. ECON. 659 (2007); Daniel A. Farber, *The Rule of Law and the Law of Precedents*, 90 MINN. L. REV. 1173 (2005) (discussing the relationship between stare decisis and the rule of law).

machine analysis of massive amounts of data, thus cutting out human bias, incompetence, and error.[2] For example, policymakers might increasingly rely on data-backed approaches and customized microdirectives, or automated regulations based on data,[3] instead of statutes and regulations.[4] And clients might rely on predictive software instead of lawyers.

In this paper we question this "Big Data" paradigm. Our argument is three-fold. First, Big Data's asserted objectivity[5] is a myth. Data require theory in order to be interpreted and applied, and any single interpretation of data is rarely conclusive. Second, the behavioral and predictive models that employ Big Data are incapable of adapting to the creative evolution of the legal system.[6] These models cannot measure or predict all of the relevant variables that may influence the legal system, such as changes in values or novel uses of the law. Furthermore, large-scale data analysis makes behavioral prescriptions using averages, without taking into consideration the extremes, and the underlying heterogeneity (e.g., in beliefs and judgments) that animates and enriches the legal system as a whole. Finally, as a result of these shortcomings, rather than Big Data reflecting the legal system, the legal system recursively will come to reflect Big Data. Prescriptions of optimal behavior will suppress the much-needed lower-level variance or heterogeneity that makes legal systems adaptive and dynamic. A legal system that is overly reliant on Big Data would stray arbitrarily and undemocratically from the legal system's underlying values.[7] The more widely Big Data is used, the more it

---

[2] *See generally* Benjamin Alarie et al., *Regulation by Machine*, Dec. 1. 2016, http://www.mlandthelaw.org/papers/alarie.pdf; Anthony J. Casey & Anthony Niblett, *The Death of Rules and Standards*, 10–11 (Univ. of Chi. Pub. Law Theory Working Paper Grp., Paper No. 550, 2015).

[3] *See* Lawrence G. Baxter, *Adaptive Financial Regulation and RegTech: A Concept Article on Realistic Protection for Victims of Bank Failures*, 66 DUKE L.J. 567, 598 (2016).

[4] While some scholars have discussed Big Data from the perspective of targeted policymaking, *see supra* notes 2–3, others have focused on Big Data from the legal services or consumer perspectives, *see, e.g.*, John O. McGinnis & Russell G. Pearce, *The Great Disruption: How Machine Intelligence Will Transform the Role of Lawyers in the Delivery of Legal Services*, 82 FORDHAM L. REV. 3041 (2014). "Personalized law" marries the two perspectives, encompassing law's customization from the perspective of both policymakers and consumers of legal services. *See infra* Part I.

[5] *See, e.g.*, Chris Anderson, *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*, WIRED MAGAZINE, June 2008, at 1 ("For instance, Google conquered the advertising world with nothing more than applied mathematics. It didn't pretend to know anything about the culture or conventions of advertising – it just assumed that better data, with better analytical tools, would win the day."); *see also* DANIEL BOLLIER & CHARLES M. FIRESTONE, THE PROMISE AND PERIL OF BIG DATA 20–33 (The Aspen Institute, 2010) (describing various applications for Big Data).

[6] *See* infra Part III.

[7] Our argument here is reminiscent of our thesis that attempts to "design" legal institutions, such as the U.S. Constitution, in order to achieve particular preconceived objectives will inevitably fail. *See* Caryn Devins et al., *Against Design*, 47 ARIZ. ST. L.J. 609, 612–14 (2015).

will imbue and prescribe a sense of optimality and artificial inevitability to legal development.

What Big Data offers is, in many ways, opposed to rule of law traditions. While law is semantic, Big Data is syntactic. Law is abstract, values-based, and built on compromise. Big Data is empirical, algorithmic, and deterministic. Also, Big Data is inherently acontextual. Big Data cannot interpret itself, nor can it discern the indeterminate boundaries of legal principles.[8] Moreover, Big Data cannot discern or create novelty, unlike humans, who can update their "frames" or paradigms as their environment changes.[9] Big Data cannot innovate beyond the paradigms imposed by its creators. Even the most sophisticated machine learning techniques cannot tell us what factors might become relevant in response to new challenges.[10]

Big Data fundamentally differs from the common law system, which evolves creatively and unforeseeably beyond its original purposes.[11] The law "sprawls" unpredictably as adaptive agents within society—in their local contexts—find new meanings, even loopholes, in laws that enable new patterns of action and consequent risk or reward, that in turn spur new legal adaptations. This evolution confounds the rigidity of Big Data, which is limited to analyzing the past. Big Data tends toward a form of behavioral optimization, which inherently seeks to reduce variance—the precise variance that makes the legal system adaptive and dynamic over time. This adaptation and dynamism is inherent to the evolution of common law, which is largely bottom-up and decentralized. For example, judges gradually modify legal doctrines in response to changing conditions through individual cases.[12] However, Big Data

Like design efforts generally, Big Data assumes a fixed frame of reference for solving problems; it is incapable of accounting for all relevant variables, much less how those variables will evolve in the future.

8 *See e.g.*, Peter Goodrich, *Rhetoric and Modern Law*, *in* THE OXFORD HANDBOOK OF RHETORICAL STUDIES 613, 617–18 (Michael MacDonald ed., 2014) (exploring the limits of empirical data in the law); *see* SANDRA B. HALE, THE DISCOURSE OF COURT INTERPRETING: DISCOURSE PRACTICES OF THE LAW, THE WITNESS AND THE INTERPRETER, 4–8, 31–32 (2004).

9 *See* Asim Zia et al., *The Prospects and Limits of Algorithms in Simulating Creative Decision Making*, 14 EMERGENCE 89, 97 (2012) (discussing the limits of algorithms in interpreting affordances).

10 *See id.*

11 "The life of the law has not been logic: it has been experience." OLIVER W. HOLMES, JR., THE COMMON LAW, 1 (1881).

12 *See* Nicola Gennaioli & Andrei Shleifer, *The Evolution of Common Law*, 115 J. POLI. ECON. 43, 43–47 (2007) (providing various scholarly perspectives on the evolutionary benefits of common law); FRIEDRICH HAYEK, THE CONSTITUTION OF LIBERTY, 167–73 (1960); FRIEDRICH HAYEK, *The Changing Concept of Law*, *in* Law, Legislation, and Liberty 72, 85–88 (1973) (arguing that common law acted as a way for jurists to discover underlying canons of justice). *See generally* RICHARD A. POSNER, THE ECONOMIC ANALYSIS OF LAW (1973) (arguing that common law is a mechanism for discovering economically efficient legal rules).

analysis is centralized and top-down, and thus neither adaptive nor responsive to local experimentation and interpretation.

Importantly, in the search for objectivity, Big Data fails to meaningfully address a fundamental purpose of our legal system: resolving "competing conceptions of the good"[13] through institutions that derive their legitimacy from the consent of the governed. Rather, Big Data would impose an algorithm-based methodology that could introduce well-intentioned but highly problematic behavioral uniformity, and a disturbing lack of transparency and accountability to the legal system as a whole.

By relying on Big Data, policy-makers risk basing their decisions on mere correlations and averages identified in the data, with little to no understanding of the relevance of those correlations or the underlying causal relationships.[14] Because all relevant facts cannot be defined, let alone included, the models used to interpret Big Data are inherently biased in unknown and arbitrary ways. In turn, these models influence the conclusions yielded by the seemingly-objective data. Such decision-making would deemphasize traditional attributions of responsibility, harm, or risk[15] and create an intolerable risk of relying on unconstitutional prejudices. Shielded by the illusion of objectivity and "evidence-based" science, Big Data-based approaches could supersede the role of judges or elected officials and exercise outsized, yet poorly understood, influence over the legal system.

Big Data thus poses the risk of going "viral" as the algorithms' inputs and outputs influence each other recursively. Law would evolve in circular and arbitrary ways increasingly unmoored from both ever-evolving social life and the legal system's underlying values and purposes. Similar to herding behaviors in financial markets, Big Data systems would increasingly rely on self-reinforcing informational cascades, rather than substantive legal evolution.[16] Lacking a mechanism to comprehensively update its frame in accordance with evolving conditions, Big Data might not only impair democratic accountability and the rule of law, but also hinder wise legal change.

The algorithmic propagation of legal "arguments" using Big Data is catastrophic as law wanders away from human meanings and actions in

---

13 Michael Rosenfeld, *The Rule of Law and the Legitimacy of Constitutional Democracy*, 74 S. CAL. L. REV. 1307, 1350 (2001).

14 In this Article we use the term "correlation" broadly to mean any statistical measure or technique that might identify, in the language of the OED, a "mutual relation of two or more things." This broad meaning includes not only linear correlation, but also measures such as mutual information.

15 For a discussion of the role of causal norms in the law, *see* Antony Honore, *Causation in the Law*, in THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta, ed., 2010), https://plato.stanford.edu/entries/causation-law/.

16 *See* Ian Ayres & Joshua Mitts, *Anti-Herding Regulation*, 5 HARV. BUS. L. REV. 1, 18–20 (2015) (discussing problem of informational cascades).

the real world. Of course, Big Data might be used as a tool to inform judgment. But Big Data cannot decide questions of meaning, equity and justice—though it risks doing so under the guise of objectivity, evidence and science. We thus seek to reframe Big Data's more limited contribution and question the widespread optimism about its potential uses in the legal system.

In all, Big Data is likely to have unintended negative effects on legal interpretation, common law evolution, and distributions of authority.[17] In Part I, we predict that Big Data will generate "personalized" or customized law. In Parts II and III, we caution against overly optimistic assumptions regarding Big Data's supposed objectivity and predictive power. Finally, in Part IV we argue that these limitations of Big Data, if not taken seriously, could produce a pernicious reflexivity in the law, undermining the legal system's evolutionary capabilities and democratic accountability. In trying to save the law, Big Data could destroy it.

I. BIG DATA, LITTLE PERSONS: THE RISE OF BEHAVIORAL OPTIMIZATION AND "PERSONALIZED" LAW

The advent of Big Data has given rise to a "Big Data paradigm." This paradigm is based on the belief that theory is no longer necessary because applied mathematical and statistical techniques—based on algorithms—can "analyze" data and find optimal solutions, better than human programmers. Big Data evokes the mythical omniscient actor from rational choice theory, accounting for all available information, probabilities of events, and potential costs and benefits in determining courses of action.[18] While models of Big Data do not necessarily assume neoclassical omniscience on the part of actors, they nonetheless assume that data analysis and algorithms will provide the basis for prescribing optimal behaviors. These approaches assume that we can engineer mistakes and errors out of decision-making, and usher in a form of omniscience and rationality in legal decision-making. Thus the Big Data

[17] Our approach thus differs from that of the many other scholars who have critically examined Big Data in a legal context by focusing on the privacy implications of data collection. *See e.g.*, Omer Tene & Jules Polonetsky, *Privacy in the Age of Big Data: A Time for Big Decisions*, 64 STAN. L. REV. ONLINE 63 (2012); Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93 (2014); Ira S. Rubinstein, *Big Data: The End of Privacy or a New Beginning?*, 3 INT'L DATA PRIVACY L. 74 (2013).

[18] *See* Roger Koppl et al., *Economics for a Creative World*, 11 J. INSTITUTIONAL. ECON. 1, 4 (2013) (describing mainstream economic thinking as treating the system as "law governed" in the sense that the system can be described using a master set of equations). Under this conception, economic theory "can be loosely compared to a computer that has been programmed to execute this master set of equations." *Id.* *See also* Wolfgang Pietsch, *Aspects of Theory-Ladenness in Data-Intensive Science*, 82 PHILOSOPHY OF SCIENCE 905, 911 n. 5 (2015).

paradigm, as we will discuss, shares linkages with the overall goals of behavioral law and economics, which seeks to engineer optimality.[19] While these evidence-based, nudge-related and objective approaches are of course well-intentioned on the part of their proponents—and seek to improve human welfare and optimize decision making—we argue that they feature unintended consequences and pernicious feedback loops with dangerous consequences for law.

The Big Data paradigm is fatally flawed for the same reason that efforts to "design" institutions to fulfill predefined objectives are flawed. Social systems are fundamentally creative and evolutionary and will thus inevitably evolve in unforeseeable ways, beyond their underlying purposes. Like metaphors, law and data have indefinite meanings, which adaptive agents interpret and utilize in novel ways to achieve their own objectives. It is necessary to have a "frame" or paradigm to interpret these affordances and determine which ones are relevant or useful in a given context. Thus, the act of interpreting data inherently requires theory.

A. *The "Big Data" Paradigm*

As our society grows more complex, interconnected, and technologically advanced, data is generated that reflects this societal change, and potentially allows us to better understand this complexity.[20] With technology, every individual's movements, decisions, and purchases—every recordable detail of their lives—is captured and memorialized in the electronic realm. Due to the exponentially increasing efficiency of storage, the data is collected in centralized servers, stored, and analyzed in ways never before possible.

Big Data's power lies in capturing the massive reserves of data that are incidentally (as well as purposefully) generated through the increasingly detailed electronic documentation of our everyday lives.[21] Algo-

---

19 *See, e.g.*, On Amir & Orly Lobel, *Stumble, Predict, Nudge: How Behavioral Economics Informs Law and Policy*, 108 COLUM. L. REV. 2098, 2100–01 (2008); Christine Jolls et al., *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471, 1473–76 (1998); Russell B. Korobkin & Thomas S. Ulen, *Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics*, 88 CALIF. L. REV. 1051, 1059 (2000); Richard H. Thaler and Cass R. Sunstein, *Libertarian Paternalism*, 93 AM. ECON. REV. 175, 175 (2003). But see Gregory Mitchell, *Why Law and Economics' Perfect Rationality Should Not Be Traded for Behavioral Law and Economics' Equal Incompetence*, 91 GEO. L.J. 67, 77 (2002).

20 "The digital revolution is driving much of the increasing complexity and pace of life we are now seeing, but this technology also presents an opportunity. "Geoffrey West, *Big Data Needs a Big Theory to Go with it*, SCI. AM., May 1, 2013, https://www.scientificamerican.com/article/big-data-needs-big-theory/.

21 While "[t]here is no rigorous definition of big data," one conception is that "big data refers to things one can do at a large scale that cannot be done at a smaller one." VIKTOR M. SCHONBERGER & KENNETH CUKIER, BIG DATA: A REVOLUTION THAT WILL TRANSFORM HOW WE LIVE, WORK, AND THINK 7 (Reprint ed., 2013). Big Data historically referred to data sets

rithms are used to analyze these large and unconventional data streams in order to find ever-finer grained correlations between data points. Correlations generated through data mining are used to construct predictive models to assess the probability of a particular outcome based on given conditions.[22] Data analysis at such a large scale allegedly dispenses entirely with the need for theory: decisions can be made solely based on correlations, and the significance of causation is diminished.[23]

Scientists use this data to build, among other things, predictive analytical models using technologies such as genetic algorithms, machine learning, and agent-based modeling. They use applied statistics to determine patterns and predict future events, including risks and opportunities—striving for an acceptable level of reliability.[24]

The primary objective of predictive analytics is optimization, or the selection of the "best" outcome with regard to a set of available alternatives.[25] Advanced analytical techniques, such as genetic algorithms, are considered to be "intelligent" in the sense that they can "learn" solutions from the data.[26] For example, algorithms executed on modern computers enable the simulation of agents interacting in complex systems.[27] Similarly, machine learning is used to generate solutions that perform better than ones hand-coded by human programmers. Machine learning is most

---

of a very large size, but its meaning has become more complex and all-encompassing as its uses have transformed over time. *See* Gil Press, *12 Big Data Definitions: What's Yours?*, FORBES MAGAZINE, Sept. 3, 2014.

22 *See* SCHONBERGER & CUKIER, *supra* note 21, at 12 (noting that big data "is not about trying to 'teach' a computer to 'think' like humans," but instead to "apply[ ] math to huge quantities of data in or order to infer probabilities").

23 *See id.* at 7 ("Most strikingly, society will need to shed some its obsession for causality in exchange for simple correlations: not knowing *why* but only *what.*"); *see also* Simon DeDeo, *Wrong Side of the Tracks: Big Data and Protected Categories* 2 (Working Paper No. 1412.4643) (noting that our "most successful algorithms . . . do not provide causal accounts").

24 For a general description of Big Data analysis and techniques, *see* SCHONBERGER & CUKIER, *supra* note 21, at 2–12.

25 *See* Matthew A. Waller & Stanley E. Fawcett, *Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management*, 34 J. BUS. LOGISTICS 78 (2013) (explaining that "data science is the application of quantitative and qualitative methods to solve relevant problems and predict outcomes"); *see also* Casey & Niblett, *supra* note 2, at 10–11 (explaining how predictive analytics could be used to create "microdirectives" that optimize regulations to the individual).

26 "The wealth of new data, in turn, accelerates advances in computing — a virtuous circle of Big Data. Machine-learning algorithms, for example, learn on data, and the more data, the more the machines learn." Steve Lohr, *The Age of Big Data*, N.Y. TIMES, Feb. 11, 2012, http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html?mcubz=1.

27 *See* Zia et al., *supra* note 9.

effective in situations where accurate models are lacking and optimal solutions are difficult to identify.[28]

It is even claimed that the potent combination of computers, algorithms, and Big Data could generate more useful results than specialists who rely on hypotheses, models, and theories.[29] In a world of Big Data, companies like Google don't have to "settle" for models, theories, or mechanistic explanations of any sort. Instead, scientists "can throw the numbers into the biggest computing clusters the world has ever seen and let statistical algorithms find patterns where science cannot."[30] Applied mathematics and troves of data replace all theories of human behavior, and causal relationships need not matter because "[c]orrelation is enough."[31]

B. *Behavioral Optimization and "Personalized Law"*

Big Data strives for objectivity. Not coincidentally, Big Data's popularity is associated with movements toward empiricism in the law, such as behavioral law and economics and "evidence-based law."[32] For example, behavioral law and economics focuses on the biases and errors of legal actors and agents. It seeks to provide nudges and remedies to ensure optimal behavior.[33] The move to make legal reasoning and decision making more scientific, objective, and evidence-based of course should be lauded. But when it leads to the types of paternalism that removes meaningful choices from the system, it can have pernicious consequences.[34] In short, the universality or generality assumption behind

---

28 *See* Torbjørn S. Dahl et al., *A Machine Learning Method for Improving Task Allocation in Distributed Multi-robot Transportation*, in COMPLEX ENGINEERED SYSTEMS: SCIENCE MEETS TECHNOLOGY 2 (Braha et al. eds. 2006).

29 *See* Mark Graham, *Big Data and the End of Theory?*, THE GUARDIAN (Mar. 9 2012), https://www.theguardian.com/news/datablog/2012/mar/09/big-data-theory.

30 *Id.*

31 *See* Anderson, *supra* note 5, at 1.

32 "Evidence-based practices" can be described as "professional practices that are supported by the 'best research evidence,' consisting of scientific results related to intervention strategies. . . derived from clinically relevant research . . . based on systematic reviews, reasonable effect sizes, statistical and clinical significance, and a body of supporting evidence." Roger Warren, *Evidence-based Practices to Reduce Recidivism: Implications for State Judiciaries,* 20 CRIME & JUSTICE INST., 18–19 (2007). In law, advocacy for empiricism can be traced back to Felix Cohen's calls for a 'functional approach' to judicial decision-making and Louis Brandeis's use of social science research to bolster his argument in favor of finding social welfare legislation constitutional. *See* Felix Cohen, *Transcendental Nonsense and the Functional Approach,* 35 COLUM. L. REV. 809, 830–32 (1935).

33 *See* Joshua D. Wright & Douglas H. Ginsburg. *Behavioral Law and Economics: Its Origins, Fatal Flaws, and Implications for Liberty.* 106 NW. U. L. REV. 1033, 1050–51 (2012).

34 *See* Tom Ginsburg et al.. *Libertarian Paternalism, Path Dependence and Temporary Law,* 81 U. CHI. L. REV. 291, 341 (2014).

models, which try to engineer error out of systems, is suspect and highly problematic.

These moves toward behavioral optimization, coupled with Big Data, undoubtedly accelerate the trends of automation and disintermediation in the legal profession. For example, a trend towards "personalized law" is formed by providing individuals with an alternative to engaging professionals for their legal needs.[35] Indeed, Big Data has already begun to transform law firm practice by providing the tools to, among other things, predict legal costs and case outcomes, manage data for regulatory compliance, and reduce document review costs.[36] Moreover, these trends have culminated in the hiring of the first artificial intelligence lawyer, "ROSS."[37] As a piece of artificial intelligence software, ROSS offers opinions based on its analysis of huge batches of data, such as judicial cases.[38] Big Data is also spreading beyond law firms. Predictive modeling has transformed areas of law ranging from financial regulation to pre-trial release and sentencing determinations in criminal cases.[39] Big Data has also proven popular in local governance, from crime prevention to health initiatives.[40] And Big Data is commonly linked to excitement

---

[35] Although we use the term "personalized law" to refer to the use of data to customize legal rules and regulations, as analogized to personalized medicine, others have used the terms "microdirectives," or "personalized default rules." *See* Ariel Porat & Lior Strahilevitz, *Personalizing Default Rules and Disclosure with Big Data*, 112 MICH. L. REV. 1417, 1448 (2014); *see also* Cass R. Sunstein, *Impersonal Default Rules vs. Active Choices vs. Personalized Default Rules: A Triptych* (Nov. 5, 2012) (unpublished manuscript, on file with author).

[36] *See, e.g.*, McGinnis & Pearce, *supra* note 4, at 3041; *See also* Daniel M. Katz, *Quantitative Legal Prediction—or—How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry*, 62 EMORY L.J. 909, 914–15 (2013) (discussing how Moore's law affects the legal industry); Michael G. Bennett, *A Critical Embracing of the Digital Lawyer*, *in* EDUCATING THE DIGITAL LAWYER, § 12–01 (Oliver Goodenough & Marc Lauritsen, eds., 2012) https://repository.library.northeastern.edu/files/neu:332782/fulltext.pdf. (noting that "the cost savings of a brick-and-mortar-less practice alone practically assure" the embrace of digital lawyers and data-based legal practice).

[37] *See* Karen Turner, *Meet Ross, the Newly Hired Legal Robot*, WASH. POST (May 16, 2016), http://wapo.st/27kXLKj?tid=ss_mail&utm_term=.fa4b0615b5ca.

[38] *See id.*

[39] *See* SCHONBERGER & CUKIER, *supra* note 21, at 2–12 (describing various applications of Big Data, including in financial regulation); *See also* Sonja B. Starr, *Evidence Based Sentencing and the Scientific Rationalization of Discrimination*, 66 STAN. L. REV. 803, 842–46 (2014) (criticizing the effectiveness of using Big Data analytics to predict recidivism rates for sentencing purposes); *See also* Andrew G. Ferguson, *Big Data and Predictive Reasonable Suspicion*, 163 U. PA. L. REV. 327, 353–60 (2014) (discussing the implications of Big Data with regards to the "reasonable suspicion" standard).

[40] *See* Monica Davey, *Chicago Police Try to Predict Who May Shoot or be Shot*, N.Y. TIMES (May 23, 2016), https://nyti.ms/2lmUbiV; *See also* BOLLIER & FIRESTONE, *supra* note 5, at 20–33 (describing applications of Big Data).

about "evidence-based law,"[41] and as discussed above, behavioral law and economics.[42]

To illustrate Big Data's potentially pervasive effects, examples can be drawn from many legal perspectives, including civil litigants and banking regulators. For example, there are several ways in which law can be "personalized" using Big Data. The personalized law business model would involve synthesizing large amounts of data regarding the course and resolution of all manner of legal issues. Sophisticated predictive analytic software would be used to analyze data and compare it to the facts of a client's case. The results would be used to provide the universe of options others have taken in similar situations and to forecast the probability that a particular course of action would be favorable to the client.[43] Private technology such as software apps could also provide simple directives for legal consumers to comply with the law without having to weigh the reasonableness of their actions or search for the content of specific laws.[44] Moreover, "personalized law" could extend beyond the legal system and into personalized dispute resolution more generally. Individuals could, based on data, consider the efficacy of options outside of the traditional legal system, such as alternative dispute resolution.

In addition, the personalized law model might be used to automate the application of generalized legal rules with far more customization and nuance than can be found in automated legal document services today. For example, personalized law could set the default terms in contracts, wills, property deeds, and other situations, rather than adhering to one-size-fits-all statutory mandates.[45] Under such a personalized approach, individuals would be assigned default terms "tailored to their own personalities, characteristics, and past behaviors."[46] Instead of adhering to the default hierarchy of property distribution from the estate of an individual who dies intestate—to their spouse, their parents, other kin, or to the state—these default presumptions could be personalized to the individual in order to take into account their valued personal relationships, history, and characteristics.[47]

---

[41] *See generally* Jeffrey J. Rachlinski, *Evidence Based Law*, 96 CORNELL L. REV. 901 (2010).

[42] *See* Cass Sunstein, *Choosing Not to Choose*, 64 DUKE L.J. 1, 4–5 (2014).

[43] Law firms have already employed predictive software in settlement negotiations and e-discovery. *See, e.g.*, Don Philbin, *Improve Negotiation Outcomes with Analytics*, AMERICAN BAR ASSOC., July 30, 2015; *See also* McGinnis & Pearce, *supra* note 4, at 3041 (predicting that "[c]omputational services are on the cusp of substituting for other legal tasks—from the generation of legal documents to predicting outcomes in litigation").

[44] *See* Casey & Niblett, *supra* note 2, at 1.

[45] *See generally* Porat & Strahilevitz, *supra* note 35.

[46] *Id.* at 1417.

[47] *Id.*

Personalization of the law will only increase as more data becomes available regarding individuals' preferences. Due to the ever-increasing availability of data, "choice architects, or social planners, can . . . establish accurate default rules—in extreme cases, default rules that are tailored specifically to each member of the relevant population."[48]

Under the personalized law model, the client would not pay for the judgment of an experienced attorney so much as access to the collective experiences of thousands, if not millions, of other individuals. Big Data could thus create a process of disintermediation[49] where, instead of relying on a professional's judgment and experience, individuals and firms would use statistics to evaluate the probability that a particular course of action would be optimal under the unique set of circumstances.[50] In Coasean terms, Big Data could reduce transaction costs by making information more cheaply available to individuals lacking knowledge in legal specialties, thus allowing for more efficient resolution of disputes. This process has already begun with online legal service providers like Legal Zoom, but Big Data would further tailor dispute resolution to the individual using sophisticated predictive analytical techniques.

The process of disintermediation will also reshape law from a policymaker's perspective. Lawmakers will be able to create "microdirectives" or a "catalog of precisely tailored laws, specifying the exact behavior that is permitted in every situation."[51] These microdirectives would anticipate contingencies using data, in order to remain calibrated to their purpose without being over- or under-inclusive.[52] Additionally, these microdirectives—if couched as scientific evidence—might even displace legal decision-makers. Lawmakers would no longer have to create laws, and judges would no longer have to decide cases.[53]

Personalized law already has a precedent in the development of personalized and so-called "evidence-based" medicine, in which Big Data is

---

48 CASS SUNSTEIN, CHOOSING NOT TO CHOOSE: UNDERSTANDING THE VALUE OF CHOICE 205 (2015).

49 *See* Koppl et al., *supra* note 18, at 14 (describing intermediaries as individuals with "specialized knowledge" who transfer such knowledge to clients or otherwise helps them to cope with novelty in that area of specialization); *See also* Jeremy Howells, *Intermediation and the role of Intermediaries in Innovation*, 35 RESEARCH POL'Y. 715, 715–17 (2006) (discussing the role of intermediaries in facilitating information transfer in the innovation process).

50 *See generally* Bennett B. Borden & Jason R. Baron, *Finding the Signal in the Noise: Information Governance, Analytics, and the Future of Legal Practice*, 20 RICH. J.L. & TECH. 7 (2014) (discussing application of Big Data in e-discovery and other legal contexts); *see also* Bennett, *supra* note 36 (describing the world of "digital lawyers").

51 Niblett & Casey, *supra* note 2, at 10.

52 *Id.* at 1; *see also* Baxter, *supra* note 3, at 597–98 (describing the need for data-driven regulation to keep pace with increasingly-automated finance).

53 Niblett & Casey, *supra* note 2, at 10–11.

being used to customize treatments to the individual.[54] Such customization of medicine is a sea change from prior protocols, which base prescriptions of medications on a hierarchy of protocols established for particular conditions associated with the apocryphal "average patient."

The gold standard for evidence-based medicine has long been randomized clinical trials (RCT). In RCTs, treatments are tested through randomized, double-blind studies in which patients are divided into treatment and control groups. The validity of RCTs, however, has been increasingly called into question by rigorous statistical analyses that have demonstrated their empirical shortcomings.[55] Randomized studies attempt to isolate the causal effect of the studied treatment by eliminating confounding factors, but "such controlled stratification cannot be applied to the thousands of possible factors that influence the outcome when we do not know what those factors are."[56] Moreover, randomized clinical trials produce protocols for an idealized "average" person and do not take into account individuals' unique characteristics that may be outliers from the model. Finally, the regulatory "best-practice" formulary of hospitals, along with comparable standards of care in tort law, have the effect of encouraging homogeneity in treatment and, as a result, discouraging experimentation and innovation in medicine.[57]

Personalized medicine turns the paradigm of diagnosis based on the "average person" on its head by placing the patient's particular condition and needs at the center of the treatment protocol and comparing that patient with other patients across a variety of health matrices.[58] Applying probability analysis techniques to vast pools of patient data yields conclusions regarding the probability that a particular treatment would be effective for a particular individual, given his or her unique circumstances, and the risks that the treatment may pose. Based on this data, the doctor and patient may then evaluate which treatment would be appropri-

---

54 BOLLIER & FIRESTONE, *supra* note 5, at 25 (explaining that "[i]dentifying new correlations in data can improve the ways to develop drugs, administer medical treatments and design government programs").

55 Iodannis argues that the majority of medical research papers present conclusions that are inadequately supported by the underlying data due to small sample sizes, poorly designed research protocols, lax statistical standards and other problems. *See* John Iodannis, *Why Most Published Research Findings Are False*, 2 PLOS MED. 696, 698–700 (2005). But while some of these deficiencies in the statistical analysis could, at least in theory, be corrected, there are deeper problems with the assumptions underlying RCTs and the structure of the studies. If the causal "action" is in the confounds, exclusive reliance on RCTs will block progress.

56 Stuart Kauffman et al., *Transforming Medicine: A Manifesto*, SCI. AM. WORLDVIEW 28, 28–31 (2014).

57 *See id.*

58 *See generally* Margaret A. Hamburg & Francis Collins, *The Path to Personalized Medicine*, 363 NEW ENGLAND J. MED. 301, 301–14 (2010); *see also* Colin Hill, *Can Big Data Save my Dad from Cancer?* FORBES (Dec. 18, 2012), https://www.forbes.com/sites/colinhill/2012/12/18/can-big-data-save-my-dad-from-cancer/.

ate based on its relative benefits and risks. Similar to personalized medicine, Big Data aims to replace traditional legal services with customized, data-driven legal solutions.[59]

In a data-driven legal system, empirical analysis would overtake the judgment of experts. In the context of criminal sentencing, for example, it has been argued that "relying upon gut instinct and experience is no longer sufficient. It may even be unethical—a kind of sentencing malpractice."[60] In many situations, not only would judgment no longer be required, it would be considered poor legal practice.

Like all experts, judges and lawmakers may produce incorrect opinions. This risk is heightened where experts enjoy a monopoly over their area of expertise and may choose for others instead of merely advising them.[61] Behavioral research suggests experts have various motivations aside from truth-seeking, and their opinions may be skewed by self-interest, institutional incentives, or observer effects such as representativeness bias, availability bias, and adjustment and anchoring bias.[62] Moreover, the cognition of experts is limited and erring. Realist legal scholarship has identified extralegal factors, such as a judge's personal political biases, that may influence legal decision-making.[63]

Just as personalized medicine discredited one-size-fits all RCTs, personalized law may undermine many "gold standards" of law that lack a rigorously-studied empirical basis.[64] For example, the use of juries to determine guilt in a criminal trial, guaranteed by the Sixth Amendment, has been revealed to be tragically flawed in cases where DNA evidence proved that innocent people were convicted.[65] Part of the reason for the unpredictability, even seeming arbitrariness, of jury trials is that the governing rules are based on dubious empirical assumptions. In the context

---

59 *See* SUNSTEIN, *supra* note 48, at 208 (arguing that "personalized default rules are the wave of the future" and "[w]e should expect to see a significant increase in personalization as greater information becomes available about the informed choices of diverse people").

60 Richard Redding, *Evidence-Based Sentencing: The Science of Sentencing Policy and Practice*,1 CHAP. J. OF CRIM. JUST. 1, 1–2 (2009).

61 ROGER KOPPL, *The Rule of Experts*, *in* THE OXFORD HANDBOOK OF AUSTRIAN ECONOMICS 344–46 (Peter Boettke & Christopher Coyne eds. 2015).

62 *See id* at 350.; *see also* Daniel Kahneman, *Maps of Bounded Rationality: Psychology for Behavioral Economics*, 93 AM. ECON. REV. 1449, 1450–54 (2003) (describing how different forms of cognition and accessibility of information affect judgments).

63 *See generally* Tracey E. George & Lee Epstein, *On the Nature of Supreme Court Decision Making*, 86 AM. POL. SCI. REV. 323 (1992) (describing the "legal" and "extralegal" models of judicial decision making and proposing an integrated model "that contemplates a range of political and environmental forces and doctrinal constraints").

64 *See* Alex Kozinski, *Criminal Law 2.0*, 44 GEO. L.J. ANN. REV. CRIM. PROC. 1, 1 (2015) ("Although we pretend otherwise, much of what we do in the law is guesswork.").

65 *See* INNOCENCE PROJECT, https://www.innocenceproject.org/all-cases/#exonerated-by-dna (last visited Sept. 11, 2017). (explaining that as of March 2016, there were 340 people previously convicted of serious crimes in the United States who had been exonerated by DNA testing).

of the Rules of Evidence, Judge Posner has famously criticized the complicated web of hearsay exceptions as not being based on any kind of empirical evidence, and for not sufficiently permitting a judge's common sense consideration of the reliability of the hearsay testimony within the circumstances of the case.[66]

Further, these wrongful convictions have rested, in part, on "gold standards" in forensic science that have also been revealed to rest on dubious assumptions. Studies have questioned the reliability of basic forensic techniques, including everything from bite-mark analysis to fingerprinting.[67] Indeed, a recent report found that FBI analysts had falsely testified to the accuracy of hair analysis techniques over a period of decades.[68]

Data analysis, it is hoped, will reveal superior methods of seeking the truth, particularly when the stakes are so high. We caution, however, against the assumption that Big Data is capable of transcending those imperfections. As we shall see, Big Data requires that data be gathered to a central server where it is analyzed by algorithms designed by largely anonymous experts with unchecked power. It may thus be more subject to "expert failure" than the legal processes its enthusiasts would have it supplant.[69]

II. BIG DATA'S LACK OF OBJECTIVITY AND THE NECESSITY OF THEORY

Big Data's supposed objectivity and predictive power are overstated, at least when applied to highly complex evolutionary systems such as the legal system. Data always require interpretation, which ne-

---

66 *See United States v. Boyce*, 742 F.3d 792, 801 (7th Cir. 2014) (Posner, J., concurring) (criticizing the present-sense impression exception to the hearsay rule as having "neither a theoretical nor an empirical basis").

67 Kozinski, *supra* note 64, at 2–3 (describing various flaws in forensic science); *see also* Radley Balko, *How the Flawed "Science" of Bite Mark Analysis Has Sent Innocent People to Prison*, WASH. POST (Feb. 13, 2015), https://www.washingtonpost.com/news/the-watch/wp/2015/02/13/how-the-flawed-science-of-bite-mark-analysis-has-sent-innocent-people-to-jail/?utm_term=.19e03c041fea; Michael J. Saks & Jonathan J. Koehler, *The Coming Paradigm Shift in Forensic Identification Science*, 309 SCI. 892, 895 (2005); Andy Newman, *Fingerprinting's Reliability Draws Growing Court Challenges*, N.Y. TIMES (Apr. 7, 2001), http://www.nytimes.com/2001/04/07/us/fingerprinting-s-reliability-draws-growing-court-challenges.html?mcubz=1.

68 *See* Spencer S. Hsu, *FBI Admits Flaws in Hair Analysis Over Decades*, WASH. POST (Apr. 18, 2015), https://www.washingtonpost.com/local/crime/fbi-overstated-forensic-hair-matches-in-nearly-all-criminal-trials-for-decades/2015/04/18/39c8d8c6-e515-11e4-b510-962fcfabc310_story.html?utm_term=.03dc9bb56187 (explaining that "over a more than two-decade period before 2000, nearly every FBI examiner gave flawed forensic hair testimony in almost all trials of criminal defendants"); *see also* Roger Koppl, *How to Improve Forensic Science*, 20(3) EUROPEAN J.L. & ECON. 255, 257 (2005).

69 For an economic theory of experts *see* ROGER KOPPL, EXPERT FAILURE, Cambridge University Press. (forthcoming)

cessitates theory and, correspondingly, evaluative judgment by humans. Further, Big Data cannot foresee the fundamentally creative, non-algorithmic evolution of the legal system, and its predictive power is limited.

Data is inherently both subjective and incomplete, rather than objective and determinant.[70] Without being filtered and theoretically-driven, mere data only produces a meaningless sea of correlations and must be simplified in order to be understood. This act of simplification (and aggregation), like legal interpretation, requires theory. Even the very act of deciding what data to gather in the first place—what to measure and observe, when and how—necessitates a theory.

Moreover, the partially defined, partially foreseeable interpretations of law and data function as affordances, which adaptive agents interpret and utilize in novel and unpredictable ways in order to achieve their own objectives. This creative evolutionary process leads the system to constantly update its "framing" of legal problems in non-algorithmic ways that cannot be predicted by Big Data. In other words, any change in law or legal practice induced by Big Data will become a tool for unknown persons to use in unknown, unknowable, and unprestateable ways for unknown, unknowable, and currently unimaginable ends. The unknowability of future ends and affordances is an ineradicable limit on the predictive power of Big Data.

From the perspective of a legal consumer, the flaws of Big Data may not seem to matter so long as the results are useful. But from the standpoint of the legal system, widespread use of Big Data threatens in the long run to undermine the integrity and efficacy of the law.

A. *Why Data, Observation and Measurement Need Theory*

The relationship between data and theory requires more careful attention. One of the central myths associated with data is that it is somehow impartial and pure—and that theory only muddles. It is argued that questions of truth and reality should be empirically and observationally resolved, rather than resorting to theory. But the notion of pure data, measurement or objective observation is a myth. Far from obliterating the need for theory, Big Data in fact makes the role of theory ever more important. As we'll discuss later, the law itself is also theory-laden, or as Dworkin puts it, legal reasoning is inherently "theory-embedded."[71]

---

70 LISA GITELMAN, INTRODUCTION TO "RAW DATA" IS AN OXYMORON 1, 7 (Lisa Gitelman ed. 2013).

71 Ronald Dworkin, *In Praise of Theory*, 29 ARIZ. ST. L. J. 353, 354 (1997).

1. Data and Observation is Theory-Laden

The problem with giving priority to data over theory is that any observation and measurement is inherently "theory-laden." The very notion of data—or the singular "datum" (etymologically: a thing given or fact)—is a contested term.[72] As noted by the philosopher Karl Popper, "all observation [or data, we might insert] involves interpretation in light of theories."[73] Thus we do not have any form of "direct access" to data, facts or reality through observation or perception.[74] Neither the human perceptual apparatus (vision)—nor any technologies we might utilize to observe or measure the world—provide us with direct access to reality. Theory is always needed.

The importance of theory relative to data and observation is readily illustrated by contexts where very little data is available (or even needed), but insights nonetheless emerge. Thus it is not the *amount* of data, as implied by "Big" Data, that somehow gives us a definitive understanding of the world or the nature of reality. Rather, theories guide us toward certain data and observations, and suggest interpretations that help us understand the world. Some of the most fundamental findings in science were not established with a large dataset or some overwhelming number of observations. Quite the opposite. For example, a single data point or observation—a so-called *experimentum crucis*—was the basis for confirming Newton's theory about the nature of light. Similarly, Einstein's general theory of relativity was confirmed with a single observation, as specified, guided and predicted by the data. The central insights came from the theory, not the data. Only minimal data (in fact, one observation: "small" data) was necessary to verify Einstein's theory, when in 1919 the astronomer and physicist Arthur Eddington observed the solar eclipse on the island of Principe off the coast of Africa. The theory provided guidance for what data to look for and expect, and how to interpret the finding.[75] It is only with the guidance of that theory that data and observation made sense.

While data may not conclusively "verify" a theory in some strong sense, such as that of 20th century Logical Empiricism, a crucial observation may well bring scientists to accept a theory into the body of results considered, as it were, "established until further notice." But the theory does the bulk of the intellectual work. Now, naturally there have been many replications of these findings (which originally may have

---

72 *See* GITELMAN, supra note 70, at 9–10.

73 KARL R. POPPER, *A Realist View of Logic, Physics, and History*, *in* OBJECTIVE KNOWLEDGE AN EVOLUTIONARY APPROACH 285, 295 (8th ed. 1994).

74 *See* CARL FRITH. MAKING UP THE MIND: HOW THE BRAIN CREATES OUR MENTAL WORLD 81 (2007).

75 Michael Polanyi, *Genius in Science*, 34 ARCHIVES DE PHILOSOPHIE 593, 601 (1971).

used small samples), and much additional theoretical and empirical development. And theories of course are not conclusive. And some theories require large amounts of data, and in other cases empirical anomalies lead to further understanding. But theory always plays a central role, as observation and data itself is theory-laden. Thus the most significant advances in science and understanding have come from theory, not from data.

Theories are needed not only for interpretation, but also to generate hypotheses, conjectures and ideas about what to observe and measure in the first place. Theories—whether intentional or not—are inherent to data, as the very fact of specifying which data to gather (and how) is driven by some expectation about what can and should be observed and measured. Whether something is even seen and registered (say, by scientific instruments or by an observer), and thus can be considered and gathered as data, has much to do with the theories that guide expectations and direct observation and data gathering. As put by Einstein, "whether you can observe a thing or not depends on the theory which you use. It is the theory which decides what can be observed."

The problem of measurement—and indeterminacy of data—has frequently been discussed by physicists and philosophers of science. In "Against Measurement," the physicist John Bell highlights how measurement is always observer-dependent and thus inherently not objective.[76] Observation always comes from *a* perspective, a point of view, and any given data point, or even large collection of data points, can be true to one way of seeing things. But this one way, the so-called truth or reality, does not somehow rule out alternative explanations or interpretations. Even the introduction of scientific instruments and methodologies does not erase the effect.[77] The very act of measurement can be said to distort the "purity" of the data that is being gathered. But ironically, the theories that guide this measurement—while distorting—also provide the very mechanism that makes understanding and explanation possible in the first place.

The role of theory, then, is to tell us where to look: which data to look for, which data to gather and why—and how to interpret it. The philosopher Karl Popper discusses the idea of "search light" versus "bucket" approaches to knowledge.[78] Bucket theories implicitly assume that we can somehow gather "the data" in a great bucket of "facts"—by simply absorbing or assimilating stimuli and our surroundings (some-

---

[76] John Bell, *Against Measurement*, 3 PHYSICS WORLD 30, 30–33 (1990).

[77] Teppo Felin et al., *Rationality, Perception, and the All-Seeing Eye*, 24 PSYCHON. BULL. REV. 1040, 1041–44 (2017).

[78] Michael ter Hark, *Popper's Theory of the Searchlight: A Historical Assessment of Its Significance in* Rethinking Popper 175 (Zuzana Parusniková & Robert S. Cohen eds. 2009).

thing which Big Data certainly facilitates)—and *then* build up our understanding simply by letting the facts speak for themselves. From this perspective the mind then can be seen as a massive computational device or camera that stores information about its surroundings and makes correct inferences. However, a searchlight approach to the mind recognizes that "facts" can be "seen" only when we shine the light of inquiry on them, and it is theory that tells us where to point the searchlight of inquiry.[79]

Bucket theories assume that scientists or technologies can somehow automatically and objectively capture data and process it, which will naturally lead to insights about the world. Bucket theories thus fit nicely into the Big Data paradigm, as the focus is on capturing—objectively in camera-like fashion—as much information as possible, with the assumption that the data and observations themselves will yield answers and understanding. However, a search light theory of knowledge focuses on the role that theories and attention play in directing us toward what data to look for in the first place and why. Theories reveal the types of observations and data that are relevant, and provide intuition for how to interpret the data.

These types of philosophical issues may, of course, seem far removed from the practice of Big Data—particularly as it applies to law—but they are very germane since they raise questions about what exactly can be established with data, and the role of theory in guiding observation and the gathering and analysis of data. In the very least, the idea that "theory is dead"—suggested by some Big Data advocates—can certainly be questioned. We discuss these implications next, as they apply to varied aspects of the law.

2. Law and Evidence: Seeing What You Believe

Law of course is heavily focused on perception, facts, and data—factors such as evidence, observation and witness testimony. For example, the *Federal Rules of Evidence*, adopted in 1975, establishes the law and standards of proof for evidence. It specifies such matters as what counts as evidence, who can be considered an expert or witness, what evidence or testimony is admissible and why, what represents fact versus hearsay, and so forth.

As we have explained, all of this legal activity related to evidence is guided by underlying assumptions about the nature of observation, data and even science—and how these can be used to establish facts, causality and responsibility. In short, the *Federal Rules of Evidence* seek to establish truth. That is, they seek to ensure "that the truth may be ascertained

[79] *Teppo Fellin et al, supra* note 78, at 1051.

and proceedings justly determined" (Rule 102).[80] In criminal law, the adversarial system of the law provides a mechanism for vetting the guilt or innocence of an accused. The Sixth Amendment's Confrontation Clause allows the accused to be confronted by witnesses and evidence, and a chance for counter-evidence and cross-examination.

But the data, witnesses, and evidence brought to bear in this process of course don't speak for themselves—it is necessarily animated and brought to life by the human actors who gather, analyze, present, aggregate, and judge based on that data and evidence.[81]

While we might want evidence, observation and witnesses to objectively assess matters related to the law (say, guilt or innocence, responsibility), all of this activity is necessarily guided by the underlying expectations, theories, and interests of the different legal and other actors involved. The data is not necessarily neutral or objective. Prosecutors have a constitutional duty to act as "neutral and detached magistrate[s]."[82] But in practice they often see any evidence or potential witness in terms of the culpability of the accused.[83] Defense attorneys have a constitutional duty of "vigorous and effective advocacy" of their clients.[84] They must therefore see matters through the lens of innocence. And they probably do, in general, when the defense is both well-paid and chosen by the accused. But public defenders often lack adequate incentives and resources to provide a vigorous defense.[85] Now, no bias or any form of maliciousness is necessarily involved in this process, though of course it is possible. But the incentives faced by the opposing parties may diverge radically—and the data or evidence-gathering and success of the respective parties is based on their *ex ante* mission (or role) to either successfully prosecute or defend the accused, *independent* of what actually happened.

Of course, even without any bias or malicious intent, the whole process of gathering evidence and data is, *a priori*, given by one's relationship to a particular client—based on defined roles and the sought-after

80 An important forerunner to the *Federal Rules of Evidence* was James Thayer's treatise. JAMES B. THAYER, A PRELIMINARY TREATISE ON EVIDENCE AT THE COMMON LAW (1896).

81 Brian Leiter & Ronald J. Allen, *Naturalized Epistemology and the Law of Evidence*, 87 VA. L. REV. 1491, 1499–1505 (2001); *cf.* Elanor Swift, *One Hundred Years of Evidence Law Reform: Thayer's Triumph*, 88 CAL. L. REV. 2437, 2451–55 (2000) (discussing the impact of the cognitive process in the context of a proposed evidentiary scheme).

82 *Coolidge v. New Hampshire*, 403 U.S. 443, 449 (1971).

83 Roger Koppl & Meghan Sacks, *The Criminal Justice System Creates Incentives for False Convictions*, 32 CRIM. JUST. ETHICS 126, 148–49 (2012).

84 *Jones v. Barnes*, 463 U.S. 745, 754 (1983).

85 *See* Stephen Schulhofer & David D. Friedman, *Rethinking Indigent Defense: Promoting Effective Representation through Consumer Sovereignty and Freedom of Choice for All Criminals*, 31 AM. CRIM. L. REV. 73, 77–80 (1993); Stephen Schulhofer & David D. Friedman, *Reforming Indigent Defense: How Free Market Principles Can Help to Fix a Broken System*, 666 POL'Y ANALYSIS 1, 3–5 (2010).

outcome. Thus, when one is, in essence, "primed" to look for confirmatory data and evidence for a particular outcome, one is likely to find such evidence.[86] In fact, the very same piece of evidence, based on this priming, may be favorably interpreted to advance the cause. Psychological and perceptual experiments highlight how our expectations, or initial impressions, can drive individuals to search for, perceive and find confirmatory evidence for particular interpretations.

Of course, the above argument might suggest that letting data speak for itself would yield better outcomes, rather than relying on biased and incentive-primed human actors to make sense of what happened in, say, a criminal trial. But as we've argued, the data rarely speaks for itself, but always requires interpretation and context. Thus the adversarial system of prosecution and defense can provide a check on this system. But there are no algorithms or data-driven alternatives which somehow might supplant this system.

The overall emphasis on data, and the bucket theory of mind (discussed above) as it relates to the law, misses some fundamental facts related to evidence gathering and (more broadly) perception.[87] This might aptly be illustrated by a "Sherlock" model of evidence (similar to Popper's aforementioned "searchlight" model), which can readily be contrasted with a naïve, camera-like model of evidence, similar to the approach suggested by Big Data.[88] To make this point concrete: imagine a crime scene investigation where we might contrast the approach taken by a naïve policeman versus a prototypical investigator like Sherlock Holmes. The naïve policeman might seek to collect, process and perhaps photograph everything possible at the crime scene—in effect, to generate a form of big data that captures as much information as possible. However, the problem of course is that almost any bit of data at the crime scene could be relevant to establishing what actually happened. The crime scene is constituted by all kind of facts, innumerable data, and *potential* evidence. Anything could be relevant. Though only some, (extremely) small portion of this information is actually relevant for the case at hand. This then creates a general problem—relevant to any context—between trying to assess which data or facts are relevant and important and which are not. The naïve policeman seeks to capture everything,

---

86 In the cognitive sciences and psychology there are active debates about what priming means and how the human attentional and perceptual mechanism works. *See, e.g.*, Edward Awh et al., *Top-down Versus Bottom-up Attentional Control: A Failed Theoretical Dichotomy*, 16 TRENDS IN COGNITIVE SCI. 437 (2012) (summarizing some of the key issues and debates regarding priming).

87 *Cf.* Felin et al., *supra* note 78.

88 Jan J. Koenderink, *Geometry of Imaginary Spaces*, 106 J. PHYSIOLOGY 173, 176–77 (2012). For further discussion, see Nick Chater et al., *Mind, Rationality and Cognition: An Interdisciplinary Debate*, 24 PSYCHON. BULL. REV. 25–26 (2017, forthcoming).

while Sherlock's investigation is driven by a plot or theory. As perception scholar Jan Koenderink puts it, the problem for the naïve policeman, then, is that "the size of this file [of "all" the potential data from the crime scene] is potentially limitless, for the world is infinitely structured. There is no end to which fact, perhaps even on a molecular scale (think of DNA traces), might eventually prove to be important. [But] facts are not 'evidence,' they are simply facts."[89] And herein lies the problem with the Big Data or bucket model. The simple gathering of vast amounts of data is meaningless without some kind of plot or theory about what may have happened, about what might be relevant. There is no automatic way to process the scene. Some kind of searchlight or theory is needed—supplied by Sherlock or human intuition. The data itself may of course suggest and provides clues for what happened, but even the identification of these clues and data must be theoretically informed. Thus, in science as in law, data is not some kind of panacea for understanding what happened or for solving problems. Theoretical intuition needs to guide this activity.

The present obsession with data, and the associated suggestions that we now live in a post-theoretical world, miss the fact that data is merely an input for higher levels of understanding. This has been recognized by some in the field of information science.[90] Data sits at the very lowest rung of the so-called data-information-knowledge ("DIKW") hierarchy or pyramid. There are no higher levels of understanding without the questions, probes, and theories which guide us from lower-level observations and data to higher-level knowledge and wisdom. Thus we might quote the poet T.S. Eliot who seemingly foresaw the increased emphasis on the lower levels of the hierarchy, at the expense of wisdom and understanding: "Where is the wisdom we have lost in knowledge? Where is the knowledge we have lost in information?"[91]

In the legal context, the common law legal system can be seen as providing precisely this type of wisdom, moving us from a raw input like data, toward information, knowledge and wisdom. The legal system can be seen as an accrual system for wisdom, where pockets of wisdom are found through the contextual information, local knowledge and overall wisdom that emerges as disparate legal actors argue and interact over time. In other words, the interaction of heterogeneous agents with disparate motivations and bits of data, information and knowledge lead to the accrual of various forms of wisdom which respond to local circum-

---

89 *Id.* at 176.

90 *See, e.g.*, KENNETH BOULDING, "NOTES ON THE INFORMATION CONCEPT" (1955); *see also* CHRISTINE L. BORGMAN, BIG DATA, LITTLE DATA, NO DATA (MIT Press ed., 2015); ROB KITCHIN, THE DATA REVOLUTION: BIG DATA, OPEN DATA, DATA INFRASTRUCTURES AND THEIR CONSEQUENCES (Robert Rojek ed., 2014).

91 T.S. ELLIOT, CHORUSES FROM THE ROCK (1936).

stances. This process is scarcely computational, but rather requires human intuition and theoretical processing which cannot be left to algorithms and data alone.

B. *The Indefinite Meanings of Law and Data: Law as Metaphor, Data as Compression*

The word "data" assumes an unwarranted objectivity in Big Data. Etymologically, a datum is "something given."[92] But data are not given once and for all; rather, they are not only interpreted but *constructed* by the coding process and inherently symbolic nature of some underlying reality. The mind evokes new applications of data in each iteration, thus allowing the data to generate an indefinite number of interpretations of events. Although data are typically thought of as "objective," they are, like language and laws, subject to creative uses and interpretation in the same way metaphors are indeterminate in application.

1. Law as Metaphor and Language Game

A metaphor is far more profound than simply "saying one thing and meaning another."[93] Rather, the power of metaphors is in their rule-changing creativity, applying language in novel ways. The generative capacity of metaphors also "provides the foundation for the more general emergence of novelty in social and economic settings."[94]

Indeed, all interactions in society may be construed as Wittgensteinian "language games," which are the rules for how we think, talk, and act in different situations.[95] These rules, however, are not algorithmic. They exist in our habits, practices, and customs. We know how to follow the rules in an indeterminate range of situations, including, importantly, novel situations. Metaphors and language games share the protean quality of being applicable in novel ways or to novel situations while conforming perfectly to their defining rules. This common trait follows from a common element: language.

Law can also be characterized as a type of language game: "A new law, cause of action, human right must first be publicly named within a sentence pointing to new levels of action and recovery. Only then has society created for itself new law."[96] Moreover, there is a diversity of

---

92 *See Data (n.)*, ONLINE ETYMOLOGY DICTIONARY, at http://www.etymonline.com/index.php?term=data

93 James E. Murray, *Understanding Law as Metaphor*, 34 J. LEGAL EDUC. 714, 715 (1984).

94 Koppl et al., *supra* note 18, at 25.

95 *See* Roger Koppl & Richard Langlois, *Organizations and Language Games*, 5 J. OF MGM'T AND GOVERNANCE 287, 288 (2001).

96 Murray, *supra* note 93, at 716.

potential meanings that drives diverging theories of legal interpretation.[97] Language thus has a dual nature; while it may be used to construct systems of logic, its substance is indeterminate, contextual, and creatively evolving.[98] Law, being built from language, is partially logical but also partially open-ended and indeterminate.[99] The holdings of judicial opinions operate not like formulas or algorithms, but instead like partially-defined principles that operate at a more abstract level than the facts to which they are applied.

In this partially open-ended sphere, judges reason by analogy from past cases, identifying underlying principles that unify these cases and suggest a particular outcome.[100] Artificial technology, as Sunstein has explained, is only capable of retrieving cases and identifying analytical similarities and differences among them.[101] However, analogizers in law do not simply ask which case has "more" similarities to the case at hand, but whether a case has *relevant* similarities to the case at hand. And whether the similarities between cases are relevant "depends on the *principle* for which the initial case is said, on reflection, to stand."[102] Therefore, reasoning by analogy involves identifying principles that justify a claim that certain cases should be treated alike or differently.

Put another way, a judicial opinion or legal doctrine, like a metaphor, is a kind of theory-making device. Making law is much more like a search-light theory than a bucket theory. This protean quality of metaphor distinguishes it from mere analogy or simile. As Dworkin argues, "analogy without theory is blind. An analogy is a way of stating a con-

---

97 *See* Richard H. Fallon, *The Meaning of Legal "Meaning" and its Implications for Theories of Legal Interpretation*, 82 U. CHI. L. REV. 1236, 1239 (2015).

98 Even "definitions" of words are really metaphoric, in the sense that a leap of intuition is required to accept that a word is logically equivalent to its ostensible definition. *See* W.V.O. Quine, *Two Dogmas of Empiricism*, 60 PHIL. R. 20, 20-43 (1951). The famous metaphor "Juliet is the sun," for example, is a statement that is neither true nor false, yet its meaning is understood. Although metaphors are not themselves analytic, they provide the foundation for logical systems. From metaphors come propositions, or objects of belief, and from propositions come syllogisms: "All men are mortal, Socrates is a man, therefore Socrates is mortal." The syllogism is logically consistent, and yet metaphoric in the sense that each word represents a concept: to know whether the syllogism is "true," we must know what is man, what is mortal, what is Socrates. Logical reason, already propositional, cannot define them.

99 *See* Jan G. Deutsch, *Law as Metaphor: A Structural Analysis of Legal Process*, 66 GEO. L. J. 1339, 1346 (1977) (arguing that common law precedents "communicate interpersonally but cannot be reduced to objectively verifiable doctrinal formulae").

100 *See generally* EDWARD LEVI, AN INTRODUCTION TO LEGAL REASONING (1949).

101 Cass R. Sunstein, *Of Artificial Intelligence and Legal Reasoning*, 18 PUB. L. & LEGAL THEORY WORKING PAPERS 1, 6 (2001).

102 *Id.* at 5; *see also* Dworkin, *supra* note 71, at 355–56 (explaining that theory requires that judges "justify legal claims by showing that principles that support those claims also offer the best justification of more general legal practice in the doctrinal area in which the case arises").

clusion, not a way of reaching one, and theory must do the real work."[103] While a metaphor does not purport to predict all of its possible applications in advance, there is also no particular limit to the number and variety of iterations it can take.[104]

These iterations are not entirely open-ended, but instead depend on the evolution of prior precedents and the social context in which they are interpreted. There may be a "core of settled meaning," but there are also "debatable cases in which words are neither obviously applicable nor obviously ruled out."[105] Hart gives the example of a legal rule that forbids you to take a vehicle into the public park. He states, "plainly this forbids an automobile, but what about bicycles, roller skates, toy automobiles? What about airplanes? Are these to be called 'vehicles' for the purpose of the rule or not?"[106]

There is a spectrum of legal determinacy. Some precedents are more deterministic while others are open-ended in unsettled areas of law.[107] Some cases are sufficiently one-sided as to make the application of judgment practically trivial; other cases seem to present multiple contradictory resolutions that can equally satisfy the letter of the law.[108] Legal formalism's attempts to bring determinism to the law have failed to quash interpretative disputes over the meaning of statutory provisions, common law precedents, and constitutional texts.[109]

---

103 Dworkin, *supra* note 71, at 355–36.

104 *Id.* Judge Posner has rejected Dworkin's conception of constitutional theory, advocating instead that the legal academy work towards "fuller participation in the enterprise of social science, and by doing this make social science a better aid to judges' understanding of the social problems that get thrust at them in the form of constitutional issues." Richard Posner, *Against Constitutional Theory*, 73 N.Y.U. L. R. 1, 11–12 (1998). In Judge Posner's view, empirical questions, rather than theoretical ones, should play a larger role in the process of constitutional interpretation. *Id.*
While it is certainly important for judges to understand "the social realities behind the issues with which they grapple," *id.* at 13, the process of asking empirical questions (Posner asks, for example, "How influenced are judges in constitutional cases by public opinion?"), gathering data to answer those questions, interpreting the data, and assessing the importance of the answers gleaned, requires judgment and, implicitly, theory at a number of junctures. Thus, we contend that there is no way to sanitize lawmaking of "theory" based on purely "objective" social science, as even the most rigorously empirical approaches require "framing" of the problem studied.

105 H. L. A. Hart, *Positivism and the Separation of Law and Morals*, 71 HARV. L. REV. 593, 607 (1958).

106 *Id.*

107 *See id.*; *see also* Pierre J. Schag, *Rules and Standards*, 33 UCLA L. REV. 379, 385 (1985) (By describing the "distinction between permissible and impermissible conduct in evaluative terms," standards permit "individualized judgments about the substantive offensiveness or nonoffensiveness" of conduct.).

108 *See generally* Ronald Dworkin, *Hard Cases* in TAKING RIGHTS SERIOUSLY (1977).

109 Indeed, scholars have noted that the meaning of legal meaning, and commensurately the process of legal interpretation, is itself highly contested, with an astonishing diversity of theories. *See* Fallon, *supra* note 97, at 1305 (arguing that in light of the diversity of possible approaches to legal interpretation, judges should decide, on a "case-by-case basis" which out-

In navigating legal ambiguities, a judge must consider not only how a principle applies to a given case, but how it fits in with ever-broader swathes of legal doctrine, leading to higher levels of generality.[110] Further, as judges discover incongruences among legal principles or respond to unanticipated developments, they must be prepared "to reexamine some part of the structure from time to time."[111] A simple example is in the law of surrogacy. When the technology for surrogate motherhood arrived, the legal system was unable to decide who "the mother" of the child was. The law had to ramify the idea of motherhood to distinguish "genetic mother" from "birth mother"—a novel, unforeseeable and previously unnecessary distinction.[112]

No matter how settled the area of law, judgment is required in order to apply the law to the infinite factual variations that arise in the context of particular cases. For example, the legal standard regarding when traffic stops are constitutionally permissible under the Fourth Amendment is fairly settled,[113] and yet broad standards like "reasonable suspicion" accommodate an indefinite number of unforeseeable factual variations. A similar standard is the requirement of "probable cause" for arrest, search or seizure, which is "known" to judges abstractly though undefined in its particulars.[114] These types of "reasonableness" standards are ubiquitous in the law, and their purpose is to provide flexibility in applying broad legal standards to specific factual circumstances.

---

come would best promote "substantive desirability, consistency with rule of law principles, and promotion of political democracy, all things considered"); *but see* William Baude & Stephen Sachs, *The Law of Interpretation*, 130 HARV. L. REV. 1079, 1081 (forthcoming) (arguing that "legal interpretative rules are conceptually possible, normatively sensible, and actually part of our legal system").

110 *See generally* Dworkin, *supra* note 71.

111 *Id.* at 359–60; *see also id.* at 356–57 ("When we raise our eyes a bit from the particular cases that seem most on point immediately, and look at neighboring areas of the law, or maybe even raise our eyes quite a bit and look in general, . . . we may discover that [a] principle is inconsistent with . . . some other principle that we must rely on to justify some other and larger part of the law.").

112 See Pavel Kuchal, *The Birth of Surrogate Motherhood Law: An Economic Analysis of Institutional Reform*, RESEARCHGATE 1, 2–3 (2014) (concluding that changing beliefs about the legitimacy of surrogate motherhood determined the boundaries of institutional adaptation in the legal system).

113 *See Terry v. Ohio*, 392 U.S. 1, 33 (1968) (establishing the reasonable suspicion standard).

114 *See, e.g.*, *Illinois v. Gates*, 462 U.S. 213, 235 (1983) (describing probable cause as "a fluid concept-turning on the assessment of probabilities in particular factual contexts-not readily, or even usefully, reduced to a neat set of legal rules"); *Brinegar v. United States*, 338 U.S. 160, 175 (1949) ("In dealing with probable cause, however, as the very name implies, we deal with probabilities. These are not technical; they are the factual and practical considerations of everyday life on which reasonable and prudent men, not legal technicians, act.").

2. Data as Compression

This interpretative process can be starkly contrasted with a Big Data approach. Unlike a judge, Big Data cannot *decide* whether reasonable suspicion exists to support a *Terry* stop. At most, Big Data can *predict* the probability that a judge would find reasonable suspicion.

Big Data algorithms can be used to mine prior precedents or other relevant data for correlations between variables, such as by finding common factors in a judge's prior decisions that are predictive of future outcomes. In areas of law that are relatively settled, Big Data techniques could advance to a point where it is possible to predict how a case will be decided with a high degree of accuracy. But while Big Data may tell us *what* will be decided, it cannot tell us *why*. Nor is it likely to somehow computationally establish the actual truth or facts behind a case or trial.

Because Big Data is syntactic and deterministic, it assumes away the open texture of the law—ambiguous cases where the proper outcome is debatable. Big Data is inherently backward-looking; it can make predictions based on past decisions, but it cannot articulate novel possibilities. Moreover, the application of legal rules depends on determinations of meaning that require evaluative judgments.

Given Big Data's syntactic nature, there is no reason to believe that Big Data is capable of making these evaluative judgments.[115] Indeed, the algorithm-driven[116] view of data assumes not only that it is unnecessary to interpret data, but that doing so introduces an undesirable bias to our understanding of the objectively "real" world. Yet even the most sophisticated algorithms cannot define the criteria used to determine what is "optimal" or the set of alternative decisions or strategy spaces.[117]

As we have described, "data" as a concept is not "objective" but rather representational and theory-laden.[118] Put another way, data has no meaning without being "compressed" into a theory or shorter description as explained by the algorithmic information theory of mathematician Gregory Chaitin.[119] In Chaitin's metaphor, a theory is like a computer program whose output describes the system's behavior. The simpler the system is, the shorter the program can be. The point of theory, in this view, is to have a short description of the system.

115 Sunstein, *supra* note101, at 6.

116 An algorithm can be described as a "finite procedure, written in a fixed symbolic vocabulary, governed by precise instructions, moving in discrete steps, 1, 2, 3 . . ." Zia et al., *supra* note 9, at 92. Put more simply, "'algorithmic' means 'deterministic and predictable.'" Koppl et al, *supra* note 18, at 6.

117 Teppo Felin et al., *Economic Opportunity and Evolution: Beyond Landscapes and Bounded Rationality*. 8 STRATEGIC ENTREPRENEURSHIP JOURNAL 269, 269–82 (2014).

118 GITELMAN, *supra* note 70.

119 *See* GREGORY CHAITIN ET AL., GODEL'S WAY: EXPLOITS INTO AN UNDECIDABLE WORLD 62 (2011).

Chaitin argues that some systems are so complex that such compression is impossible. Simplifying the system is impossible, and thus explaining or identifying the system requires a lengthy description. In these complex cases, any compression that reduces description length will give a false and distorted picture of the system and its behavior.

Chaitin's notion of compression applies to Big Data. Indeed, the idea of compression is, at it were, baked into Big Data. The very idea is to gather vast amounts of data to a central location and extract from it salient summary descriptions by, essentially, identifying statistically significant correlations in the data. But in many instances, including the law, the underlying phenomena are too complex to allow meaningful compression of all that data. All compressions distort in such cases. Big Data must "compress" the underlying data into a series of discrete correlations or relationships, but these correlations necessarily oversimplify the true nature of the complex system.[120]

With Big Data, the more data that can be mined, the more reliable and specific the predictions that can be made, based on the analysis of past patterns in the data. But when Big Data is compressed into a theory, the high volume of data "demands that information be stripped of context and revealing ambiguity."[121] Correlations tell us nothing about underlying causal relationships, and it is well-known that correlations do not imply causation.[122] Big Data thus ignores the inherently semantic, context-dependent nature of data.

Without taking context into account, Big Data appears destined to produce spurious correlations.[123] Big Data's characteristically large data sets are especially vulnerable to spurious correlations because "large deviations are vastly more attributable to variance (or noise) than to information (or signal)."[124] Therefore, the "central issue" with Big Data "is that the needle comes in an increasingly larger haystack."[125] Big Data cannot tell us which correlations are relevant or important, and the more

---

120 *See id.*; Gregory Chaitin, *The Limits of Reason*, 294 SCI. AM. 74, 74–81 (2006).

121 DeDeo, *supra* note 23, at 3. In describing the follies of stripping data from the past of all context and using it to predict future behaviors, Nassim Taleb uses the parable of a Thanksgiving turkey. The turkey is well-fed and contented in the months leading up to the holiday, and assumes based on its past experiences that its human captors are well-intentioned and trustworthy. All of that changes, as we know (but the unsuspecting turkey does not), on the day of the feast. NASSIM TALEB, THE BLACK SWAN: THE IMPACT OF THE HIGHLY IMPROBABLE 40–42 (2d ed. 2010).

122 *See generally* C.S. Calude & G. Longo, *The Deluge of Spurious Correlations in Big Data*, FOUNDATIONS SCI. 1 (2015).

123 *See id.*

124 Nicholas Taleb, *Beware the Dangers of Big Data*, WIRED MAGAZINE, Feb. 8, 2013, https://www.wired.com/2013/02/big-data-means-big-errors-people/.

125 *Id.*

correlations we have, the more difficulty in discerning which correlations are relevant.

The supposed irrelevancy of causation is especially alien to the legal system, where attributions of responsibility are often premised on causal grounds.[126] Whether someone is liable to be punished or pay compensation often depends on whether that person has caused harm of the sort that the law seeks to avoid.[127] Not only must the events in question be identified in terms of the time, place, and persons involved, but a causal link between those events must "be specified in such a way as to show that it falls within the relevant legal categories," such as negligence and physical injury.[128] Law is often concerned with whether a causal link exists between actions and resulting harm, and also whether that causal link is relevant and material to the elements of causes of action, such as causation in fact and proximate cause.

Ultimately, Big Data leaves us in a paradox. On the one hand, we gather more data and evidence in order to gain an ostensibly more accurate and complete understanding of the phenomenon we seek to influence. On the other, the more data we have, the more we have to simplify it in order to gain any useful insights; "a theory has to be simpler than the data it explains, otherwise it does not explain anything."[129] But in compressing the data, we lose nuances of the data that are simply too complicated to be compressed into a theory. Such a procedure, as applied to law, could discard causal nuances and result in arbitrary decision-making.

We are reminded of Borges' "On Exactitude in Science."[130] Borges speaks of an Empire where the art of cartography had advanced to the point where "the map of a single Province occupied the entirety of a City, and the map of the Empire, the entirety of a Province." Eventually, cartography advanced to a point where a map's size "was that of the Empire, and . . . coincided point for point with it." As one might expect, the following generations were not so reverent of cartography and realized the obvious—that the vast, point-by-point map of the Empire was useless. The map was thus abandoned, leaving only "Tattered Ruins . . . inhabited by Animals and Beggars."

Like Borges' point-by-point map of the Empire, Big Data may provide too many data points to be useful. The data simply is not comprehensible without a theory to understand it. Compression of data operates

---

126 *See generally* Honore, *supra* note 15; J.L. Mnookin, *Atomism, Holism, and the Judicial Assessment of Evidence*, 60 UCLA L. REV. 1524 (2013).

127 *See* Honore, *supra* note 15.

128 *Id.*

129 Chaitin, *supra* note120, at 76.

130 JORGE LUIS BORGES, ON EXACTITUDE IN SCIENCE 1, 1 (1946) (fictitiously quoting from Suarez Miranda, *Viajes de Varones Prudentes*, Libro IV, Cap. XLV, Lerida (1658)).

the same way a metaphor works in language and law. The meaning of data, like a metaphor, is partially open-ended and indeterminate, unlike any one particular propositional compression or application of a metaphor, which is by definition specific. Thus, Big Data does little to bring us closer to objective decision-making than traditional legal methods.

3. Law in a Normative Universe

A temptation would be to use Big Data to attempt to devise or inform the "best" rules without acknowledging the implicit theories, compressions, or judgments that shape those rules. Such an attempt would overlook that our social and organizational lives are highly complex and conflicting, and would try to reduce policy decisions to data to, in essence, "take it out of our (biased) hands," make decisions neater and cleaner. Big Data, however, cannot choose between competing possible compressions or interpretations of law.[131] Nor can it resolve conflicting value judgments.

Efforts to define or optimize the law's purpose lead only to greater complexity. "Almost every area of law is filled with conflicting purpose. Tort and contract law both embrace efforts to achieve efficiency, but also fairness. Criminal law balances rights of the accused with society's broader need to control crime."[132] Moreover, legal rules may interact in complex and contradictory ways.[133]

Like a complex system with multiple equilibria or possible solutions, conflicting themes in the law make it impossible to "optimize" between several competing criteria. Indeed, the limitations to optimization have been explored in diverse areas of thought, including Kurt Gödel's incompleteness theorem in mathematics.[134] Similarly, Kenneth Arrow's impossibility theorem suggests that algorithms, like governments, cannot simultaneously accommodate multiple societal objectives

---

131 This raises a more general point about "personalized" law—law will be personalized not only for the consumer's benefit, but also to bolster the profits for the service providers. Currently, Big Data is used to customize Google searches or advertisements, so that more expensive products are targeted to wealthier consumers. Similarly, legal "solutions" will likely be targeted to consumers not solely based on the merits of their case, but to fulfill other objectives.

132 Rachlinski, *supra* note 41, at 918.

133 *See* Devins et al., *supra* note 7, at 665.

134 Gödel demonstrated that there is a kind of incompleteness in formal mathematics. The axioms of the system let you state true mathematical theorems that you cannot prove using just those axioms. In this sense, no axiom system can "capture" the full richness of mathematics. Gödel's original proof relied on a version of the liar's paradox: "This statement is false." The statement contains a contradiction: it can only be true if it is false, and it is therefore neither true nor false. See the discussion in Chaitin et al., *supra* note 120. *See also* Kurt Gödel,, *On Formally Undecidable Propositions of Principia Mathematica and Related Systems*, *in* FROM FREGE TO GÖDEL: A SOURCE BOOK IN MATHEMATICAL LOGIC 1879–1931, 596, (Jean van Heijenoort, Ed., 2002).

while adhering to basic principles of fairness.[135] Even if Big Data could somehow "choose" between competing objectives or preferences, it could not predict how these multiple competing objectives will shift and evolve over time.

The resolution of competing, conflicting purposes is central to the rule of law, which provides a framework for resolving "competing conceptions of the good" so that pluralistic societies may achieve "political cohesion with minimum oppression."[136] The open texture of law thus has a fundamentally moral component.[137] Although scholars have examined various factors that enable societies to successfully adopt the rule of law,[138] one essential factor is the law's production of *meaning* that is sufficiently accepted within society to constitute binding authority. Law is thus inescapably normative; it "becomes not merely a system of rules to be observed, but a world in which we live."[139]

For example, society universally values *fairness*, as embodied in Rawls' veil of ignorance—the designing of rules where "no one knows his place in society, his class position or social status; nor does he know his fortune in the distribution of natural assets and abilities, his intelligence and strength, and the like."[140] Yet the legal system's attempts to achieve this ideal are consistently revealed as illusory due to the multiple inconsistent possible interpretations of fairness in any given situation. What is considered a fair sentence for a criminal defendant who assaults without intent to kill, but the victim dies? Perhaps a fair sentence would reflect his lessened culpability, or perhaps it is only the consequence of his actions that matters. If a jury nullifies a defendant's conviction, is it a noble appeal to individual morals or a rejection of democratic values?[141]

---

135 While an extensive discussion of Arrow's theorem is beyond the scope of this paper, the theorem states, in essence, that when voters have three or more distinct options, no ranked order voting system can always convert the ranked preferences of voters into an aggregate, community-wide ranking while also satisfying a certain set of "fairness" criteria. In other words, Arrow articulated a serious limitation in the ability of representative governments to fulfill multiple alternative voter preferences. *See* Kenneth J. Arrow, *A Difficulty in the Concept of Social Welfare*, 58 J. POLI. ECON. 328 (1950).

136 Rosenfeld, *supra* note 13, at 1310.

137 *Id.*; *see also* Deutsch, *supra* note 99, at 1346 (noting that "precedents can be defined as constituting moral injunctions [which are] persuasive because of the factual descriptions from which they are devised").

138 *See generally* Fallon, *supra* note 1.

139 Robert M. Cover, *Foreword: Nomos and Narrative*, 97 HARV. L. REV. 4, 68 (1983).

140 JOHN RAWLS, A THEORY OF JUSTICE 118 (1999).

141 For a variety of views on this question, *see generally* Paul Butler, *Racially Based Jury Nullification: Black Power in the Criminal Justice System*, 105 YALE L. J. 677 (1995); Jack Weinstein, *Considering Jury Nullification: When May and Should a Jury Reject the Law to do Justice*, 30 AM. CRIM. L. REV. 239 (1992); Gary Simson, *Jury Nullification in the American System: A Skeptical View*, 54 TEXAS L. REV. 488 (1975).

Even such a universally-respected virtue as "fairness" is subject to conflicting interpretations, none of which can be objectively verified as "right" or "wrong." The law's efforts to achieve ideals of justice and fairness inevitably fall prey to the messy, conflicted process of resolving the competing narratives surrounding these ideals, especially where decision-making is democratic and there are multiple persons imagining what might go on behind the veil. Moreover, notions of fairness may change over time. A "one-time Rawlsian bargain behind the veil of ignorance is insufficient in a world where institutions themselves evolve beyond the intentions of the designers."[142] This messy process of resolving conflicting values is a feature, not a flaw, of the legal system, with its purpose to resolve conflicts through democratically legitimate processes rather than by violence.

In sum, we question the assumption that Big Data is "objective." Data is subject to an indefinite number of possible compressions and interpretations, and is thus inherently theory-laden. Moreover, because Big Data is syntactic and acontextual, it cannot interpret or decide legal questions, nor is it capable of resolving the competing, conflicting purposes of law.

III. LAW AND BIG DATA'S ILLUSORY PREDICTIVE POWER

Some might object that Big Data is not used to *make* legal judgments, only to *inform* those judgments. But this distinction is illusory. The acts of collecting and interpreting data, and of interpreting and applying law, are inherently theory-laden.

As we will discuss in this section, the acts of interpreting and applying data and law are also deeply affected by "creativity, newness, novelty, surprise, and ignorance."[143] The indeterminate nature of law and data allows them to be deployed in novel, creative ways that may not have been previously anticipated. By contrast, algorithms are formulaic; their "execution requires no insight, cleverness, intuition, intelligence, or perspicuity."[144] Without human intervention, Big Data cannot update its "frame" to account for novelty, and thus cannot account for the creatively evolving nature of law.

In this section, we describe the frame problem and argue that predictive analytics fails in the legal context, using the examples of risk assessment models in sentencing and financial regulation.

---

142 Devins et al., *supra* note 7, at 620.

143 Zia et al., *supra* note 9, at 95.

144 *Id.*

A. *The "Frame" Problem: Algorithms in a World of No Entailing "Laws"*

Big Data maps and analyzes systems using real data in real time. The idea is that, with the ability to analyze massive troves of data, the structure of the system or, perhaps, the "frame" of the problem being analyzed can be measured instead of being merely hypothesized. The Big Data paradigm thus views systems as governed by stable, predetermined entailing laws. The "possible paths the system can take are all predetermined before the dynamics unfold."[145]

Some of the more ambitious hopes for Big Data could be fulfilled only if all of the possible variables that may affect the development of a system were known and accounted for in the mathematical model. This approach has parallels in general equilibrium from economics, which views the economy as a bounded system in which "actors omnisciently calculate and compare all possible actions, including future ones."[146] Thus, standard economic theory is comparable to "a computer that has been programmed to execute a master set of equations,"[147] where the effects of a given policy can be calculated based on a set of hypothetical initial conditions.

But complex systems such as the legal system are not deterministic, let alone predictable. Rather, "the full range of relevant variables that may affect a system's development—or the frame of the system—cannot be ascertained either when the system is created or as it changes over time."[148] The very dimensions of the system are subject to unprestatable change, as creative agents adapt to unforeseen and unforeseeable possibilities and opportunities.[149]

Each law has a set of affordances, or possible uses and interpretations, which "adaptive actors, from lawyers to regulators to business and lay people, exploit in order to fulfill their own purposes, creating new systemic behaviors that may diverge radically from the underlying purposes behind the law." This full set of possible uses for a law cannot be anticipated in advance.[150]

The concept of affordances has parallels in the nature of metaphor. Laws and other data, like metaphors, have an indefinite, unforeseeable set of possible interpretations and uses, which evolve in order to adapt to

---

145 *Id.*

146 Koppl et al., *supra* note 18, at 7. For a more complete examination of competitive equilibrium, *see generally* Yves Balasko & John Geanakoplos, *Introduction to General Equilibrium*, 147 J. ECON. THEORY 400 (2012); Kenneth Arrow & Gerard Debreu, *Existence of an Equilibrium for a Competitive Economy*, 22 ECONOMETRICA 265 (1954).

147 Koppl et al., *supra* note 18, at 4.

148 Devins et al., *supra* note 7, at 622.

149 *Id.*

150 *Id.* at 624.

shifting social contexts. Moreover, affordances are highly contextual.[151] The affordances of a set of data, or of laws, are situated in the evolutionary agents' "frames" or mental models;[152] these agents are constantly updating their "frames" by finding new affordances that, in turn, expand the potential "phase space"—the space of pertinent observables and parameters, within which the system unfolds—in unforeseeable ways.[153] Learning implies novelty: What agents in the system learn—or might learn—is not listed or listable from previous phase spaces of the system.

Legal evolution can be characterized as an unending series of "frame" changes which alter the landscape of legal reasoning. Before the Civil War Amendments, for example, slavery was legally sanctioned and slaves were considered to be property protected under the Fifth Amendment.[154] Although the Civil War Amendments prohibited slavery, about thirty years later, the "separate but equal" doctrine permitted racial segregation in public facilities.[155] In 1954, the Supreme Court changed the legal landscape yet again by overruling the separate but equal doctrine and abolishing racial segregation in classrooms.[156] A Big Data approach could not anticipate these developments—it would merely reiterate and reinforce past doctrines, beginning from an arbitrary fixed point in time.

Yet, these frame changes are essential to the law's evolution because they allow the law to be adaptive, flexible, and innovative. We see this in private law such as contracts. In supply chains, there are many unknown issues that are resolved by being adaptive. For example, a supply chain may be "jury-rigged", meaning that supplies available for certain purposes are utilized in novel ways to solve new problems that arise. We cannot write a contract ahead of time for a jury-rigged supply chain. Rather, we leave open-ended contract terms, and then afterwards decide what is "fair." Courts act with similar flexibility when they deem contract terms to be ambiguous. Courts determine what the parties' reasonable expectations would have been by looking to past practices of parties, industry practice, and other external evidence to determine the parties' reasonable expectations.[157]

151 Hugo Letiche & Michael Lissack, *Making Room for Affordances,* 11 EMERGENCE: COMPLEXITY & ORG. 1, 1 (2009) (Affordances, like language games, represent "dynamic reciprocal relationships between animate persons and their environments. Affordances are in-between—their cognition is situated and contextual.").

152 *See* Arthur T. Denzau & Douglass C. North, *Shared Mental Models: Ideologies and Institutions,* 47 Kyklos 3, 10 (1994).

153 Koppl et al., *supra* note 18, at 4–5.

154 *See Dred Scott v. Sanford,* 60 U.S. 393, 453 (1857).

155 *See Plessy v. Ferguson,* 163 U.S. 537, 544–46 (1896).

156 *See Brown v. Board of Eduction,* 347 U.S. 483, 494 (1954)

157 *See generally* Richard A. Posner, *The Law and Economics of Contract Interpretation,* U. CHI. L. & ECON., Olin Working Paper No. 229 (2004) (analyzing economic tradeoffs in-

This open-ended environment produces a "combinatorial explosion" in the form of endless possible affordances which creative adaptive agents can exploit in novel ways, yielding a diverse and constantly-expanding range of new opportunities. In the economy, for example, "the Turing machine ultimately has yielded the World Wide Web, selling on the World Wide Web, Web browsers, and iPads"—none of which was prestatable in Turing's time.[158] Likewise, the list of traded goods has grown from a small handful when biologically modern man appeared to the multitude of goods traded in the modern global economy.[159]

In the American legal system, the length of statutes has exploded over time and administrative agencies have crafted reams of rules to govern the implementation of those statutes.[160] The number of laws and regulations "has exploded in tandem with the explosion of diversity in the economy—which not only enables the creation of new activities that could potentially be regulated, but also innovations in the method of regulation."[161]

The affordances of law thus act as enabling constraints that guide, without determining, the development of legal and economic systems. Laws do not merely constrain agents by defining required behaviors, but they also enable future innovations, such as regulatory arbitrage.[162] But because we cannot predict all potential actions a particular law will enable, we cannot entirely prevent the unintended consequences of laws. We have thus argued that legal institutions cannot be designed because "these institutions evolve over time and outstrip the original design to the point that the original institutional configurations may become unrecognizable."[163]

The co-evolution of agent frames and object affordances produces emergent phenomena that could not have been predicted from a prior knowledge of the system's frames and affordances. Thus, a "spontaneous order of law emerges from the innumerable interactions of judges, lawyers, policy makers, regulated entities and the society at large."[164] No single individual has complete knowledge or understanding of the

volved in application of principle doctrines of contract interpretation); E. Allan Farnsworth, *Meaning in the Law of Contracts*, 76 YALE L. J. 939 (1967).

158 Koppl et al., *supra* note 18, at 7.

159 *Id.*

160 Devins et al., *supra* note 7, at 653.

161 *Id.*

162 *See, e.g.*, Victor Fleischer, *Regulatory Arbitrage*, 89 TEX. L. REV. 227, 229 (2010) (describing theory of regulatory arbitrage and conditions that give rise to it); Bruce Yandle, *Baptists and Bootleggers: The Education of a Regulatory Economist*, 7 AEI J. ON GOV'T & SOC'Y REG. 12, 13–14 (1983) (describing how seemingly opposing parties, such as Baptists and bootleggers during the Prohibition era, can be bolstered by the same legal provisions).

163 Devins et al., *supra* note 7, at 623.

164 *Id.* at 624.

"evolving legal code and its sprawling enforcement mechanisms."[165] The implicit, evolving "frame" for a system thus emerges from the distributed decision-making of individuals.

Justice Sotomayor evocatively captured the emergent, evolving dynamics of law in her dissenting opinion in *Utah v. Strieff*.[166] In that criminal procedure case, the majority held that although an officer had unlawfully initiated a traffic stop of a suspect, the officer's subsequent discovery of a valid arrest warrant attenuated the unconstitutional taint of the unlawful stop. Therefore, the evidence he later seized while searching the suspect incident to arrest was admissible. In her dissenting opinion, Justice Sotomayor reasoned that the majority's holding would allow law enforcement to check innocent people for search warrants and use those warrants to excuse illegal traffic stops in order to allow the admission of evidence obtained incident to arrest.[167]

Importantly, Justice Sotomayor explained that law enforcement officers would not interpret the majority's reasoning in isolation. Her dissent illustrates how, in the aggregate, the Court's criminal procedure jurisprudence operates as a mosaic of affordances for law enforcement:

> The officer's control over you does not end with the stop. If the officer chooses, he may handcuff you and take you to jail for doing nothing more than speeding, jaywalking, or driving your pickup truck with your 3–year–old son and 5–year–old daughter without your seatbelt fastened. At the jail, he can fingerprint you, swab DNA from the inside of your mouth, and force you to shower with a delousing agent while you lift your tongue, hold out your arms, turn around, and lift your genitals. Even if you are innocent, you will now join the 65 million Americans with an arrest record and experience the "civil death" of discrimination by employers, landlords, and whoever else conducts a background check.[168]

Justice Sotomayor's dissent demonstrates the contrasting perspectives of judges and law enforcement officers. While judges decide cases incrementally, examining one issue at a time, law enforcement officers utilize courts' jurisprudence in an emergent, aggregated manner. Various legal holdings provide a diverse array of niches or tools for police to use in order to accomplish their objectives. These holdings amount to a

---

[165] *Id.*

[166] 136 S. Ct. 2056, 2064 (2016) (Sotomayor, J., dissenting).

[167] *Id.* at 2068–69.

[168] *Id.* at 2070 (internal citations and quotation marks omitted).

series of frame changes in the law, each of which provided a new adjacent possible niche for law enforcement, and none of which could be anticipated algorithmically.

Over time, small shifts in the frame may accumulate and become very large, emergent frame changes. In dispersed systems of decision-making, these legal entrepreneurs thus solve the frame problem by creating distributed, non-algorithmic frame adjustments. Yet, the algorithmic analysis of Big Data can only accommodate a single frame at a time.

B. *Due to the Frame Problem, Predictive Analytics Fails in Law*

By definition, an algorithmic system cannot change its frame of analysis. As Turing describes in "Lady Lovelace's objection," an "Analytical Engine [such as a computer] has no pretentions to *originate* anything. It can do *whatever we know how to order it* to perform."[169]

Turing himself rejected Lovelace's Objection, reasoning that there is nothing really new under the sun, and computers can be programmed to "learn" everything out there. But in a creative world, where the system must constantly adapt to unforeseen and unforeseeable changes that represent new opportunities in the adjacent possible, the system must be capable of "adaptive, non-algorithmic change."[170]

Programmers have not been able to solve the frame problem. In the field of robotics, programmers typically create meta-algorithms to make algorithms adaptive. Meta-algorithms are essentially adaptive learning systems that use analysis of past data to determine how algorithmic rules would change in response to novel or surprising decision problems.[171] This approach, however, is subject to the problem of infinite regress between levels of meta-algorithms. Each level of the meta-algorithm "would require some boundary conditions to be set for *framing* the novel decision problems at *lower* levels."[172] Therefore, layers of meta-algorithms, though somewhat adaptive, remain subject to the frame problem.

Just as we cannot design a legal system that reliably and consistently achieves certain objectives, like a clock we wind up and let go, we cannot reduce the legal system to a predictive, data-driven model. First, there is a measurement problem. It is impossible to know in advance precisely which variables, whether intrinsic or extrinsic, may have an impact on the legal system's development. Without knowing in advance

169 Alan Turing, *Computing Machinery and Intelligence*, 49 Mind 433, 447 (1950).

170 *Id.*

171 Zia et al., *supra* note 9, at 13.

172 *Id.*

which variables to measure, the increasing universe of data remains limited and is likely to be skewed by preconceived bias.[173]

This problem is analogous to the "multitask problem" in economics.[174] Any task has multiple dimensions, only some of which can be monitored at reasonable cost, or even at all. If decisions are made solely based on the measured dimensions, the unmeasured dimensions, whether important or not, will be neglected. The classic example is creating and teaching curriculum centered around test-taking when test scores determine teacher salaries.[175]

Big Data gives us something similar. We measure N dimensions of M, where N<M. If we make our decisions based on Big Data, we neglect the unmeasured dimensions without knowing whether those dimensions have a meaningful impact. Conversely, we may overvalue the dimensions we *do* measure simply because they are measurable, or place too much importance on correlations that are simply noise.

Further, there is a design problem. The Big Data paradigm assumes that we can passively measure the world and that it simply won't react to such measurement. In the real world, however, societal entities will react and adapt to decisions being made based on the measurements, and the process of measurement will influence the underlying data. The data thus becomes recursive and, accordingly, the system's evolution becomes self-reflexive and circular. Moreover, the set of possible interpretations and uses of the data is also unforeseeable. Those attempting to design predictive models for the law based on Big Data may thus encounter setbacks at both the measurement and design stages. The data they use to design the institutions will be inherently incomplete and will evolve in unforeseeable ways; and the institutions themselves will continue to evolve in unforeseeable ways.

1. Risk Assessment Tools in Sentencing

Notwithstanding these limitations, predictive analytics have made striking inroads in criminal law. This development has some precedent in the history of the Federal Sentencing Guidelines ("Guidelines"), which illustrates the tensions between algorithmic and discretionary modes of

---

173 This sort of bias is sometimes called the "ludic fallacy," or more simply, "unknown unknowns." TALEB, *supra* note121, at 127.

174 *See* Bengt Holmstrom & Paul Milgrom, *Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design*, 7 J. L., ECON. & ORG. 24, 50 (1991) (noting that performance measures may have the effect "aggregat[ing] highly disparate aspects of performance into a single number and omit[ting] other aspects of performance that are essential if the firm is to achieve its goals").

175 *See id.* at 25; *see generally* Jane Hannaway, *Higher Order Skills, Job Design, and Incentives: An Analysis and Proposal*, 29 AM. EDUC. RES. J. 3 (1992) (describing the multitask problem in education, particularly how emphasis on testing leads to curriculum imbalance).

legal decision-making. Congress created the Guidelines as part of a bipartisan effort to produce uniformity and certainty in sentencing.[176] The Guidelines were intended in part to alleviate policy makers' concern about unjustified variances in sentences nationwide.[177]

The resulting Guidelines provided a deterministic approach, more or less, to sentencing, which judges were required to follow.[178] The Guidelines are essentially a chart that produces a score or Total Offense Level for a particular convicted person based on their Base Offense Level and Criminal History, which each receive a score. The Offense Level has points added or subtracted for aggravating or mitigating circumstances, such as a supervisor role in a drug trafficking offense, which would increase the offense level, or acceptance of responsibility, which would decrease the offense level.[179] The system, as originally conceived, operated much like an algorithm; inputs formulated sentences.

While reducing judicial discretion, the Guidelines had the unintended consequence of increasing Congressional and executive power over sentencing policy, which in turn created a "one-way upward ratchet" in the severity of federal sentences.[180] For example, the Guidelines formed a niche that provided prosecutors with a range of tools to exercise power over criminal defendants. Because the Guidelines are complex, rigid, and heavily fact-dependent, they enabled prosecutors to determine many of the facts and significantly contribute to sentencing determinations before the case was even heard by a judge.[181]

Moreover, the deterministic approach of the Guidelines proved incompatible with the historically more open-ended, discretionary nature of sentencing. In particular, many judges viewed the binding nature of the Guidelines as unjust in a number of cases where the Guidelines required unduly harsh sentences without allowing discretion to account for relevant mitigating factors in the case.[182]

---

176 *See generally* Stephen F. Breyer, *The Federal Sentencing Guidelines and the Key Compromises on Which they Rest*, 17 HOFSTRA L. REV. 1, 8–25 (1988).

177 Frank O. Bowman, III, *The Failure of the Federal Sentencing Guidelines: A Structural Analysis*, 105 COLUM. L. REV. 1315, 1318 (2005); Jeffrey S. Parker & Michael K. Block, *The Limits of Federal Criminal Sentencing Policy; Or, Confessions of Two Reformed Reformers*, 9 GEO. MASON L. REV. 1001, 1006 (2001).

178 *See* Breyer, *supra* note176, at 24.

179 *See* USSG Sentencing Table (providing sentencing grid based on Offense Level and Criminal History Category).

180 Parker & Block, *supra* note177, at 1004 ("[F]ederal sentencing reform . . . taught Congress how to micro-manage the sentencing process and facilitated its interventions, by supplying a template for congressional dabbling in the details of the sentencing process").

181 Bowman, *supra* note 1777, at 1338–39.

182 *See, e.g.*, John Nochols, *Judge Resigns Over Congressional Meddling in Sentences*, THE NATION (June 25, 2003), https://www.thenation.com/article/judge-resigns-over-congressional-meddling/ (describing district judge's decision to resign due to "Congressional meddling" in federal sentencing decisions); *Criticizing Sentencing Rules, U.S. Judge Resigns*, N.Y.

In *Booker*[183] (and *Apprendi*[184] at the state level), the Supreme Court held the Guidelines to be advisory, thereby permitting the sentencing judge to tailor the sentence in light of other statutory factors.[185] The Court concluded that the structure of the Guidelines, which required a judge to increase the defendant's sentence beyond the statutory maximum if a preponderance of the evidence proved aggravating facts, violated the defendant's Sixth Amendment right to have a jury decide those facts.[186] As a remedy for this constitutional violation, the Court made the Guidelines system advisory. The effect of the Court's decision was to reassert the discretion of judges in criminal sentencing and provide greater institutional balance.[187]

The mandatory Federal Sentencing Guidelines illustrate the inadequacies of data-based, deterministic sentencing policy. The Guidelines attempt to reduce sentencing to an algorithmic model with predefined inputs and outputs. However, these mandatory Guidelines fail because just sentencing requires judicial discretion to consider the individual holistically, to weigh the competing purposes of sentencing, and to consider factors not accounted for by the Guidelines. In other words, the "frame" of sentencing determinations is fluid and requires case-by-case evaluations. The variables that were important in one sentencing proceeding may be less influential in another. These types of discretionary determinations are inherently not reducible to rigid criteria or models.

While federal sentencing generally is moving away from the mandatory Guidelines system to a more discretionary system, the use of predictive analytics is moving the system back towards a more deterministic system. At sentencing, judges are increasingly relying on risk assessment models to determine the likelihood of recidivism based on an offender's criminal history, personal characteristics, offense conduct, and other factors.[188] Using these models, judges decide which among a spec-

---

TIMES (Sept. 30, 1990), http://www.nytimes.com/1990/09/30/us/criticizing-sentencing-rules-us-judge-resigns.html?mcubz=1 (describing federal district Judge Lawrence Irving's resignation over the harshness of federal sentencing guidelines); *see also Eva S. Nilsen, Indecent Standards: The Case of U.S. versus Weldon Angelos*, 11 ROGER WILLIAMS UNIV. L. REV. 537, 544–45 (2006) (describing a case in which twenty-nine former federal judges and prosecutors filed amicus briefs requesting that a judge find mandatory minimum sentencing guidelines unconstitutional).

183 543 U.S. 220 (2005).

184 530 U.S. 466 (2000).

185 *Booker*, 543 U.S. at 245–46.

186 *Id.*

187 Susan Klein & Sandra Thompson, *DOJ's Attack on Federal Judicial "Leniency," The Supreme Court's Response, and the Future of Criminal Sentencing*, 44 TULSA L. REV. 519, 542 (2008) (arguing that *Booker* "brought to a halt the drive to shift the power to punish away from the judiciary and put it in the hands of federal prosecutors").

188 *See, e.g.*, State v. Loomis, 881 N.W.2d 749, 765–69 (Wis. 2016) (holding that although risk assessments cannot be determinative, a sentencing court may use them to, inter

trum of punishments, from probation to imprisonment, best responds to a particular offender's risk profile.[189] The objective is to employ "corrections practices that have been proven through scientific corrections research 'to work' to reduce offender recidivism."[190]

Lee Ellis provides an unusually vivid and worrying example of the dangers we should associate with science-based risk assessments in the criminal justice system. He says that biosocial approaches to crime assume that, "genetic factors contribute to criminality. Therefore, according to such biological theories, curtailing the reproduction rates of persons with 'crime-prone genes' relative to persons with few such genes should reduce a country's crime rates."[191] He explicitly labels this strategy a "Eugenic Approach[ ]" to crime fighting.[192] Noting that the use of anti-androgen drugs "is also called chemical castration," he further argues that "administering anti-androgens to young postpubertal males at high risk of offending, especially regarding violent offenses, should help to suppress the dramatic surge in testosterone in the years immediately following puberty. Males with the greatest difficulty learning may need to be maintained on anti-androgen treatment for as much as a decade."[193]

Further, these types of risk assessment models have been criticized as flawed, in both theory and practice, by former Attorney General Eric Holder, among others. Holder has questioned whether predictive models reinforce racial biases in the criminal justice system.[194] When judges consider factors such as poverty and lack of community resources to determine probability of recidivism, this raises the risk of condemning defendants for factors beyond their control, including group-level

---

alia, divert low-risk offenders from prison and impose terms and conditions of probation and supervised release); Pamela M. Casey et al., *Using Offender Risk and Needs Assessment Information at Sentencing: Guidance for Courts from a National Working Group*, NAT'L CTR. FOR ST. CTS., 1, 1 (2011).

189 *See* John Monahan & Jennifer Skeem, *Risk Assessment in Criminal Sentencing*, 12 ANNU. REV. CLIN. PSYCH. 489, 500–01 (2016); Redding, *supra* note 60, at 2 (arguing that "selecting the sentencing option(s) . . . is a scientific question that should be informed by the science of . . . "evidence-based practices").

190 Warren, *supra* note 32, at 20.

191 Lee Ellis, *Reducing Crime Evolutionarily*, *in* EVOLUTIONARY FORENSIC PSYCHOLOGY: DARWINIAN FOUNDATIONS OF CRIME AND LAW 249, 259 (Joshua D. Duntley & Todd K. Shackelford eds., 2008).

192 *Id.* at 258.

193 *Id.* at 255.

194 Eric Holder, Attorney General, Remarks at NAT'L ASS'N CRIM. DEF. LAW. (Aug. 1, 2014), http://www.justice.gov/iso/opa/ag/speeches/2014/ag-speech-140801.html; *see generally* Starr, *supra* note 39, *but see* Jennifer Skeem & Christopher T. Lowenkamp, *Risk, Race and Recidivism: Predictive Bias and Disparate Impact*, 54 CRIMINOLOGY 680 (2016) (finding that although risk assessment tools resulted in disparate results between blacks and whites, the difference was likely attributable to criminal history rather than racial bias).

characteristics,[195] rather than punishing them as individuals based on their offense conduct and characteristics.[196]

Ultimately, where algorithms do not provide causal accounts, the ethics of decision-making become opaque.[197] Although the Equal Protection Clause imposes heightened scrutiny on decision-making based on an individual's status in a protected category, such as race, religion, or gender[198], Big Data makes no such distinctions. In contexts such as sentencing, protected categories like race may be highly correlated with unprotected categories, socioeconomic status or past criminal history for example, and the algorithms are often so complex that it is impossible to determine which factors were important to the outcome.[199]

Although researchers have begun to develop statistical methods that might allow us to track which features of input data lead to a particular outcome, these methods will tell us only *how* certain variables were combined, not *why* they were combined in that manner.[200] We thus remain "uncertain of the causal mechanisms that led to this or that combination of variables being a good predictor of what we wish to know."[201] Moreover, the Big Data paradigm is also implicitly, yet highly, value-laden in the sentencing context. Predictive analytical models cannot take into account the multiple competing theories of sentencing law.[202] The use of predictive analytical models implicitly takes a strong utilitarian perspective on the purposes of sentencing, prioritizing public safety issues and

---

195 *See* Brian Netter, *Using Groups Statistics to Sentence Individual Criminals: An Ethical and Statistical Critique of the Virginia Risk Assessment Program*, 97 J CRIM. L. & CRIMINOLOGY 699, 728 (2007) (criticizing application of group-level statistics to predict individual behavior through risk assessment tools).

196 *See* State v. Loomis, 881 N.W.2d 749, 764–65 (Wis. 2016) (acknowledging that "risk assessment is about predicting group behavior, . . . not about prediction at the individual level[,]" but concluding that risk assessments may help judges arrive at an individualized sentence); Starr, *supra* note 39, at 842–50 (noting that evidence-based sentencing often results in individual sentences being based on group-level predictions).

197 *See* DeDeo, *supra* note 23, at 2; *see also* Melissa Hamilton, *Risk-Needs Assessment: Constitutional and Ethical Challenges*, 52 AM. CRIM. L. REV. 231, 242–61 (2015) (discussing constitutional questions regarding the use of risk assessment tools at sentencing).

198 Richmond v. J. A. Croson, 488 U.S. 469, 493–94 (1989) (racial discrimination); Craig v. Boren, 429 U.S. 190, 198–99 (1976) (gender discrimination). Lemon v. Kurtzman, 403 U.S. 602, 614–15 (1971) (calling for close scrutiny of government action entangled with religion).

199 *See, e.g.*, Kelly Hannah-Moffat, *Actuarial Sentencing: An "Unsettled" Proposition*, 30 JUST. Q. 270, 370–74 (2013) (discussing logical and methodological limitations with risk assessment tools).

200 *Id.*

201 DeDeo, *supra* note 23, at 5.

202 *See generally* Breyer, *supra* note 176, at 8–25 (describing compromises in sentencing guidelines to address competing goals of federal sentencing); *see also* Nagel et al., *Symposium: Equality Versus Discretion in Sentencing*, 26 AM. CRIM. L. REV. 1813, 1813 (1989) (noting that judges historically had great discretion "in the sources of information they used in imposing sentences within a range, based on their own mix considerations of desert, deterrence, rehabilitation, and incapacitation, and in deciding how to weigh each of these factors").

the likelihood of recidivism over other factors, such as the culpability of the offender, the need for proportionality among offenders, and rehabilitative concerns, among others.[203] Moreover, priorities in sentencing law can, and do, change over time.

Although judges could rely on the results of the predictive models as a mere starting point,[204] behavioral research indicates that judges tend to "anchor" their views of an appropriate sentence to benchmarks such as sentencing guidelines.[205] Therefore, it is likely that the models' "recommendations," even if not binding, would have a high level of influence on offenders' sentences. The criminal justice system should not acquiesce to risk assessment models, and the values implicitly embedded within them, simply because they seem to be more "objective" or "evidence-based." Rather, as with the Guidelines, we should question whether the models reflect the proper purposes of sentencing and whether they are effective at determining offender risk.[206] Ultimately, Big Data cannot resolve the debate between utilitarian and retributivist views of sentencing policy. At most, Big Data may be able to inform our judgments as to the effectiveness of those polices and whether they fulfill certain predetermined societal objectives.

2. Risk Assessment Models in Financial Regulation

The flaws of predictive models also can be seen in financial regulation. While financial systems are subject to risk, which is governed by probability distributions, they are also rife with uncertainty, which is in-

---

203 *See* 18 U.S.C. § 3553(a)(2010) (providing numerous factors for courts to consider in sentencing); *see also* John Monahan, *A Jurisprudence of Risk Assessment: Forecasting Harm Among Prisoners, Predators, and Patients*, 92 VA. L. REV. 391, 435 (2006) (arguing that sentencing should be focused on culpability and that risk tools that measure future risk should not be relevant to sentencing); Paul H. Robinson, *Punishing Dangerousness: Cloaking Preventive Detention as Criminal Justice*, 114 HARV. L. REV. 1429, 1440 (2001) (arguing that focusing on risk factors for future violence without indexing for blameworthiness is offensive to justice).

204 *See* State v. Loomis, 881 N.W.2d 749, 764 (Wis. 2016) (permitting use of risk assessment tools as "starting point" in sentencing).

205 *See* United States v. Ingram, 721 F.3d 35, 40 (2d Cir. 2013) (Calabresi, J., concurring) (arguing that "[t]he so-called 'anchoring effects' long described by cognitive scientists and behavioral economists show why the starting, guidelines-departure point matters, even when courts know they are not bound to that point." (citing Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 Sci. 1124 (1974))).

206 *See id.* Casey and Niblett argue that both consequentialists and non-consequentialists can use predictive technology. *See generally* Casey & Niblett, *supra* note 2. For example, lawmakers who want to prohibit certain behavior deemed immoral, regardless of the consequences, could identify examples of such behavior to feed into a machine. The machine can then use analytic and pattern recognition technology to determine whether other examples of behavior would be deemed immoral according to the lawmaker's standards. In our view, this scenario assumes away the very problem it identifies—that multiple possible optima exist to resolve legal problems, and that Big Data cannot decide among them.

herently incalculable.[207] In the face of uncertainty, we do not know the possibility space and cannot assign probabilities. Consequently, adaptive agents in the economy are driven to solve problems through experimentation with different methods until stumbling across an effective one.[208] The more complex the system, the number of adjacent possible niches expands and, correspondingly, the space of possibilities for experimentation, making the system even less predictable.

For these reasons, we are skeptical of efforts to base legal and regulatory decisions on predictive analytical models that extrapolate from past data to predict future trends. For example, risk-based regulation, which regulates activities based on their perceived risk,[209] has been prominently featured in financial regulation but has also exhibited some prominent failures.[210]

Both Dodd-Frank and the Basel III agreement adopted enhanced regulations for Systemically Important Financial Institutions (SIFIs). These regulations include heightened capital requirements, leverage limits, liquidity requirements, monitoring and supervision, and regular stress testing to ensure that the SIFIs can withstand financial crises.

The risk-assessment models used in connection with Dodd-Frank and Basel III, however, have been criticized for their reliance on dubious assumptions such as stability of credit and determinability of asset prices during a crisis—assumptions loosely analogous to the deterministic "phase space" in physics.[211] Models based on overly optimistic or deterministic assumptions based on past trends can create complacency and incentivize further risk-taking, in contradiction to the model's purpose to assist in minimizing risk. Even if these erroneous assumptions were corrected, econometric "VAR" models cannot account for "Black swans" or high-impact yet rare events, Knightian uncertainty, and unquantifiable

---

207 *See generally* FRANK KNIGHT, RISK, UNCERTAINTY, AND PROFIT (1921) (distinguishing between risk, which can be calculated, and uncertainty, which cannot). Taleb calls Knight's distinction "artificial," since so-called "computable risks" are "largely absent from real life." TALEB, *supra* note 121, at 128. Rather, "[i]n real life you do not know the odds; you need to discover them, and the sources of uncertainty are not defined." *Id.* at 127.

208 *See infra* Part IV.

209 *See* Philip Maymin & Zakhar Maymin, *Any Regulation of Risk Increases Risk*, 15 FIN. MKT. & PORTFOLIO MGMT. 299, 304–08 (2012) (discussing the paradoxical effects of regulations intended to reduce risk on the market).

210 The financial crisis itself exposed the vulnerability of predictive models, as Taleb painstakingly described in The Black Swan. *See generally* TALEB, *supra* note 120. As a case in point, in 2004, economist Ben Bernanke declared that modern economics had diminished macroeconomic volatility and rendered economic crises obsolete, a phenomenon he termed the "Great Moderation." *See* Ben S. Bernanke, Governor, FED. RES. BOARD. at the EASTERN ECON. ASS'N. (Feb. 20, 2004). Less than four years later, a financial crisis nearly plunged the world into a deflationary depression.

211 *See* Lawrence Baxter, *Betting Big: Value, Caution and Accountability in an Era of Large Banks and Complex Finance*, 31 REV. BANKING & FIN. L. 765, 843–44 (2011).

yet relevant factors like management quality.[212] In short, we cannot address these problems simply by making "better" models.[213]

Further, these models highlight contradictory motivations in banking regulation. Policymakers recognize that banks generally must be allowed to fail, but it is equally certain that some banks are too big to fail, and therefore, must be stringently regulated to avoid the systemic consequences of failure. As a society, we must decide whether banks are too big to fail, and thus should be treated like public utilities, or whether banks are free market entities subject to the forces of creative destruction. This is a question that predictive models cannot answer.

Finally, econometric models cannot pre-state, or account for, economic innovations that drastically alter the economic and financial landscape. Such innovations range from the then-novel box stores several decades ago that often drove out small retailers, to the mushrooming of derivatives that were intended to hedge financial risk but inadvertently underwrote the crash of 2008. In fact, "Big Data" itself is just such a transformative innovation, unprestatable before Turing and Von Neumann invented the digital computer.

In sum, Big Data is radically incomplete and its potential predictive value is thus limited. While we recognize that Big Data can be effective in analyzing systems where habits are deeply entrenched and past behavior is highly predictive of future trends, social systems such as the legal system rarely act so predictably, especially when larger time scales are considered. Rather, the legal system is at least partially indeterminate, and it is impossible to foresee how legal doctrines may evolve over time away from their original premises in order to suit new purposes.

IV. LAW GONE VIRAL—BIG DATA'S UNINTENDED CONSEQUENCES

Our position on Big Data may seem overly pessimistic in light of the magnificent capabilities already demonstrated by predictive technologies. For example, machines outperform humans in many areas of life, including predicting consumers' taste, advising clients on financial opportunities, and predicting the likelihood of cancer.[214] As the capacity of computers to collect, store, and process data continues to increase, these capabilities will likely grow exponentially.

While we do not question Big Data's formidable capabilities, we have argued that these capabilities are inherently limited. In this section,

---

212 *Id.* at 842–45.

213 *Id.* Many of these issues are discussed in Koppl's *From Crisis to Confidence: Macroeconomics after the Crash*, which outlines his interpretation of the crisis, including the role of both the rating agencies and monetary authorities. ROGER KOPPL, FROM CRISIS TO CONFIDENCE: MACROECONOMICS AFTER THE CRASH (2014).

214 Niblett & Casey, *supra* note 2, at 29.

we will argue that these limitations have the dangerous potential to generate perverse unintended consequences for the legal system by undermining the evolution of the law and concentrating political power in undemocratic ways.

Big Data is not a passive instrument but rather, an active (and entirely opaque) participant in our social lives. Take, for example, the recent controversy over Facebook's News Feeds. Conservatives accused Facebook editors of vetting the "trending topics" section on Facebook in order to highlight stories with a liberal bias. Facebook defended itself by explaining that the process was "neutral" because the stories were "surfaced by an algorithm."[215]

The claim that algorithms cannot be biased paints a deceptively naïve portrait of algorithms. While teams of liberal programmers did not actively conspire to rig the News Feeds, it cannot be said that the algorithms were "neutral"—they were, at a minimum, designed to optimize outputs to parameters chosen by the company, such as increased comments and sharing of news stories. Thus, without being biased, the optimization instructions for the algorithms produced a skewed result. Despite Facebook's claims that its algorithms were entirely neutral, the company later announced a change in policy that would limit posts in News Feeds from other Facebook pages, instead prioritizing posts from users' families and friends—a change in the algorithms that will also undoubtedly impact the resulting user content.[216]

The danger of Big Data lies in its complex and opaque effects on the systems it is used to analyze. Big Data does not merely "predict" behaviors, it also influences those behaviors, and in the process, affects distributions of power. Big Data, as with social media, is used to discover pre-existing interconnected distributed networks—whether financial, social, health, or others—and to connect new networks. At the same time, the data from these distributed networks is compiled in concentrated servers. Whoever controls the server has the power to control how the data is collected and interpreted, and ultimately has the power to influence the system being measured. Thus, Big Data is at once centralized and highly distributed, depending on whether it is viewed from the perspective of the network or the perspective of the server.

---

[215] *See* Zeynep Tufekci, *The Real Bias is Built In at Facebook*, N.Y. TIMES OPINION, May 19, 2016, https://www.nytimes.com/2016/05/19/opinion/the-real-bias-built-in-at-facebook.html.

[216] *See* Parul Guliani, *Facebook's New Algorithm Change is a Blessing in Disguise*, FORBES, June 30, 2016, https://www.forbes.com/sites/parulguliani/2016/06/30/why-facebooks-new-algorithm-change-is-a-blessing-in-disguise/#533a4e6b5d68.

A. *Big Data and the Knowledge Problem: The Stifling of Legal Evolution*

The Big Data paradigm would take the opposite approach of common law systems, as common law systems recognize, at least implicitly, the role of incremental learning and evolution in the law. Common law principles are informed by social conventions and practices that arise organically from the learned experiences of individuals and organizations. In his advocacy for flexible legal standards such as reasonableness in tort law, American jurist Oliver Wendell Holmes Jr. emphasized that "[t]he law embodies the story of a nation's development through many centuries, and it cannot be dealt with as if it contained only the axioms and corollaries of a book of mathematics."[217] The role of unintended consequences in the development of social institutions was central to Holmes' jurisprudence.[218]

Although the common law system, and judicial review in particular, have been criticized as anti-democratic and slow to adapt to changes in public opinion (or quick to inspire public backlash),[219] the wisdom of the common law system lies in its retention of learned wisdom and collective societal knowledge.[220] Hayek argued that ignorance drives individuals to solve problems through experimentation, trying different methods until stumbling across an effective one.[221] The ultimate correct answer may prove wildly different from the answer originally presumed, and the gap between the two could only be closed through the blind process of creative experimentation. Human progress is achieved, therefore, largely through experimentation and innovation, not through predetermined designs.

A legal system based on "Big Data" would thus undermine the main source of legal innovation in our common law system—evolution of the law via judicial decision-making on a case by case basis.[222] Indeed, it is

217 OLIVER W. HOLMES, THE COMMON LAW 1 (Mark D. Howe ed., 1881).

218 Oliver Wendell Holmes, *The Path of the Law*, 10 HARV. L. REV. 457, 478 (1897).

219 *See, e.g.*, Jeremy Waldron, *The Core of the Case Against Judicial Review*, 115 YALE L.J. 1346 (2006); Reva Siegel & Robert Post, *Roe Rage: Democratic Constitutionalism and Backlash*, 42 HARV. C.R.-C.L. L. REV. 373 (2007).

220 For example, David Hume argued that the formation of the common law justice system was a form of spontaneous order, a logical outgrowth of the community's mutual recognition of human irrationality. Since people tend to be prejudiced towards short term interests, they developed independent judicial institutions to serve as neutral enforcers of long term principles, stability and certainty (1739: bk. III, sec VII 534–39). Common law principles, such as private property rights, are informed by social conventions that arise organically from the learned experiences of societies (Id.: 490).

221 *C. f.* F. A. Hayek, *The Use of Knowledge in Society*, 35 AM. ECON. REV. 519, 521–24 (1945) (discussing different forms of knowledge, and stressing the importance of knowledge particularized to the circumstances in decision-making).

222 Of course, Big Data and technology more generally are not the only source of the decline in judicial adjudication of disputes. The increase in settlements in both the civil and

essential that the law "sprawls" unprestatably as new meanings, sometimes even new loopholes, are found in the law that enable new patterns of action and risk or reward, that may then call forth new law. This sprawling *is* the evolution of the law. It is visible in British common law and its temporal evolution. New practices such as jury nullification, previously unforeseeable, arise. These practices confound the rigidity of Big Data, analyzing the past, and hinder the wise evolution of the law.

The evolutionary wisdom of common law has been resoundingly affirmed by behavioral research on the process of search and invention. In the context of technological invention, Maggitti et al. have found that the "search and discovery process . . . is inherently complex: nonlinear and disjointed rather than linear and cumulative."[223] Maggitti et al.'s work builds on Kauffman's modeling of evolution via exploration on NK landscapes.[224] An NK fitness landscape is an evolutionary model that models the interactions of adaptive agents and the evolutionary environment.

The NK landscape may be analogized to the common law system, where decision-makers such as litigants, lawyers, and judges adapt their litigation strategies and judicial doctrines, respectively, as the "landscape" of precedents and novel factual situations shifts. The NK landscape contains a number of individual components in the system and the level of interaction between the components—represented by the "N" and "K" variables. Adaptive agents are programmed to take "random walks" across the landscape, simulating the blind process of evolution. In the process, the agents encounter peaks and valleys in the landscape, which represent the different possible fitness levels that can be achieved. If the agents wander too far down into a valley, however, they are selected out of the landscape because their fitness is too low to survive. Fitness is thus achieved through experimental, trial-and-error traverses over the peaks and valleys.

The process of searching for novel solutions is deeply contextual. Inventors employ routines "including casting for information and re-categorizing that information."[225] This categorization process occurs "at the intersection of existing, yet seemingly disparate, landscapes that require

---

criminal context, as well as widespread use of alternative dispute resolution, have contributed as well.

[223] Patrick G. Maggitti et al., *The Complex Search Process of Invention*, 42 RES. POL'Y 90, 90 (2013) (explaining that "the search and discovery process of technological invention is inherently complex: nonlinear and disjointed rather than linear and cumulative").

[224] Stuart Kauffman & Simon Levin, *Towards a General Theory of Adaptive Walks on Rugged Landscapes*, 128 J. THEORETICAL BIOLOGY 11, 29–32 (1987); For further background on NK modeling *see generally* STUART KAUFFMAN, AT HOME IN THE UNIVERSE (1995); Lee Altenberg, *NK Fitness Landscapes*, *in* HANDBOOK OF EVOLUTIONARY COMPUTATION (eds. Thomas Bäck et al., 1997).

[225] Maggitti et al., *supra* note222, at 91.

[inventors] to manage interdependencies and react to a complex system of continuously changing internal and external factors."[226] Similarly, judges navigate a complex, sometimes conflicting system of precedents in order to apply judicial rules to novel fact patterns.[227]

In making decisions, agents categorize information based on their memories, which comprise a repertoire of past situations. The agent draws on memories of past situations and perceived similarities between different situations.[228] Similarly, judges rely on past legal decisions in order to reason by analogy in future cases.[229]

By contrast, the information collected for the Big Data models is limited in scope to "hard" information and excludes tacit and soft information, thus ensuring that only a fraction of the relevant knowledge is actually collected. In complex systems, individuals' use of knowledge is guided by intuitive rules that emerge through natural selection; rules that lead to successful social order supplant the rules least adapted to the prevailing social environment.[230] These rules constitute political, cultural, and other social traditions accumulated over time which become intuitive to social relations without conscious imposition or design.[231] A rational reconstruction of the system of social rules in its entirety is beyond the grasp of individuals.[232] Further, the applications of the rules often cannot be articulated ahead of time, and only become evident within specific contexts.

Knowledge changes over time in a process of variation, selection, and retention. This evolution produces a stream of novelty in human knowledge.[233] An essential part of the problem with Big Data is translating knowledge that exists tacitly as habit and judgment into some sort of coded form. Big Data gives us the illusion of knowledge we do not truly have. If change is non-algorithmic, as we have argued, then Big Data would seem to be a conservative force resisting salutary change.

Moreover, in the real world, evolution occurs not via singular agents, but through the cooperation of interconnected agents sharing distributed knowledge. In other words, knowledge is synecological, meaning that "knowing" units are often not individuals but rather a collection of interacting individuals. This idea harkens back to Adam Smith's no-

---

226 *Id.*

227 *See supra* Part II.B.1.

228 *See* Giovanni Gavetti & Massimo Warglien, *A Model of Collective Interpretation*, 26 ORG. SCI. 1263, 1269 (2015).

229 *See* Sunstein, *supra* note 101, at 5.

230 F. A. HAYEK, *The Theory of Complex Phenomena*, *in* STUDIES IN PHILOSOPHY, POLITICS, AND ECONOMICS 22, 33 (1967).

231 *Id.* at 28–9.

232 *See generally* F.A. HAYEK, THE CONSTITUTION OF LIBERTY (1976).

233 Koppl et al., *supra* note 18.

tions of economic specialization and diversification as contributing to increased societal wealth.[234] As Leonard Reed taught us, no one person knows how to make a pencil. The pencil-making knowledge exists in the system as a whole.[235]

"Interpretation" of reality is thus often a collective endeavor resulting from interactions among individuals. In a system of distributed knowledge, different members contribute unique knowledge and the system's functionality results from "the patterns of interconnections among the system's elements."[236] Through interactions and information-sharing, otherwise partially ignorant actors can accurately interpret complex situations. In other words, "well-functioning collectives can be reliably effective in contexts that are so challenging that individual agents alone would likely make interpretative errors."[237]

This approach is corroborated by a study by Eppstein et al. comparing the effectiveness of randomized clinical trials in medicine to that of an alternative approach called "team learning," in which teams of care providers exchange experiences and information and discuss how to optimize treatment protocol without use of a formal RCT study.[238] The study concluded that the alternate methods performed differently depending on the complexity of the problem. In simple problems with independent causal factors, RCTs slightly outperformed team learning.

As interconnectedness of causal factors increased, however, meaning that outcomes tended to be based on multi-causality, team learning outperformed RCTs. Moreover, the greater the multi-causality of factors that influenced a particular condition, the better the relative performance of team learning.[239] Eppstein et al.'s work suggests that a team learning approach may be superior to analytical, centralized ones in solving problems involving complex systems with multi-causal pathways.

In short, evolutionary, innovative processes such as judicial decision-making are not based on logical reasoning within a predetermined frame. Rather, decision-makers expand and adapt frames in order to discern new opportunities and possibilities within the adjacent possible. These blind experimentation and search processes are inherently non-algorithmic and thus in many ways antithetical to a Big Data approach.

Whereas common law sees the landscape of jurisprudence as constantly evolving and expanding to accommodate new situations, Big Data

---

[234] *See generally* ADAM SMITH, AN INQUIRY INTO THE NATURE AND CAUSES OF THE WEALTH OF NATIONS (Edwin Cannan ed., The Modern Library 1937).

[235] *See* LEONARD READ, I, PENCIL 10 (1958).

[236] Gavetti & Warglien, *supra* note 228, at 1265.

[237] *Id.* at 1264.

[238] Margaret J. Eppstein et al., *Searching the Clinical Fitness Landscape*, 7 PLOS ONE 1, 2–3 (2012).

[239] *Id.* at 3

sees the legal landscape as static and deterministic. Whereas common law is contextual and associative, relying on precedents and analogies, Big Data relies on "serial, systematic search[es]"[240] of information and reduces complex relationships to simple correlations devoid of temporal context or systemic, interconnected relationships. Further, Big Data is inherently centralized and relies on a server; its top-down approach is the opposite of distributed knowledge and team-learning that characterizes evolution in the real world.

The centralized nature of Big Data puts it in conflict with the rule of law. As Fallon notes, "the Rule of Law should allow people to plan their affairs with reasonable confidence that they can know in advance the legal consequences of various actions."[241] With Big Data, decisions such as personalized default rules for inheritance are made in a center that relies on information gathered up from many sources. Each person has only a sliver of that big data set. Lacking the information behind the data-driven choices made in the center, individuals cannot have "reasonable confidence" about the "legal consequences of various actions."[242] Simple default rules such as equal division of estates can be known and understood. Opaque algorithms generating personalized "rules" are unknowable and unpredictable. Thus, one of the things often touted as the great strength of Big Data, personalized law, abrogates the rule of law and subjects the people to needlessly heightened uncertainty.

The absurd phrases "personalized law" and "personalized rule" are oxymorons, at least within the context of the rule of law as articulated by Dicey, Fallon and others.[243] Dicey says the rule of law, "means, in the first place, the absolute supremacy or predominance of regular law as opposed to the influence of arbitrary power, and excludes the existence of arbitrariness, of prerogative, or even of wide discretionary authority on the part of the government"[244] The rules of "regular law" are simple and uniform across persons. To say that we are going to have "personalized law" is precisely to say that we are not going to have regular law. It is thus to say that we are not going to have the rule of law. In this sense, we may say that some of the highest ambitions for Big Data in the legal system are schemes to abandon law in favor of an impenetrable host of idiosyncratic directives issued from an all-powerful center.

---

240 Gavetti & Warglien, *supra* note 228, at 1267.

241 Fallon, *supra* note 1, at 7–8.

242 *Id.*

243 A. V. DICEY, INTRODUCTION TO THE STUDY OF THE LAW OF THE CONSTITUTION 107 (8 ed. 1982).

244 *Id.* at 198.

B. *Network Effects, Herding Effects, Big Data's Centralization of Power*

Although Big Data's promise lies in "customized" law, Big Data threatens to replace the evolutionary, diverse common law system with a sort of one-size-fits all system where the tacit "frame" of the problem is identical in all situations and innovations in the frame are inadvertently discouraged. This approach would not only suppress legal evolution, it would also centralize political power in opaque and undemocratic ways.

Because algorithms are not semantic, they cannot perceive affordances. The greater the combinatorial inflation of various affordances, the greater the computational complexity, to "a point where algorithms can end up running in perpetual loops."[245] The law could thus become self-reflexive and recursive.

Big Data functions by taking advantage of network effects, or feedback cycles that can make a network ever more influential or valuable.[246] Network effects can be rewarding or punishing; a network gains prominence through rewarding effects and maintains dominance through punishing network effects that "lock-in" the network's users to prevent them from leaving.[247]

A classic example is how Facebook grew massively through rewarding network effects; the more people joined, the more the value of the network grew. Once a certain threshold was passed, Facebook became such a prevalent tool for social networking that negative network effects deterred people from leaving—by leaving, they would lose a vital source of social connection. Correspondingly, a winner-take-all system arose in which competing social network profile websites such as MySpace and Friendster were ousted. These network effects can be quantified, as for example in Google's use of algorithms to summarize collective opinions about the rank-order value of webpages.[248]

Big Data intensifies these network effects because it uses powerful servers to gather extensive amounts of data into a single pool and analyzes and impacts these networks have within that single pool. However,

[245] Zia et al., *supra* note 9, at 97.

[246] *See* Paul Ormerod, *Hayek, The Intellectuals and Socialism, and Weighted Scale-Free Networks*, 26 ECON. AFF. 41, 43 (2006) (explaining that in a scale-free network, "[a] small number of individuals are connected to large numbers of others, but most people connect only to a small number of others"); *see generally* Werner Raub & Jeroen Weesie, *Reputation and Efficiency in Social Interactions: An Example of Network Effects*, 96 AM. J. SOC. 626, 626 (1990) (describing effects of social networks on reputation).

[247] *See generally* Joseph Farrell & Paul Klemperer, *Coordination and Lock-In: Competition with Switching Costs and Network Effects*, *in* 3 HANDBOOK OF INDUSTRIAL ORGANIZATION 1967–2072 (2007) (ebook) (explaining commercial effects of lock-in); *see also* Ormerod, *supra* note 246, at 43 (using network theory to explain why the views of a small but influential minority can come to prevail as the orthodoxy).

[248] *See* DeDeo, *supra* note 23.

the predictive analytical models rely on algorithms that are, by definition, deterministic. The algorithm cannot update its frame or perceive new affordances but can only implement the judgments or "weights" assigned by the data. Through machine learning, the algorithms update these "weights" based on their observations of changes in the underlying data.[249] This is how the algorithms "learn" solutions to complex problems, particularly those for which there are many potential solutions. Importantly, the weights or values assigned to the data do not change based on any frame adjustments that are made by the algorithm. Rather, the weights change as the data changes based on the algorithm's *conformance* to the initial weights or frame programmed within the algorithm to measure fitness or optimization.

Moreover, changes in the data observed in the model are not always a result of actual observable changes in the underlying conditions of the system being studied, but may be actually be generated by the model's previous outputs (and estimations). The model creates a new set of observations from the data, which in turn influences the inputs that the model generates. Ultimately, this process creates a feedback loop where the model's inputs and outputs influence each other, and the correlations between the inputs and the outputs grow stronger with each iteration. This creates a convergence of inputs and outputs, resulting in network effects where a small number of nodes has outsized influence on the rest of the system.

Stifling innovation will promote legal uniformity and discourage heterogeneity. Although uniformity can be a virtue in law, there are situations where reduced diversity in decision-making can generate adverse unintended consequences for the legal system. As Ayres and Mitts explain, excessive behavioral uniformity increases systemic risk, as diversity is essential for adaptability or the ability "to respond to unexpected shocks and changes in environmental conditions."[250] Moreover, diversity allows for the system to learn more effectively within a greater number of alternative states of the world.[251]

Uniformity of rules is generally a good thing. Uniformity of thought and knowledge is generally a bad thing. In decentralized systems such as traditional common law, the rules are broad and general, the

---

249 *See generally* Zhi-Hua Zhou et al., *Ensembling Neural Networks: Many Could Be Better Than All*, 137 ARTIFICIAL INTELLIGENCE 239, (2002) (describing the use of genetic algorithms that evolve randomly assigned "weights" in order to assess the fitness of a Yes-network); Thomas G. Dietterich, *Ensemble Methods in Machine Learning*, *in* 1857 LECTURE NOTES IN COMPUTER SCIENCE 1, 1 (J. Kittler and F. Roli eds., 2002) (describing algorithm learning methods).

250 Ian Ayres & Joshua Mitts, *Anti-Herding Regulation*, 5 HARV. BUS. L. REV. 1, 14 (2015).?

251 *Id.* at 17.

knowledge driving the system's evolution is dispersed, heterogeneous, and particular. With Big Data, the rules are inherently likely to become centralized, uniform, and general. Big Data's promotion of legal uniformity could thus "deprive[ ] society of the knowledge regarding payoffs conditional on the future or even differing circumstances in the present."[252] Ultimately, Big Data risks generating herding behaviors, where individuals and markets follow the "crowd" by chasing each others' behavior rather than making rational choices.[253] In that scenario, legal evolution would devolve into a Keynesian "beauty contest."[254]

Big Data servers essentially recreate the calculation problem that plagues centralized economic planning by re-enforcing "habits of control" through top-down, common-and-control governance.[255] Successful legal evolution requires that the paradigm governing legal issues—or the "frame" through which these issues are viewed—shifts as the legal system itself evolves. This requires distributed decision-making among a variety of diverse actors who are able to communicate, and capitalize on, fragmented knowledge. By contrast, Big Data means that information is collected from, and inferences drawn from, a centralized pool of data. Data mining becomes a virtue and not a vice. Because the data is centralized, decision-making based on the data is also centralized. Therefore, Big Data exacerbates the problem of centralized planning in law.

Essential to the "calculation problem" argument from the start has been the assumption that the economy and legal systems are complex adaptive systems that respond to change. If we could magically whisk away dynamic change in the "underlying data," however, calculation would no longer be a hindrance to centralized planning.

Long before Big Data, the economist Oskar Lange anticipated that technology could advance to a point where computers could program the economy.[256] In 2012, the economist Glen Weyl echoed Lange, arguing that "it is increasingly hard to see how dispersed information poses the challenge it once did to centralized planning."[257] Big Data will not alleviate the problems caused by centralized planning, however, because

---

252 *Id.* at 22.

253 This problem has parallels in the actions of Big Players generally speaking, which tend to encourage herding and irrational bubbles in financial markets, for example. *See* Roger Koppl & Leland B. Yeager, *Big Players and Herding in Asset Markets: The Case of the Russian Ruble*, 33 EXPLORATIONS IN ECON. HIST. 367, 367–83 (1996).

254 *See id.*

255 Asim Zia et al., *From the Habit of Control to Institutional Enablement*, 1 COMPLEXITY, GOVERNANCE & NETWORKS 79, 82 (2014).

256 *See generally* OSKAR R. LANGE, *The Computer and the Market*, *in* SOCIALISM, CAPITALISM AND ECONOMIC GROWTH (C. H. Feinstein ed., 1967).

257 Ali Wyne, *Empirics and Psychology: Eight of the World's Top Young Economics Discuss Where Their Field is Going*, BIG THINK, http://bigthink.com/power-games/empirics-and-psychology-eight-of-the-worlds-top-young-economists-discuss-where-their-field-is-going.

data gathering and coding cannot adjust their "frames" quickly enough to adapt to changes in the underlying data.

These problematic implications of recursion have parallels in law and economics scholarship regarding "nudges," or the use of unconventional policy tools to influence behavior. Sunstein and Thaler, among others, argue in favor of using policies such as default rules as "nudges" in order to make people better off.[258] In order to use nudges effectively, peoples' preferences towards which they should be "nudged" must be known to policymakers. The very fact that people respond to nudges, however, distorts peoples' preferences. Therefore, the use of nudges poses a dilemma: "One cannot learn which option a person prefers by looking to see which option she chooses when what she chooses varies according to . . . which way she is nudged."[259] The use of data to guide legal decisions, like the use of nudges, distorts the underlying data in ways that make it difficult to determine which option people would prefer.

Moreover, defaulting people into goods and services based on their previous choices may be an affront to individual learning, which is essential to the evolution of the common law system.[260] For this reason, Sunstein emphasizes active choosing in situations where it is desirable to encourage learning and the development of preferences, such as in democratic elections.[261] Similarly with Big Data, using past behaviors to predict the future may discourage development of heterogeneous preferences, which are essential to adaptive evolution of law.

In short, it matters very much whether the origination process of data is independent of the data-using process. If there is independence, then we may be able to use the data profitably. But if there is dependence such that our use of the data alters the data origination process, then we get self-fulfilling prophesies and analytic distortions.[262] In any event, there will always be a frame in the origination process of data that will shape the data that is ultimately collected.

A legal system governed by Big Data, then, could be one in which partially open-ended, indeterministic metaphors are reduced to algorithmic propositions, which would then be replicated *ad infinitum* in

258 *See* RICHARD THALER & CASS SUNSTEIN, NUDGE: IMPROVING DECISIONS ABOUT HEALTH, WEALTH, AND HAPPINESS 8 (2008) (ebook).

259 Jacob Goldin, *Which Way to Nudge? Uncovering Preferences in the Behavioral Age*, 125 YALE L.J. 226, 230 (2015).

260 CASS SUNSTEIN, WHY NUDGE? THE POLITICS OF LIBERTARIAN PATERNALISM 94 (2014) (citing JOHN STUART MILL, ON LIBERTY 8 (Kathy Casey ed., 2002) (1859) (stating that "the free development of individuality is . . . a coordinate element with all that is designated by the terms civilization, instruction, education, culture")).

261 *Id.*

262 *See* Robert K. Merton, *The Self-Fulfilling Prophecy*, THE ANTIOCH REV. 193, 193–210 (discussing the economic consequences of self-fulfilling prophecies).

feedback loops. Because these algorithmic propositions inherently lack the richness in meaning that semantic propositions have, the legal system would become highly distorted as these propositions are applied in ever more absurd ways, unmoored from their original meaning.

Big Data cannot discern any more meaning than we have given it the inputs to "understand." When faced with questions of meaning, such as the legal implications of one's status as birth mother or genetic mother, or the weighing of utilitarian and retributivist factors in sentencing, Big Data would likely code an outcome that would seem to us to be arbitrary. Recursion is thus problematic because it is circular and self-reinforcing. It creates the very world it purports to reflect by using past outcomes to predict the future, and discounts the creatively evolving, dynamic nature of social systems.

Ultimately, this vicious reflexivity of Big Data has broad implications for the distribution and exercise of political power. The largest nodes in a network, by definition, have the most influence over the network. This provides one definition of power, which is "both created by, and summarizes, the interactions of a society."[263] Social scientists have long studied "how the manifold interactions within a social group lead to hierarchy of status that bears some—but often not very much—relationship to the original intrinsic properties of the individuals themselves."[264]In this context, certain interactions take on more relevance and influence than others, for seemingly arbitrary reasons.

Like Big Data, social power is often recursive, as "[t]o know social power is to know more than just facts about individuals: it is to summarize innumerable facts about the thoughts individuals have about each other, and thoughts about those thoughts, and so forth."[265] Therefore, "to have power is to be seen to have power by those who are themselves powerful."[266]

Big Data would take legal power out of the hands of novelty intermediaries—judges, lawyers, politicians, and litigants who act as legal "entrepreneurs," searching the landscape of statutes, regulations and caselaw for new opportunities in the legal adjacent possible—and place that power in the hands of Big Data servers and those who own and control them.

263 Simon DeDeo, *Major Transitions in Political Order* 7 (SFI Working Paper 16-06-010, 2016).

264 *Id.*; *see generally* MICHAEL MANN, THE SOURCES OF SOCIAL POWER, VOL. 1 (2012).

265 DeDeo, *supra* note 262.

266 *Id.*

CONCLUSION

We have described the reasons why the Big Data paradigm will be problematic in the legal context. Our argument is essentially three-fold. First, law and data have multiple possible compressions, interpretations or meanings which cannot be interpreted by Big Data. Second, these compressions or meanings are indefinite and continually evolving in ways that cannot be predicted by Big Data. Thus the increased availability of data privileges, rather than obviates the need for theory. Third, the use of Big Data, accordingly, generates self-reinforcing feedback loops, essentially causing law to go "viral" with irrational, herding-like behaviors.

Our biggest concern is the potential result. Big Data could lull the legal system into accepting opaque, undemocratic decision-making far removed from the value-laden debates that should animate a free society. We have not seen any convincing evidence that these structural deficiencies of the Big Data paradigm can be overcome by superior technology. In our view, the only satisfying solution is to recognize, and adhere to, the value of human judgment in the legal system, which can be supplemented, but never be superseded by algorithms.