

UCLA

UCLA Public Law & Legal Theory Series

Title

Environmental Law, Big Data, and the Torrent of Singularities

Permalink

<https://escholarship.org/uc/item/5px0s7zt>

Journal

UCLA School of Law Public Law & Legal Theory Series

Author

Boyd, William

Publication Date

2016-11-11

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at

<https://creativecommons.org/licenses/by-nc-nd/4.0/>



LEGAL STUDIES RESEARCH PAPER SERIES

Paper Number 16-9

November 7, 2016

Environmental Law, Big Data, and the Torrent of Singularities

William Boyd

University of Colorado Law School

64 UCLA L. Rev. Disc. 544 (2016)

This paper can be downloaded without charge from
the Social Science Research Network electronic library
at: <https://ssrn.com/abstract=2866570>



Environmental Law, Big Data, and the Torrent of Singularities

William Boyd

ABSTRACT

How will big data impact environmental law in the near future? This Essay imagines one possible future for environmental law in 2030 that focuses on the implications of big data for the protection of public health from risks associated with pollution and industrial chemicals. It assumes the perspective of an historian looking back from the end of the twenty-first century at the evolution of environmental law during the late twentieth and early twenty-first centuries. The premise of the Essay is that big data will drive a major shift in the underlying knowledge practices of environmental law (along with other areas of law focused on health and safety). This change in the epistemic foundations of environmental law, it is argued, will in turn have important, far-reaching implications for environmental law's normative commitments and for its ability to discharge its statutory responsibilities. In particular, by significantly enhancing the ability of environmental regulators to make harm more visible and more traceable, big data will put considerable pressure on previous understandings of acceptable risk across populations, pushing toward a more singular and more individualized understanding of harm. This will raise new and difficult questions regarding environmental law's capacity to confront and take responsibility for the actual lives caught up in the tragic choices it is called upon to make. In imagining this near future, the Essay takes a somewhat exaggerated and, some might argue, overly pessimistic view of the implications of big data for environmental law's efforts to protect public health. This is done not out of a conviction that such a future is likely, but rather to highlight some of the potential problems that may arise as big data becomes a more prominent part of environmental protection. In an age of data triumphalism, such a perspective, it is hoped, may provide grounds for a more critical engagement with the tools and knowledge practices that inform environmental law and the implications of those tools for environmental law's ability to meet its obligations. Of course, there are other possible futures, and big data surely has the potential to make many positive contributions to environmental protection in the coming decades. Whether it will do so will depend in no small part on the collective choices we make to manage these new capabilities in the years ahead.

AUTHOR

William Boyd is Professor and John H. Schultz Energy Law Fellow at the University of Colorado Law School. He thanks participants in the April 2016 Program on Understanding, Law Science, and Evidence (PULSE) Symposium, “Imagining the Legal Landscape: Technology and the Law in 2030,” for stimulating conversation and valuable feedback on an earlier draft of this Essay. Special thanks to Richard Re and Dean Jennifer Mnookin for the opportunity to participate in the PULSE Symposium and to the editors of the *UCLA Law Review* for their excellent suggestions and editorial assistance. Thanks also to participants at the 2016 Colorado-Duke Climate Change Law and Policy Workshop for comments on an earlier draft.

TABLE OF CONTENTS

I. A VIEW FROM THE ARCHIVES	546
II. THREE AGES OF ENVIRONMENTAL LAW: A TECHNOLOGY STORY	548
III. BIG DATA AND THE TORRENT OF SINGULARITIES	554
IV. RISK AND THE SUBJECT	558
V. ACCOUNTABILITY AND EXPERTISE	565
VI. VIOLENCE, MEMORY, RESPONSIBILITY	568

*Singularity is disconcerting; what can be done with these countless individuals, their tenuous plans, their many disjointed movements?*¹

I. A VIEW FROM THE ARCHIVES

In her reflections on archival practice and the historian's craft, Arlette Farge takes us deep into the immense archives of the French criminal justice system.² Her account of the challenges of recovering individual lives and the work of producing historical knowledge might seem far afield for an Essay charged with speculating about the near future. But Farge's confrontation with the unique, singular lives of those who left traces in the archives has considerable relevance for how governments, today and in the future, will deal with massive and growing flows of data about individual lives.³ By exploring the challenges of incorporating individual lives into larger historical narratives, she raises hard questions about the tendency to naturalize certain abstractions and the violence that sometimes results when such abstractions are deployed as tools of governance.⁴ In doing so, she forces us to confront the difficulties involved in making knowledge out of a mass of singular, partially identified lives and their various collisions and conflicts with established systems of power.⁵

This Essay investigates these difficulties in the context of big data and environmental law. As used here, big data refers not only to the sheer volume of data on health and environmental conditions that is becoming available but also to new techniques in data analytics and new infrastructures for translating those data into governmental policy and practice.⁶ The premise is that over the next fifteen

1. ARLETTE FARGE, *THE ALLURE OF THE ARCHIVES* 91 (Thomas Scott-Railton trans., 2013) (1989).
2. *See generally id.*
3. *See generally id.*
4. *Id.* at 92 ("Our task is to find a language that can integrate singular moments into a narrative capable of reproducing their roughness, while underlining both their irreducibility and their affinity with other representations. . . . The human being captured in documents . . . should not simply be measured against the yardstick of the average person from his time, about whom we have little to say. Rather, these individuals should be approached with an eye toward drawing out the sequence of strategies that each person uses to make his way in the world.").
5. *Id.* at 86 ("It is fruitless to search through the archives for something, anything, that could reconcile these opposites. Because the historical event also resides in this torrent of singularities, which are as contradictory as they are subtle, sometimes even overwhelming. History is not a balanced narrative of the results of opposing moves. It is a way of taking in hand and grasping the true harshness of reality, which we can glimpse through the collision of conflicting logics.").
6. "Big data" has emerged as a keyword for the early twenty-first century and, as such, has multiple meanings and normative valences spanning the full spectrum from triumphalist to dystopian. For a general discussion, see Rob Kitchin, *Big Data, New Epistemologies and Paradigm Shifts*, Apr.–June

years or so big data will drive a major shift in the underlying knowledge practices of environmental law (along with other areas of law focused on health and safety). This change in the epistemic foundations of environmental law, it is argued, will in turn have important, far-reaching implications for the normative commitments of environmental law and for its ability to discharge its statutory responsibilities.⁷

Rather than speculating about what big data might mean for the future of environmental law, however, the Essay assumes the opposite posture: that of an historian looking back from the late twenty-first century at the implications of big data for environmental law circa 2030. By shifting the perspective in this manner, the Essay seeks to open up a way of thinking about the history of environmental law as a technology story and to situate the current moment and its implications for the near future in the longer sweep of the field.

Of course, our future historian will not have the opportunity or the need to spend time in physical archives such as those that occupy Farge. Instead, she will have virtually instantaneous access to immense digital archives and data analytics capabilities that would have been inconceivable to previous generations of historians. Part of her task will be to understand what is revealed and what is not (what is remembered and what is forgotten) in these new digital archives about people's lives and the harms they suffer. In particular, she will investigate how regulatory agencies such as the Environmental Protection Agency (EPA), the Occupational Safety and Health Administration (OSHA), the Food and Drug Administration (FDA) and other relevant organs of the federal government, dealt with vast troves of data on specific environmental conditions and trends over time and space—including the avalanche of data made available by the proliferation of personal trackers, wearables, and other tools of self-measurement that became quite common in the United States starting in the 2010s. She will want to know how these agencies used the proliferation of databases regarding health and medical records, much of it at the individual level given the boom in personalized medicine and the move by insurance companies, employers, and governments to nudge people in various ways to make their health data available

2014 BIG DATA & SOCIETY 1, 1. For a discussion and overview of big data and environmental law, see Linda K. Breggin & Judith Amsalem, *Big Data and the Environment: A Survey of Initiatives and Observations Moving Forward*, 44 ENVTL. L. REP. 10984 (2014). See generally ENVIRONMENTAL LAW INSTITUTE, *BIG DATA AND ENVIRONMENTAL PROTECTION: AN INITIAL SURVEY OF PUBLIC AND PRIVATE INITIATIVES* (2014).

7. This Essay does not focus on climate change, a problem that will clearly occupy a great deal of attention in coming decades, affecting most of EPA's programs, and one that will surely call upon the use of big data in various ways. Instead, the Essay focuses on how big data will impact EPA's core responsibilities in protecting public health against pollution and toxic substances.

to analytics firms and others for purposes of personal improvement and the development of individually tailored health care options (we will leave aside for now the privacy implications of all of this).

In short, she will want to understand how these agencies grappled with an unprecedented ability to trace the harms, insults, and consequences of environmental pollutants through multiple pathways (air, water, food) and to link these with the life histories and body burdens of large numbers of individuals. With the advent of big data, regulators had a powerful new lens through which to view the environmental histories of singular individuals—making visible for the first time the subtle harms, the diminished capacities, the slow violence that environmental pollutants visited upon real people living real lives in real places.⁸ But with these new powers of observation and surveillance came new capabilities for sorting individuals across a more precise set of risk factors and rendering those lives as suitable objects of regulation.

In carrying out her project, our future historian will need to map the field of environmental law by investigating its knowledge practices, identifying major shifts in those practices and how they have affected EPA's ability to meet its responsibilities for environmental protection. As suggested, her central claim is that during the fifteen-year period from 2015 to 2030, big data ushered in a major shift in the knowledge practices underwriting environmental regulation in the United States and other countries. This shift significantly enhanced the ability of environmental regulators to make harm more identifiable and more traceable, putting considerable pressure on environmental law's basic epistemic and normative commitments—throwing into sharp relief the limits of an older version of environmental law based on an actuarial logic of risk. In the process, traditional notions of expertise and accountability in environmental law were diminished. Important questions were raised regarding environmental law's ability to confront and take responsibility for the actual lives caught up in the tragic choices it is called upon to make.

II. THREE AGES OF ENVIRONMENTAL LAW: A TECHNOLOGY STORY

There are many ways to tell the history of American environmental law. Standard approaches have focused for good and obvious reasons on the great environmental lawmaking moment of the 1970s; the interactions between Congress, EPA, and the courts; the role of environmental advocacy groups and citizen participation; the rise of risk assessment and cost-benefit analysis; debates

8. See ROB NIXON, *SLOW VIOLENCE AND THE ENVIRONMENTALISM OF THE POOR* (2011).

over policy instruments; and the politics of environmental protection in the late twentieth and early twenty-first centuries.⁹

But what might such a history look like if told as a science and technology story? What would it look like if the focus instead were on the underlying knowledge practices—the changing suite of concepts, tools, and practices—that allow environmental law to do its work? At the risk of oversimplification, such a history might identify three important ages of American environmental law spanning the period from the 1960s to the 2030s:

- **1960s–1970s, The Age of Detection:** During the 1960s and 1970s, a revolution in analytical techniques and detection capabilities had a profound effect on health, safety, and environmental regulation. The massive expansion in animal testing that had been underway since the 1950s, for example, uncovered a much larger universe of potential carcinogens and other harmful substances than previously recognized.¹⁰ At the same time, major advances in analytical techniques and detection capabilities revealed chemical residues and other substances in food, air, water, and living tissues at far lower concentrations than was possible in previous decades. Between 1958 and 1978, for example, the sensitivity of detection capabilities for chemicals in food increased by up to five orders of magnitude.¹¹ Similar advances were made in the ability to detect trace organic compounds and other hazardous substances in the environment, with detection limits pushed into the low parts per billion and even parts per trillion range, several orders of magnitude below what had been possible only a decade earlier.¹² The period

9. For general histories of U.S. environmental law and politics, see JONATHAN Z. CANNON, *ENVIRONMENT IN THE BALANCE: THE GREEN MOVEMENT AND THE SUPREME COURT* (2015); RICHARD J. LAZARUS, *THE MAKING OF ENVIRONMENTAL LAW* (2004); RICHARD N.L. ANDREWS, *MANAGING THE ENVIRONMENT, MANAGING OURSELVES: A HISTORY OF AMERICAN ENVIRONMENTAL POLICY* (1999); SAMUEL P. HAYS, *BEAUTY, HEALTH, AND PERMANENCE: ENVIRONMENTAL POLITICS IN THE UNITED STATES, 1955–1985* (1987).

10. See Richard Wilson, *Risks Caused by Low-Levels of Pollution*, 51 *YALE J. BIO.* 37, 48 (1978) (noting that in 1958 there were only four known human carcinogens but that twenty years later scientists had identified thirty-seven human carcinogens and more than 500 animal carcinogens); Listing of D&C Orange No. 17 for Use in Externally Applied Drugs and Cosmetics, 51 *Fed. Reg.* 28,331, 28,343 (Aug. 7, 1986) (recounting developments in animal testing).

11. See Chemical Compounds in Food Producing Animals, 44 *Fed. Reg.* 17,070, 17,075 (Mar. 20, 1979) (discussing advances in detection capabilities).

12. See EPA, *MEASUREMENT OF CARCINOGENIC VAPORS IN AMBIENT ATMOSPHERES* 1 (1978) (discussing development of analytical techniques to collect and analyze air samples to detect presence of toxic and/or carcinogenic organic compounds); *Air Quality Criteria: Hearing Before the Subcomm. on Air and Water Pollution of the S. Comm. On Public Works*, 90th Cong. (1970), reprinted in 1 *CLEAN AIR ACT AMENDMENTS OF 1970* 608, 615–16 (1970) (statement of Dr. Samuel S. Epstein) (discussing development of new “high sensitive biological techniques for measuring the

also witnessed significant improvements in the ability to track the fate and transport of chlorinated organic compounds and other synthetic compounds in the global environment.¹³ Finally, new extrapolative techniques were developed and refined during this period, allowing scientists to extend dose-response relationships from the observable to the unobservable range in a quantitative manner.¹⁴ Taken together these new ways of seeing revealed a world of environmental hazards that far exceeded, in scale and scope, previous understandings. All of which created considerable challenges for environmental law and provided an important impetus for the move away from an earlier approach founded on precaution to one focused on quantitative risk assessment.¹⁵

- **1980s–2010s, The Age of Modeling and Simulation:** Starting in the 1980s, EPA greatly expanded its use of models and simulations as key tools for environmental decisionmaking. This was in large part a response to the increased analytic demands of quantitative risk assessment, an expanding set of statutory responsibilities, and an atmosphere of growing public distrust. Models of various types became central to the agency's efforts to assess the potential health effects of air and water pollution and establish health-based standards, to predict the fate and

carcinogenicity of organic extracts of atmospheric pollutants"); L.A. Wallace, *Human Exposure to Environmental Pollutants: A Decade of Experience*, 25 CLINICAL & EXPERIMENTAL ALLERGY 4, 4 (1995) (discussing "exquisite sensitivities" of gas chromatography-mass spectrometry and their use by the EPA during the early 1970s); William L. Budde & James W. Eichelberger, *Organics in the Environment*, 51 ANALYTICAL CHEMISTRY 567A, 567A (1979) (discussing new "measurements of the presence and concentration of a variety of pollutants . . . made possible by significant advances in analytical instrumentation, electronics, computer science, and analytical chemistry during the 1960s and early 1970s").

13. See George M. Woodwell, *Toxic Substances and Ecological Cycles*, 216 SCI. AM. 24 (1967) (discussing studies aimed at understanding "global, long-term ecological processes that concentrate toxic substances" in the environment); L. Harrison et al., *Systems Studies of DDT Transport*, 170 SCI. 503 (1970) (discussing use of systems models for understanding long-term impacts of DDT in ecosystems); David B. Peakall & Jeffrey L. Lincer, *Polychlorinated Biphenyls: Another Long-Life Widespread Chemical in the Environment*, 20 BIOSCIENCE 958 (1970) (documenting presence of PCBs in various environmental media and animal tissues).
14. See, e.g., N. Mantel & W.R. Bryan, *Safety Testing of Carcinogenic Agents*, 27 J. NAT. CANCER INST. 455 (1961) (proposing log-probit model for low dose extrapolation); David G. Hoel, *Statistical Models for Estimating Carcinogenic Risks From Animal Data*, PROCEEDINGS OF THE 5TH ANNUAL CONFERENCE ON ENVIRONMENTAL TOXICOLOGY (1974) (discussing statistical techniques for low-dose extrapolations to estimate cancer risk in humans associated with exposure to various environmental agents and their applications in regulatory context); NAT'L RESEARCH COUNCIL, DRINKING WATER AND HEALTH 29–59 (1977) (discussing use and associated challenges of extrapolation techniques in chemical safety assessment).
15. See William Boyd, *Genealogies of Risk: Searching for Safety: 1930s–1970s*, 39 ECOL. L. Q. 895, 944–47 (2012) (describing these developments).

transport of hazardous substances at superfund sites and define remediation targets, to evaluate the structure, activity, and potential toxicity of new industrial chemicals, to map out possible exposure pathways for pesticide residues on foods, and to manage uncertainty. Drawing on dramatic and ongoing increases in computational capacity, EPA's modeling programs expanded across all of the agency's major branches during the 1980s and 1990s. By the early 2000s, more than one hundred models were in use throughout the agency, providing a substantial part of the knowledge infrastructure supporting the agency's decisionmaking.¹⁶ Models promised to significantly reduce the costs and time of acquiring the information necessary to support regulatory action and survive legal challenge. They made it possible to go where traditional techniques of monitoring and testing could not go and allowed for new understandings of complex environmental harms. As such, they marked an important shift from earlier reliance on expert judgment to one relying more heavily on expert systems.

- **2010s–2030s, The Age of Big Data:** During the 2010s, several important developments in data science and information technology converged to usher in a major shift toward “big data” (the buzzword of the times) as a foundation for environmental, health, and safety regulation. These developments included an unprecedented ability to track sources, concentrations, and receptors (human and environmental) over time and space using geographic information technologies, ubiquitous sensing for individual and ecological exposure assessment, cheap and widespread biomonitoring of individuals to assess actual exposures, and data analytics capabilities for linking and evaluating source-exposure and exposure-disease relationships. Together these developments allowed for the first time a move to fine-scale, individualized exposure assessments over space and time. This led to sharp reductions in the uncertainty associated with past, proxy-based exposure assessments and, in the process, highlighted some of the errors in previously understood dose-response relationships.¹⁷ It also resulted in a further displacement of traditional

16. See NAT'L RESEARCH COUNCIL, *MODELS IN ENVIRONMENTAL REGULATORY DECISION MAKING* 44 (2007).

17. During the 1990s and 2000s, exposure assessment was sometimes referred to as the “Achilles heel” of environmental epidemiology because of the inability to assess exposures at the individual level and in key microenvironments. See, e.g., TOPICS IN ENVIRONMENTAL EPIDEMIOLOGY (Kyle Steenland & David A. Savitz eds., 1997) (discussing limits of proxy-based exposure assessments).

notions of expertise, raising difficult questions regarding environmental law's responsibility to protect against certain kinds of harm.

While each of these phases built upon and assimilated the tools and techniques of prior phases, each was also marked by a dominant set of knowledge practices that underwrote a particular normative approach to environmental regulation and a distinctive conception of expertise. The third phase, which is the focus of this project, was the time when EPA, like many agencies across the federal government, got serious about data science and began to integrate its data efforts across various domains. Here, for example, is how EPA described its then nascent efforts in 2016:

EPA is at the beginning of a transformative stage in information management, where there are new and enhanced ways to gather data, conduct analysis, perform data visualization and use "big data" to explore and address environmental, business, and public policy challenges. Business intelligence tools, geospatial tools and visualization tools are also converging, providing opportunities for knowledge and insight from disparate data sources.

EPA will reap huge benefits from the ability to harness this power of data corporately across the enterprise to produce knowledge on demand and drive better environmental results. The Office of Information Analysis and Access intends to work with EPA programs and regional offices to create a central platform for data analytics to evaluate information across a wide range of data collections and break through the Agency's stove-piped set of data systems, which do not make these associations easy or common today.

Work in this area will also fit with the Agency's interests in advanced monitoring technology that leverages sensors, real time data and external data sources such as NASA satellites, or financial or health data in combination with EPA data sources. Expert systems and machine learning to target violations using data acquired through electronic reporting, as well as analytics with unstructured data and/or scattered data sets across the Agency are also envisioned as part of this new, emerging program at EPA.¹⁸

Like other federal agencies, EPA hired its first Chief Data Scientist in 2015 and began in earnest to create a new data platform for environmental protection.

18. See EPA, *EPA's Cross-Agency Data Analytics and Visualization Program*, SEMANTIC COMMUNITY, http://semanticcommunity.info/Data_Science/Data_Science_for_EPA_Big_Data_Analytics#EPA.E2.80.99s_Cross-Agency_Data_Analytics_and_Visualization_Program [<https://perma.cc/GA7K-AQMC>].

During this period, EPA officials expressed considerable enthusiasm for the possibilities associated with big data. Witness the words of Cynthia Giles, EPA Assistant Administrator for Office of Enforcement and Compliance Assurance:

Just as the internet has transformed the way we communicate and access information, advances in information and emissions monitoring technology are setting the stage for detection, processing, and communication capabilities that can revolutionize environmental protection. We are moving toward a world in which states, EPA, citizens, and industry will have real-time electronic information regarding environmental conditions, emissions, and compliance.¹⁹

Realizing the promise of big data for environmental protection, of course, would take time (the better part of two decades in fact) and, as we will see, there were important yet poorly understood consequences associated with this shift. But it was a profound shift.

At the center of this shift was a new concept, known at the time as the exposome. As one report described it, the “exposome” was premised on the idea that “the totality of environmental exposures (including such factors as diet, stress, drug use, and infection) throughout a person’s life can be identified.”²⁰ By compiling and mapping individual exposures at high resolution across space and time, the exposome made it possible to identify the cumulative exposures experienced by particular bodies to multiple pollutants and other environmental factors and to match this with information on individual genomes and behaviors. By the middle of the twenty-first century, a majority of people in the United States (and other rich countries) had access to some version of a personal exposome. As we will see, the quality of and access to these exposomes was uneven, a fact that had important distributional implications for environmental protection and public health.

Much of the growing interest in the idea of a personal exposome was made possible by the growing availability of cheap, portable (and increasingly wearable)

-
19. Cynthia Giles, *Next Generation Compliance*, Sept.–Oct. 2013 ENVTL. F. 22 (2013); *see also id.* at 24 (“Monitoring devices are becoming more accurate, more mobile, and cheaper, all of which are contributing to a revolution in how we find and fix pollution problems.”).
 20. NAT’L RESEARCH COUNCIL, *EXPOSURE SCIENCE IN THE 21ST CENTURY: A VISION AND A STRATEGY* 4 (2012); *see also* Stephen M. Rappaport & Martyn T. Smith, *Environment and Disease Risks*, 330 *SCIENCE* 460, 460–61 (2010) (“The term ‘exposome’ refers to the totality of environmental exposures from conception onwards, and has been proposed to be a critical entity for disease etiology.”); Christopher Paul Wild, *Complementing the Genome With an “Exposome”: The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology*, 14 *CANCER EPIDEMIOLOGY, BIOMARKERS, & PREVENTION* 1847 (2005). Wild, who was the director of the International Agency for Research on Cancer, is credited with first introducing the term in his 2005 article.

sensors.²¹ By 2030, for example, billions of sensors were connected to the Internet, up from a mere ten million in 2007.²² Many of these sensors were incorporated in clothing and everyday devices that people used, providing massive new flows of data on individual exposures and environmental conditions.²³ This in turn required very large investments in data analytics and the development of algorithms to process, correlate, and extract information from such data. Environmental regulation would never be the same.

III. BIG DATA AND THE TORRENT OF SINGULARITIES

By the 2030s, this growing knowledge infrastructure made it possible for EPA and other agencies to identify and trace the harms associated with various forms of pollution and toxic substances in ways that had previously been possible only for certain well-identified exposure pathways and a few signature diseases.²⁴

-
21. See, e.g., Emily G. Snyder et al., *The Changing Paradigm of Air Pollution Monitoring*, 47 ENVTL. SCI. TECH. 11369, 11369 (2013) (discussing a move away from expensive, stationary air pollution monitors to “lower-cost, easy-to-use, portable air pollution monitors (sensors) that provide high-time resolution data in near real-time”).
 22. See Melanie Swan, *Sensor Mania! The Internet of Things, Wearable Computing, Objective Metrics, and the Quantified Self 2.0*, 1 J. SENSOR & ACTUATOR NETWORKS 217, 218 (2012) (citing various estimates of internet-connected devices and sensors). Some contemporary observers predicted that more than one hundred trillion sensors would be connected to the Internet by 2030. See, e.g., MICHAEL PATRICK LYNCH, *THE INTERNET OF US: KNOWING MORE AND UNDERSTANDING LESS IN THE AGE OF BIG DATA* 7 (2016).
 23. The National Research Council (NRC) called for exactly such a device in its 2012 report on exposure science:
 “There is a need for a wearable sensor that is capable of monitoring multiple analytes in real time. Such a device would allow more rapid identification of “highly exposed” people to help identify sources and means of reducing exposures. Recent advances in nanoscience and nanotechnology offer an unprecedented opportunity to develop very small, integrated sensors that can overcome current limitations.”
 NAT’L RESEARCH COUNCIL, *supra* note 20, at 9. The nonprofit Environmental Defense Fund launched a pilot project in 2015 with a company called MyExposome that gave twenty-eight people wristbands that detected daily exposure to dozens of chemicals. See *Chemical Detection Project: Pilot Results*, ENVTL. DEF. FUND, <https://www.edf.org/health/chemical-detection-project-pilot-results> [<https://perma.cc/3GPP-JHS6>].
 24. Prominent examples of such signature diseases include mesothelioma from exposure to certain types of asbestos fibers and clear cell adenocarcinoma (a cancer of the cervix and vagina) as a result of prenatal exposure to Diethylstilbestrol (DES), a synthetic form of estrogen that was prescribed to pregnant women between 1940 and 1971 to prevent miscarriage, premature labor, and related complications of pregnancy. On asbestos, see Marty S. Kanarek, *Mesothelioma From Chrysotile Asbestos: Update*, 21 ANNALS EPIDEMIOLOGY 688, 688 (2011) (reviewing “overwhelming evidence” that asbestos is responsible for “mesothelioma”); M.C. Goodwin & Juraj Jagatic, *Asbestos and Mesotheliomas*, 3 ENVTL. RES. 391, 391 (1970) (“The significant association between mesothelioma of the pleura and peritoneum and asbestos is widely accepted.”). On DES, see Arthur L. Herbst et al., *Adenocarcinoma of the Vagina*, 284 NEW ENG. J. MED. 878 (1971) (tracing increased incidence of adenocarcinoma of the vagina in young women to prenatal exposure to DES).

This challenged EPA's prevailing approach to risk assessment and the ways in which regulators had formerly considered the lives at stake in their decisions. Put more concretely, it precipitated a move from statistical lives to identified lives—from populations to singularities—that forced environmental law to confront the consequences of its choices far more directly than in the past.²⁵

With big data, environmental regulators were for the first time able to focus on particular, identified lives in a systematic way.²⁶ As the National Research Council put it in 2012, the convergence of scientific and technological advances in exposure science “raises the possibility that in the near future integrated sensing systems will facilitate individual-level exposure assessments in large populations of humans.”²⁷ Within a couple of decades, that possibility had been realized as the individual exposure histories of actual people that had been largely beyond environmental law's reach were made available as a basis for environmental decisionmaking.

This forced new understandings of the normal and the pathological and new orientations toward the future (and the past). More precise attention to individual exposures and potential harms meant that environmental law could no longer focus simply on general population risks in the future. As individuals gained access to their own personal exposomes, each with a distinct set of exposures and potential harms, environmental law thus found itself confronting a

ingested by their mothers during pregnancy); see also Nancy Langston, *The Retreat From Precaution: Regulating Diethylstilbestrol (DES), Endocrine Disruptors, and Environmental Health*, 13 ENVTL. HIST. 41, 51 (2008) (discussing link between prenatal DES exposure and rare vaginal cancers in young women).

25. See Lisa Heinzerling, *Statistical Lives in Environmental Law*, in IDENTIFIED VERSUS STATISTICAL LIVES: AN INTERDISCIPLINARY PERSPECTIVE (I. Glenn Cohen et al. eds., 2015). But these “identified lives” were not exactly concrete lives. Rather, they were “digital shadows” (not unlike the traces that Farge was working with in her archival research). As such, they were part of what Julie Cohen referred to in the early twenty-first century as the “biopolitical public domain”—a digital commons of sorts where the proliferation of data on individual lives and habits became a resource subject to a relentless commercial logic of extraction. Surplus value lived and flowed through data about consumers and their preferences. At the same time, this biopolitical public domain also provided enormous and growing opportunities (and challenges) for governments and for efforts to regulate environmental harm. See Julie Cohen, *The Biopolitical Public Domain* (Sept. 28, 2015) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2666570.

26. See Heinzerling, *supra* note 25.

27. NAT'L RESEARCH COUNCIL, *supra* note 20, at 7; see also NAT'L RESEARCH COUNCIL, *FRONTIERS IN MASSIVE DATA ANALYSIS* 12 (2013) (“[W]hat is particularly notable about the recent rise in the prevalence of ‘big data’ is not merely the size of modern data sets, but rather that their fine-grained nature permits inferences and decisions at the level of single individuals.”).

“torrent of singularities” that demanded attention to individual lives and to the accumulation of harms they experienced.²⁸

To be sure, the widespread mapping of individual exposomes proved critical in understanding the environmental contributions to various chronic diseases because it provided, for the first time, the ability to account for cumulative exposures to multiple agents over long time frames (individual lifetimes) and to map their interactions with specific genetic, behavioral, and other environmental factors. It also provided the basis for major advances in understanding how the timing of certain exposures affected health and development, particularly neurodevelopment in children and the subtle harms caused by exposure to endocrine disrupting chemicals during particular life stages.²⁹ In doing so, it raised the bar considerably for environmental protection.

Up until this time, environmental health science and environmental regulation had struggled with the limitations of existing approaches to understanding exposure and risk. Reliance on animal studies, limited and uneven toxicity testing for many chemicals, gaps in environmental monitoring, and, perhaps most importantly, the inability to map exposure pathways in any systematic way allowed for only a crude understanding of environmental health risks across populations. At the same time, however, there was a general recognition that many chronic diseases such as cancer had strong environmental components, and that the growth rates of many diseases, including asthma, autism, dyslexia, attention-deficit hyperactivity disorder, schizophrenia, obesity, diabetes, and certain cancers were too rapid to be of genetic origin.³⁰

28. Farge, *supra* note 1, at 86. These developments also had obvious and important implications for tort law, given the enhanced ability to trace specific individual harms to particular exposures and sources. While a discussion of these implications is beyond the scope of this Essay, any thorough investigation of big data and environmental law would need to attend to the ways in which the ability to make individual harms more visible and more traceable affected the viability of various causes of action under tort law as well as the broader relationship between environmental law and tort law.

29. These issues had been a major topic of concern among environmental health professionals and others since the mid-1990s. *See, e.g.*, Bruce P. Lanphear, *The Impact of Toxins on the Developing Brain*, 36 ANN. REV. PUB. HEALTH 21, 212 (2015) (noting that “evidence has accumulated over the past century that implicates ubiquitous, low-level exposures to an every-growing litany of environmental toxins in the development of diminished birth weight, shortened gestation, intellectual deficits, and mental disorders in children”); P. Grandjean & P.J. Landrigan, *Developmental Neurotoxicity of Industrial Chemicals*, 368 LANCET 2167 (2006) (discussing accumulating evidence of neurodevelopmental damage in children caused by industrial chemicals); *see also* THEO COLBURN ET AL., OUR STOLEN FUTURE (1996) (discussing public health implications of endocrine disrupting chemicals).

30. *See, e.g.*, PRESIDENT’S CANCER PANEL, REDUCING ENVIRONMENTAL CANCER RISK (2010) (finding “that the true burden of environmentally induced cancer has been grossly underestimated”); Philip J. Landrigan & Dean B. Baker, *The National Children’s Study—End or*

By linking specific exposure histories to various other genetic and behavioral factors, then, the exposome brought environmental health to the individual level. In doing so, it provided a much-needed complement to the considerable effort undertaken during the first two decades of the 2000s to sequence individual genomes and to understand the genetic contribution to many chronic diseases. Although launched with great fanfare, the genomics effort, in fact, essentially confirmed that the incidence of most diseases was primarily the result of environmental rather than genetic factors.³¹ The exposome thus emerged as critical in the effort to understand and reduce the burden of disease across large populations.³² By the 2030s, with a decade of exposome data available for millions of

-
- New Beginning?*, 372 N. ENG. J. MED. 1486, 1486 (2015) (discussing rising rates of chronic diseases in children and arguing for a strong emphasis on environmental exposures in any longitudinal study of children's health); see also Song Wu et al., *Substantial Contribution of Extrinsic Risk Factors to Cancer Development*, 529 NATURE 43, 46–47 (2016) (concluding that 70 to 90 percent of most common cancer types result from exposure to environmental or "extrinsic" factors).
31. See, e.g., David E. Adelman, *The False Promise of the Genomics Revolution for Environmental Law*, 29 HARV. ENV. L. REV. 1, 4–5 (2005) (questioning the promise of genomics for environmental law given the inherent complexity and heterogeneity of biological systems and the importance of environmental rather than genetic factors in causing disease); Yuxia Cui et al., *The Exposome: Embracing the Complexity for Discovery in Environmental Health*, 124 ENVTL. HEALTH PERSP. A-137, A-137 (2016) (observing that "for many complex human diseases such as cancer, cardiovascular diseases, respiratory diseases, and type 2 diabetes, which are among the leading causes morbidity and mortality in human populations, genetic variation only explains a portion of the disease risk, and much of the disease burden is likely attributed to differences in the environment and the interplay between an individual's genes and the environment"); Gary W. Miller & Dean P. Jones, *The Nature of Nurture: Refining the Definition of the Exposome*, 137 TOXICOLOGICAL SCI. 1, 1 (2014) ("Genome-wide association studies (GWAS) have revealed genetic associations and networks that improve understanding of disease, but these still account for only a fraction of disease risk. With the majority of disease causation being nongenetic, the need for improved tools to quantify environmental contributions seems obvious."); Stephen M. Rappaport, *Genetic Factors Are Not the Major Causes of Chronic Diseases*, 11 PLoS ONE 1 (2016) (noting that the major causes of most chronic diseases appear to be exposure-related factors rather than genetic factors).
 32. Early proponents of the exposome concept pointed to its potential to better understand the causes of disease and to enhance prevention. See, e.g., Christopher P. Wild et al., *Measuring the Exposome: A Powerful Basis for Evaluating Environmental Exposures and Cancer Risk*, 54 ENVTL. & MOLECULAR MUTAGENESIS 480, 492 (2013) ("[T]he promises are so great and the value of understanding causes and prevention so high that [exposome research] should be given far greater priority, especially with its relevance not only to cancer but also to other noncommunicable diseases."). In response, governments began to support more systematic efforts starting in the 2010s to characterize and track the exposomes of particular populations. See, e.g., Martine Vrijheid et al., *The Human Early-Life Exposome (HELIX): Project Rationale and Design*, 122 ENVTL. HEALTH PERSP. 535 (2014) (describing design of the Human Early-Life Exposome (HELIX) project, a European collaboration directed at measuring and tracking the environmental exposures of 32,000 mother-child pairs). The Emory Health and Exposome Research Center: Understanding Lifetime Exposures (HERCULES) was launched in 2013 with a grant from the National Institute of Environmental Health Sciences to provide key infrastructure and expertise to

individuals, there was less need for large, expensive cohort studies over long time frames to understand the etiology of particular diseases. Environmental health could instead focus its efforts on the specific exposure and disease profiles of individuals and could do so almost in real time.

IV. RISK AND THE SUBJECT

For several decades starting in the 1970s, environmental law conceived and regulated environmental harm largely on the basis of acceptable risk (or its mirror image, unreasonable risk).³³ Rooted in population thinking, with its emphasis on statistical lives, probabilistic reasoning, and the normal distribution, risk assessment provided a means for environmental law to impose a normative order on the world of environmental harms and a rationale for allocating regulatory resources.³⁴ Although the specifics varied depending on the particular problem or statute, the basic objective was to identify and reduce risks to levels that were considered acceptable given the broader universe of potential harms that people confronted in their everyday lives.³⁵ Sometimes these risk assessments were tied to vulnerable subpopulations or to constructs such as the “maximum exposed individual,” but in all cases they were rooted in a formal understanding of risk as the

develop and refine new tools and technologies for exposome assessment. See HERCULES EXPOSOME RESEARCH CENTER, <http://emoryhercules.com> [<https://perma.cc/3DWK-5TJU>].

33. See Boyd, *supra* note 15, at 964–81 (discussing move to acceptable risk and increasing reliance on quantitative risk assessment in U.S. environmental law during the 1970s and 1980s).
34. This form of risk thinking drew directly on a logic of aggregation and statistical regularities that had been central to actuarial science since the nineteenth century. Such thinking, in fact, was instrumental to the very conception of “society” that emerged in the nineteenth century and provided the basis for a whole suite of concepts and techniques that allowed governments to approach the health and welfare of the population in more systematic ways. See *id.* at 910–15 (discussing these developments); see also NIKLAS LUHMAN, *RISK: A SOCIOLOGICAL THEORY* 102 (Rhodes Barrett trans., 1993) (discussing the “distinct forms of social solidarity” entailed by the concept of risk); Alain Desrosieres, *How to Make Things Which Hold Together: Social Science, Statistics, and the State*, in *DISCOURSES ON SOCIETY* 195, 197 (Peter Wagner ed., 1991) (noting “the extent to which the setting up of systems of statistical recording goes hand in hand with the construction of the State”); Karl H. Metz, *Paupers and Numbers: The Statistical Argument for Social Reform in Britain During the Period of Industrialization*, in *THE PROBABILISTIC REVOLUTION* 345 (Lorenz Kruger et al. eds., 1987) (“Statistics made the social state of the nation an affair that could be measured, and in this way it contributed significantly to the bureaucratization of disease and poverty that was to pave the way for the development of the welfare state.”).
35. For endpoints such as cancer, this was often framed as a risk of one-in-one-million. See Joseph V. Rodricks et al., *Significant Risk Decisions in Federal Regulatory Agencies*, 7 *REG. TOXICOLOGY & PHARMACOLOGY* 307 (1987) (discussing efforts by the Food and Drug Administration (FDA), EPA, and the Occupational Safety and Health Administration (OSHA) to define “significant” risk thresholds in their efforts to regulate carcinogens).

expected value of an undesirable outcome distributed across a population.³⁶ All of these risk estimates, moreover, were inevitably based on a range of extrapolative techniques and models, which included within them various default assumptions and inference choices intended to deal with data limits and gaps and to manage uncertainty.³⁷ Looking back from the late twenty-first century, the entire enterprise appears rather crude. At the time, however, these exercises seemed quite sophisticated and complex, consuming enormous resources and often serving as sites of intense political conflict over the scale and scope of regulation.

The shortcomings of the standard approach to risk assessment became particularly apparent in the 2000s as various studies conducted by the National Research Council and other entities pointed to ongoing problems confronting the risk assessment enterprise. As one particularly influential report from 2009 concluded, “the regulatory risk assessment process is bogged down,” facing substantial challenges in its ability to deliver useful, credible knowledge for regulators even while it confronted an increasingly complex and unpredictable world of environmental harms.³⁸ “Uncertainty,” according to the study, “continues to lead to

-
36. In setting National Ambient Air Quality Standards (NAAQS) for criteria air pollutants under the Clean Air Act, for example, the EPA administrator bases the assessment on the health implications of specific ambient concentrations of such pollutants for sensitive populations such as asthmatics, children, and the elderly. See Clean Air Act § 109, 42 U.S.C. § 7409; *Lead Industries Assoc. v. EPA*, 647 F.2d 1130, 1152–54 (D.C. Cir. 1980) (confirming that EPA must set primary NAAQS at levels that will protect the health of sensitive populations). Likewise, the Food Quality Protection Act requires that the assessment of pesticide risks focus on the health impacts for children and other vulnerable subpopulations. See Food Quality Protection Act of 1996, § 405, 11 Stat 1489, 1514–19 (requiring EPA to set tolerances for pesticide residues on food at a level that will ensure “reasonable certainty of no harm” and to apply specific safety factors to accommodate the special sensitivities of infants and children in setting such standards). And assessments of remediation targets at hazardous waste sites under the Comprehensive Environmental Response, Cleanup, and Liability Act (CERCLA) are based on the “reasonable maximum exposure” at the site. See U.S. EPA, RISK ASSESSMENT GUIDANCE FOR SUPERFUND VOLUME I: HUMAN HEALTH EVALUATION MANUAL 3-1 (2004) (defining “reasonable maximum exposure” as “the highest exposure that is reasonably expected to occur at a site but that is still within the range of possible exposures”).
 37. This overall approach also complemented the broader logic of cost-benefit analysis that had been ascendant across the U.S. regulatory state since the early 1980s. Together, risk assessment and cost-benefit analysis provided an overarching framework for environmental decisionmaking, one that received a fair amount of criticism from environmental law scholars and others favoring a more precautionary approach to environmental regulation. See, e.g., FRANK ACKERMAN & LISA HEINZERLING, PRICELESS: ON KNOWING THE PRICE OF EVERYTHING AND THE VALUE OF NOTHING (2004); DOUGLAS A. KYSAR, REGULATING FROM NOWHERE: ENVIRONMENTAL LAW AND THE SEARCH FOR OBJECTIVITY (2010).
 38. See NAT’L RESEARCH COUNCIL, SCIENCE AND DECISIONS: ADVANCING RISK ASSESSMENT ix (2009) (“[R]isk assessment is at a crossroads. Despite advances in the field, it faces a number of substantial challenges, including long delays in completing complex risk assessments, some of which take decades to complete; lack of data, which leads to important uncertainty in risk

multiple interpretations and contribute to decision-making gridlock.”³⁹ Thus, major risk assessment exercises for formaldehyde, trichloroethylene, and dioxin dragged on for decades, with many additional chemicals waiting in the queue.⁴⁰

Much of the difficulty involved in completing these risk assessments stemmed from the seemingly irreducible uncertainties involved in such assessments.⁴¹ Abstract notions of populations and averages, extrapolation techniques, and computational models all brought with them uncertainties precisely because they represented formalized, and often highly simplified, versions of complex, open systems. Thus, the entire exercise of determining whether the risk of exposure to a particular carcinogen was “significant”—that it exceeded the one-in-one million threshold, to take the most commonly accepted measure of significance—suggested the possibility of precise quantification and relatively simple classification of risks despite the fact that the threshold itself, as various commentators had observed, could be calculated in an almost infinite number of ways depending, for example, on the choice of animal studies, interpretation of tissue samples, animal-to-human extrapolation methods, exposure data, and assumptions about exposure pathways.⁴² The standard approach to risk assessment also faced considerable challenges in addressing other, more subtle and complex harms associated with industrial chemicals such as endocrine system disruption and neurodevelopmental effects, not to mention the complexities of dealing with multiple, cumulative exposures to various environmental agents and the special sensitivities of vulnerable populations. By the 2010s, the basic risk assessment paradigm needed an overhaul.

assessments; and the need for risk assessment of many unevaluated chemicals in the marketplace and emerging agents.”).

39. *Id.* at 4.

40. *Id.* at 3–4, 17. In the case of the dioxin risk reassessment, although multiple extrapolation models appeared to fit the data equally well, they generated risk estimates that varied by three orders of magnitude. See Peter C. Wright et al., *Twenty-Five Years of Dioxin Cancer Risk Assessment*, 19 NAT. RESOURCES & ENV'T 31, 35 (2005) (discussing range of cancer risk estimates for dioxin using different standards and guidelines for extrapolation from same data).

41. See, e.g., INST. OF MED., ENVIRONMENTAL DECISIONS IN THE FACE OF UNCERTAINTY 5 (2013) (observing that EPA’s “analyses of and concerns about uncertainties have in some cases (such as in the agency’s work involving dioxin contamination) delayed rulemaking” and “some uncertainty analyses have not provided useful or necessary information for the decision at hand”); *id.* at 6 (cautioning “against excessively complex uncertainty analysis” in risk assessments).

42. See NAT’L RESEARCH COUNCIL, *supra* note 38, at 113–19 (discussing uncertainty and variability in various components of risk assessment); JOHN WARGO, OUR CHILDREN’S TOXIC LEGACY: HOW SCIENCE AND LAW FAIL TO PROTECT US FROM PESTICIDES 111–12 (1996) (discussing the “infinite number of ways” that one-in-one-million risk threshold could be calculated depending on choice of animal studies, extrapolation models, exposure data, etc.).

In the view of its proponents, big data promised a way out of this morass—a return to a more experience-based approach to environmental decisionmaking that would allow the data (and the correlations apparent in the data) to speak for themselves. With the ability to identify and trace harms with far more precision than in the past, there would be less need, so the argument went, for complex, resource-intensive risk assessment exercises as a basis for environmental decisionmaking. Rather, each individual exposome could be assessed to determine specific exposures and potential harms, which could then be compared and sorted against those of other individuals by various algorithms.

But the move to individual exposomes undermined the stability of averages and statistical regularities that had underwritten the very concept of risk in the first place. Because each individual had a unique exposure history—a unique environmental history—that could be combined with his or her specific genetic and behavioral data, it became difficult to compare and sort individuals on the basis of prior assumptions tied to the idea of a normal distribution.⁴³ Older categories and assumptions regarding populations and subpopulations gave way to a much more variable and differentiated understanding. As the variability of harm and disease became increasingly personalized—to the point where everyone who had cancer, for example, was viewed as having a distinctive form of the disease—the entire enterprise of environmental health protection lost part of its collective orientation.⁴⁴ In the process, the very notion of acceptable risk began to unravel.

With the advent of big data, then, environmental law's social imaginary and its subject—the statistical person—gave way to a more granular, increasingly desocialized conception of the individual and of the environmental insults and

43. A similar challenge had long confronted efforts to determine “acceptable risk” for pollutants and toxic substances with multiple health endpoints. Under the Clean Air Act’s NAAQS program, for example, there was no obvious way (no common risk metric) to use diverse health effects such as reduced IQ, angina, or impaired lung capacity as a basis for determining a standard that would protect public health with an adequate margin of safety. See, e.g., John Bachmann, *Will the Circle Be Unbroken: A History of the U.S. National Ambient Air Quality Standards*, 57 J. AIR & WASTE MGMT. ASS’N 652, 690 (2007).

44. As historians of science and medicine have reminded us, disease categories themselves were constructed on the basis of a series of abstractions that aggregated across a diversity of individual circumstances and experiences. See, e.g., Charles E. Rosenberg, *The Tyranny of Diagnosis: Specific Entities and Individual Experience*, 80 MILLBANK Q. 237, 252–53 (2002) (“Agreed upon disease pictures are configured in contemporary medicine around aggregated clinical findings—readings, values, thresholds—whereas therapeutic practice is increasingly and similarly dependent on tests of statistical significance. Yet men and women come in an infinite variety, a spectrum rather than a set of discrete points along that spectrum. An instance of cancer exists, for example, along such a continuous spectrum; the staging that describes and prescribes treatment protocols is no more than a convenience, if perhaps an indispensable one. In this sense, the clinician can be seen as a kind of interface manager, shaping the intersection between the individual patient and a collectively and cumulatively agreed-upon picture of a particular disease and its optimal treatment.”).

harms he or she suffered over his or her lifetime. As billions of data points became available for individuals and the environments through which they moved in their daily lives, governments and their proxies were able to track in almost real time these environmental insults and the harms and consequences that resulted.⁴⁵

The distributional consequences of this shift did not go entirely unnoticed, but for the most part government reports and official proclamations gave little more than lip service to the ethical issues associated with these developments.⁴⁶ Not surprisingly, much of the early exposome data confirmed that exposures and harm often tracked disparities of wealth. Big data thus provided an important means for making environmental injustices visible. But it also converged with and reinforced ongoing tendencies toward exclusion and “responsibilization,” intensifying and deepening existing inequalities in the process.⁴⁷

In many ways, the convergence of the exposome with the widespread embrace of self-tracking and self-measurement should not be surprising. The “quantified self” movement, which started as a path to self-improvement, brought with it new opportunities for environmental health surveillance and new logics of exclusion and control.⁴⁸ As individuals began to compile their own

45. Cf. Andrew Lakoff, *Real-Time Biopolitics: The Actuary and the Sentinel in Global Public Health*, 44 *ECON. & SOC'Y* 40 (2015).

46. See, e.g., NAT'L RESEARCH COUNCIL, *supra* note 20, at 5 (“The availability of the massive quantities of individualized exposure data that will be generated might create ethical challenges and raise issues of privacy protection.”).

47. Cf. Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 *CALIF. L. REV.* 671 (2016). The notion of “responsibilization” was prominent in the extensive literature that emerged in the 1990s around Michel Foucault’s concept of governmentality (“the conduct of conduct”) and its relevance for advanced liberal economies in the late twentieth and early twenty-first centuries. See, e.g., Nikolas Rose, *Government and Control*, 40 *BRIT. J. CRIMINOLOGY* 321, 324 (2000) (“There has been a fragmentation of ‘the social’ as a field of action and thought. . . . The idea of a unified solidary social domain and a single national culture is displaced by images of multiple communities, plural identities, and cultural diversity. A whole range of new technologies—‘technologies of freedom’—have been invented that seek to govern ‘at a distance’ through, not in spite of, the autonomous choices of relatively independent entities. . . . As far as individuals are concerned, one sees a revitalization of the demand that each person should be obliged to be prudent, responsible for their own destinies, actively calculating about their futures and providing for their own security and that of their families with the assistance of a plurality of independent experts and profit-making businesses from private health insurance to private security firms. This alloy of autonomization and responsibilization underpins shifts in strategies of welfare, in which substantive issues of income distribution and poverty have been displaced by a focus upon processual issues that affiliate or expel individuals from the universe of civility, choice and responsibility, best captured by the dichotomy of inclusion and exclusion.”).

48. A few critics warned of some of the problems attending these developments. See Kate Crawford et al., *Our Metrics, Ourselves: A Hundred Years of Self-Tracking From the Weight Scale to the Wrist Wearable Device*, 18 *EUR. J. CULTURAL STUD.* 479, 493–94 (2015) (“[I]n the age of networked connectivity, systems that are supposedly built for you are, in fact, systems about you and the data you provide. . . . [W]hen people start using these devices they enter into a relation that is an

quantified risk profiles, they faced new incentives to further “improve” their lives in yet another example of the pervasive logics of economization and entrepreneurialization that some identified as core to neoliberal governance.⁴⁹ The quantified self soon went from hobby for the gadget obsessed to a virtual prerequisite for access to health care, insurance, and employment (to say nothing of happiness and spiritual fulfillment).⁵⁰

Indeed, as personalized health and environmental data became available for millions of individuals, new programs and incentives were established by employers, health care providers, insurance companies, and governments to encourage people to share their data as a basis for precisely tailored health care options, individualized insurance plans, and personal improvement strategies. This further strained the implicit solidarity that underwrote basic tenets of the welfare state.⁵¹ The whole idea of the public—public health, public welfare, the public interest—and of individuals as members of society was thinned out and diminished, replaced with a new “individualism of singularity” that carried with it a

inherently uneven exchange—they are providing more data than they receive, and have little input as to the life of that data—where it is stored, whether it can be deleted and with whom it is shared. They are also becoming part of an aggregated data set that is compared against other forms of data: medical and otherwise. Just as the Association of Life Insurance Medical Directors of America devised standardized height and weight charts in the late 1800s, now companies like Vivametrica are vying to become the standard-setters for wearable device data. There is considerable power in becoming the standard-setter of what makes a ‘normal’ user. . . . While self-knowledge may be the rhetoric of wearable device advertising, it is just as much a technology of *being known by others*. With more detailed information, far more individualized and precise interventions can be conducted. . . .”).

49. See, e.g., WENDY BROWN, *UNDOING THE DEMOS: NEOLIBERALISM'S STEALTH REVOLUTION* 36 (2015) (“This subtle shift from exchange to competition as the essence of the market means that all market actors are rendered as little capitals (rather than as owners, workers, and consumers) competing with, rather than exchanging with each other. Human capital's constant and ubiquitous aim, whether studying, interning, working, planning retirement, or reinventing itself in a new life, is to entrepreneurialize its endeavors, appreciate its value, and increase its rating or ranking.”).
50. Deborah Lupton, *The Diverse Domains of Quantified Selves: Self-Tracking Modes and Dataveillance*, 45 *ECON. & SOC'Y* 101, 103 (2016) (“Self-tracking at first glance appears to be a highly specialized subculture, confined to the chronically ill, obsessives, narcissists or computer geeks or simply people who are already interested in optimizing their health, physical fitness, and productivity. . . . [But] this form of dataveillance is now being used in situations where the choice to participate may be limited. The concept and practices of self-tracking are now dispersing rapidly into multiple social domains.”).
51. See PIERRE ROSANVALLON, *THE SOCIETY OF EQUALS* 211 (Arthur Goldhammer trans., Harvard Univ. Press 2013) (2011) (“The implicit principle of justice and solidarity inherent in the welfare state rested on the idea that risks were equally shared and essentially random.”).

hyper-individualized understanding of inequality and a strong normative commitment to individual responsibility.⁵²

In the process, however, big data and the information derived from it about individual exposomes and disease profiles did create the basis for new affinity groups and novel political agendas.⁵³ Similar in some ways to past patient activism organized around rare diseases, this new “politics of singularization” drew upon the shared sense of vulnerability and harm that came with new, more granular understandings of individual risk.⁵⁴ Big data also held out the possibility of leveraging previous modes of “citizen science” and “popular epidemiology” in the service of environmental justice claims advanced by disadvantaged communities.⁵⁵ But only rarely did these forms of activism coalesce into broader movements for systemic reform.

-
52. *Id.* at 227 (“Inequalities are now as much the result of individual situations, which are becoming more diverse, as of social conditions, which reproduce themselves. . . . They are ambiguous in nature. At times, they are more bitterly resented than other inequalities because they are attributed to personal failure or lack of ability. They lack the clearly objective and therefore psychologically ‘reassuring’ character of traditional inequalities of condition. Even if they can also be blamed on injustice or misfortune, they are still associated with an idea of responsibility. Indeed, responsibility automatically comes to the fore in a society that values singularity, as both constraint and positive value.”); see also Francois Ewald, *Omnes Et Singulatim. After Risk*, 7 CARCERAL NOTEBOOKS 77, 77–83 (2011) (discussing shift from a notion of solidarity as a basis for risk and the welfare state to a world of differentiation and singularity in the era of big data).
 53. See, e.g., Kim Fortun et al., *Pushback: Critical Data Designers and Pollution Politics*, July–Dec. 2016 BIG DATA & SOCIETY 1 (discussing ways in which big data can support new political agendas and advance efforts to understand, map, and control pollution). During the 1990s and 2000s, there was also a wide-ranging discussion of new “biosocial identities” made possible by genomics technologies. The advent of big data and attention to the exposome pushed these tendencies further, giving rise to new affinity groups and new bases for activism. See, e.g., Daniel Nevon, *Genomic Designation: How Genetics Can Delineate New, Phenotypically Diffuse Medical Categories*, 41 SOC. STUD. SCI. 203, 219 (2011) (discussing “a new kind of bioscientific-social fact, the genomic designation of medical conditions, whereby genetic mutations can give rise to new, otherwise unthinkable kinds of people”); Ian Hacking, *Genetics, Biosocial Groups, and the Future of Identity*, 135 DAEDALUS 81, 91 (2006) (“A set of people with a risk factor is a biological, not social, group. But people at risk for the same disease will clump together for mutual support, joint advocacy, and, in many cases, activism. The emergence of these advocacy groups will be one of the most important topics for any history of medicine in late twentieth-century America.”).
 54. See Vololona Rabearisoa et al., *From ‘Politics of Numbers’ to ‘Politics of Singularisation’: Patients’ Activism and Engagement in Research on Rare Diseases in France and Portugal*, 9 BIOSOCIETIES 194 (2014).
 55. See Alice Mah, *Environmental Justice in the Age of Big Data: Challenging Toxic Blind Spots of Voice, Speed and Expertise*, ENVTL. SOCIOLOGY 2 (2016) (“Environmental justice advocates have already adopted many citizen science techniques that could be classified, broadly speaking, as part of the big data phenomenon.”).

V. ACCOUNTABILITY AND EXPERTISE

At the same time that big data was undermining traditional notions of risk and reorienting the practice of risk assessment, it was also reinforcing an ongoing shift in conventional understandings of accountability and expertise in administrative law. Put simply, the explosion of new facts and the proliferation of new modes of fact-making made possible by big data—marked most prominently by the increasing reliance upon complex (often proprietary) algorithms and expert systems to translate massive volumes of data into actionable information about the problem at hand—further diminished received understandings of expert judgment, reason, and deliberation.

To be sure, this trend had been underway well before the advent of big data and was hardly exclusive to environmental law.⁵⁶ In many respects, the entire move to quantitative risk assessment that began in the late 1970s represented an effort to replace individual expert judgment with more rule-bound and more formal approaches to decisionmaking.⁵⁷ As EPA and other agencies turned to increasingly sophisticated computational models to carry out their responsibilities during the 1980s and 1990s, moreover, the whole notion of expertise was further diminished and displaced.⁵⁸

Big data underwrote a further intensification of this ongoing process of displacement.⁵⁹ By substituting “data-knowledge” for explanation and judgment, it

56. In fact, these trends drew on deeper tendencies toward calculation and control that had been underway throughout the second half of the twentieth century—during the so-called Cold War—manifest in an ongoing formalization of knowledge practices across multiple disciplines and an expanding administrative state. See, e.g., PAUL ERICKSON ET AL., *HOW REASON ALMOST LOST ITS MIND: THE STRANGE CAREER OF COLD WAR RATIONALITY* 2–4 (2013) (describing Cold War rationality as formal, algorithmic, reductionist, ahistorical and acontextual, and amenable to rules that could be applied by computers); see also Julie E. Cohen, *The Regulatory State in the Information Age*, 17 *THEORETICAL INQUIRIES L.* 369 (2016) (discussing challenges facing regulatory institutions in the world of “informational capitalism”).

57. See Boyd, *supra* note 15, at 964–83 (tracing these developments).

58. See, e.g., NAT’L RESEARCH COUNCIL, *supra* note 23, at 193 (noting that “ever-larger and more-sophisticated models may not necessarily make better regulatory tools” and discussing “the possibility that pursuing larger and more-sophisticated models make them less and less able to be evaluated and more impenetrable to the public and decision makers”); see also Paul Humphreys, *The Philosophical Novelty of Computer Simulation Methods*, 169 *SYNTHESE* 615, 617 (2009) (“For an increasing number of fields in science, an exclusively anthropocentric epistemology is no longer appropriate because there now exist superior, non-human epistemic authorities. So we are now faced with a problem, which we can call the *anthropocentric predicament*, of how we, as humans, can understand and evaluate computationally based scientific methods that transcend our own abilities.”).

59. This process of displacement reflected a more general unease over the ways in which deliberative reason was being crowded out and superseded by a stricter rule-governed

marked a culmination of these trends—a definitive and humbling end to the technocratic ideal that had informed basic understandings of expertise and the regulatory state throughout the second half of the twentieth century.⁶⁰

As regulatory decisionmaking was subsumed by a proliferation of expert systems and algorithms, a handful of critics raised questions about the implications for transparency and accountability. What, these critics asked, did transparency really mean in a world of big data? Who exactly was accountable for regulatory science when the algorithms and data analytics tools exceeded the evaluative capacities of the regulators and the scientists? How were courts going to ensure that agencies had in fact taken a hard look when the agencies themselves were not able to fully evaluate the tools they used?⁶¹

Despite EPA's efforts to rationalize and consolidate its data enterprise, it had little choice but to rely on third parties and proprietary algorithms for much of its data analysis. This proved to be one of the more important paradoxes of big data and environmental law. On the one hand, big data promised to dramatically reduce the uncertainties that had plagued prior approaches to risk assessment by developing a much more precise, empirically grounded understanding of environmental exposure and harm. But on the other hand it put considerable, largely unreviewable power in the hands of algorithms.

In an odd and somewhat ironic sense, then, there was a double movement at work. Big data and algorithmic rationality allowed for more fine-grained attention to the singularity of individual lives at the same time that they were

rationality. The well-known twentieth century philosopher Robert Nozick had already remarked upon this very possibility in 1993:

In the study of reliable processes for arriving at belief, philosophers will become technologically obsolescent. They will be replaced by cognitive and computer scientists, workers in artificial intelligence, and others. . . . This will be useful to us—machines will be produced to do intricate tasks—but it will not be what philosophers earlier had hoped for: rules and procedures that we ourselves could apply to better our own beliefs, surveyable rules and procedures—I take the term from Ludwig Wittgenstein—that we can take in and understand as a whole and that give us a structurally revealing description of the nature of rationality.

ROBERT NOZICK, *THE NATURE OF RATIONALITY* 76 (1993); see also Lorraine Daston, *Simon and the Sirens: A Commentary*, 106 *ISIS* 669, 675 (2015) (citing Nozick and noting that “the rules and procedures that constitute rationality would be valid and efficacious but algorithmic—efficiently and reliably executed by a machine but opaque to human understanding”).

60. A handful of legal and social science scholars remarked on this phenomenon as early as the 2010s. See, e.g., Annelise Riles, *Market Collaboration: Finance, Culture, and Ethnography After Neoliberalism*, 115 *AM. ANTHROPOLOGIST* 555, 561 (2013) (identifying the turn to big data as a “reflection of a loss of confidence in prior forms of expertise on which models, simulations, and samples are premised”). This growing displacement of expertise, including loss of “control even over the hypothesis,” represented a “profoundly humbling moment in the history of expertise.” *Id.*

61. Similar questions regarding the unreviewability of certain decisions (and the broader frameworks for decisionmaking) emerged in other legal domains—from criminal law to financial regulation.

displacing the ability of human beings (experts, regulators, the public) to understand and review the rules and procedures that revealed these details.⁶² Put another way, the deep “black-boxing” of the ways in which knowledge about individual lives was being produced for environmental decisionmaking made the very idea of accountability (and reviewability) tenuous at best, raising profound questions about the meaning of expertise and “epistemic competence.”⁶³

Contemporary observers, of course, did raise concerns about the proliferation of algorithms (what some called “algorithmic power”) and the considerable influence that third-party analytics firms enjoyed as gatekeepers to the insights available with big data.⁶⁴ Scholars working in the emerging field of “critical data studies” also questioned the epistemic and normative implications of these trends.⁶⁵ But the power of these tools and the manner in which they fit with ongoing tendencies to privatize and outsource core government responsibilities went largely unheeded by policymakers.⁶⁶ By the 2030s, earlier conceptions of expertise, accountability, and agency review had limited purchase in the face of tools and techniques that exceeded the evaluative capacities of the experts themselves.⁶⁷ The delicate and sometimes fraught relationship between expertise and the public that had animated foundational

-
62. See Neil M. Richards & Jonathan H. King, *Three Paradoxes of Big Data*, 66 STAN. L. REV. ONLINE 41, 42–43 (2013) (“This is the Transparency Paradox. Big data promises to use this data to make the world more transparent, but its collection is invisible, and its tools and techniques are opaque, shrouded by layers of physical, legal, and technical privacy by design.”).
 63. Cf. Jennifer L. Mnookin, *Expert Evidence, Partisanship, and Epistemic Competence*, 73 BROOK. L. REV. 1009 (2008).
 64. See, e.g., Daniel Neyland, *On Organizing Algorithms*, 32 THEORY CULTURE & SOC’Y 119 (2015); Malte Ziewitz, *Governing Algorithms: Myth, Mess, and Methods*, 41 SCI. TECH. & HUM. VALUES 3 (2016).
 65. See, e.g., Andrew Illiadis & Federica Russo, *Critical Data Studies: An Introduction*, July–Dec. 2016 BIG DATA & SOCIETY 2; Kate Crawford et al., *Critiquing Big Data: Politics, Ethics, Epistemology*, 8 INT’L J. COMM. 1663 (2014).
 66. There is a large literature, going back to the late twentieth century, on government outsourcing. For an introduction and overview, see GOVERNMENT BY CONTRACT: OUTSOURCING AND AMERICAN DEMOCRACY (Jody Freeman & Martha Minow eds., 2009); see also Jon D. Michaels, *Privatization’s Pretensions*, 77 U. CHI. L. REV. 717 (2010).
 67. Although this trend had been underway for some time, big data created new challenges for traditional understandings of expert knowledge and judgment. See, e.g., John Symons & Ramon Alvarado, *Can We Trust Big Data? Applying Philosophy of Science to Software*, July–Dec. 2016 BIG DATA & SOCIETY 1, 8–9 (discussing problems of “epistemic opacity” created by big data’s reliance on software and computational methods that are inaccessible to traditional forms of human-centered inquiry).

questions about the roles and functions of government throughout the twentieth century seemed strange and out of place in a world where our tools for producing knowledge had become semiautonomous.⁶⁸

VI. VIOLENCE, MEMORY, RESPONSIBILITY

The early promise of U.S. environmental law was reflected in the protective, precautionary language of statutes such as the Clean Air Act, in decisions by the newly established EPA to ban certain pesticides and to impose tough pollution controls on industry, and in early appellate court decisions giving EPA wide discretion to carry out its responsibilities.⁶⁹ At its most fundamental, that promise was to protect public health from unreasonable risk; to ensure that, in the face of uncertainty, standards would be set and regulations imposed at levels necessary to provide an adequate margin of safety. Protecting human health (human life) from poorly understood, even unknown potential harms was seen as well within the bounds of EPA's responsibilities. Precaution in the face of uncertainty provided the dominant impulse.⁷⁰

Over time, and in the face of an increasingly complex and expansive world of potential environmental hazards, the strong precautionary impulse was replaced with a more formal approach to quantifying risk as a basis for making threshold determinations of significance prior to regulating.⁷¹ This reflected in part the influence of the regulated community and a growing distrust of regulatory agencies such as the EPA, but it also reflected fundamental changes in the

-
68. In his reflections on the role of experts in public life, John Dewey observed that:
[E]xpertness is not shown in framing and executing policies but in discovering and making known the facts upon which the former depend . . . It is not necessary that the many should have the knowledge and skill to carry on the needed investigations; what is required is that they have the ability to judge of the bearing of the knowledge supplied by others upon common concern.
JOHN DEWEY, *THE PUBLIC AND ITS PROBLEMS* 208–09 (1927). Is this possible in a world of big data?
69. See Boyd, *supra* note 15, at 948–63 (discussing these developments).
70. See Kysar, *supra* note 37.
71. See Boyd, *supra* note 15, at 964–83 (tracing this shift); see also *Industrial Union Dep't, AFL-CIO v. Am. Petroleum Inst.*, 488 U.S. 607, 614–16 (1980) (requiring OSHA to make a threshold finding of “significant risk” before regulating). The *Benzene* decision, as it is known, is widely viewed as giving a major boost to quantitative risk assessment in health safety, and environmental law. See, e.g., THOMAS MCGARITY, *THE STORY OF THE BENZENE CASE: JUDICIALLY IMPOSED REGULATORY REFORM THROUGH RISK ASSESSMENT IN ENVIRONMENTAL LAW STORIES* 143 (Lazarus & Houck eds., 2005) (noting that the *Benzene* decision “came at a critical inflection point in the historical flow of U.S. environmental law and . . . had a profound impact on that body of law”).

underlying knowledge practices and the need for new tools in the face a vast expansion of statutory responsibilities.

The limits of this approach became increasingly apparent by the 2000s as EPA's risk assessment efforts struggled in the face of seemingly irreducible uncertainties and a lack of information on actual exposure pathways and the environmental etiology of disease. Over the next quarter century, the risk assessment framework gradually gave way to a new approach founded on big data. As the highly formalized, abstract models of environmental systems and human health that had underwritten risk assessment in the past were replaced with massive flows of data on actual individual lives and harms, big data promised to put environmental law on more solid, empirical ground.

But as the subtle harms and slow violence of various environmental exposures became apparent for millions of individuals, the fragile solidarity that had supported environmental law for more than half a century began to fray. Questions about structural inequalities and their connections to environmental injustice that had once resonated as profound ethical challenges to environmental law were clouded and confused by the explosion of data showing a much more heterogeneous and individually variable world of environmental harms. This variability, in fact, rendered the very concept of public health—and the idea that environmental protection should be viewed as a collective undertaking—increasingly tenuous.

The normative implications of this shift became the subject of intense debate during the 2030s. Where some argued in favor of expanding environmental regulation to deal more comprehensively with the subtle harms visited upon individuals as a means to restore the *public* in public health, others argued that these individuals should take primary responsibility for improving their own lives and avoiding exposures. In the middle were those promoting a more pervasive form of paternalism—one that replaced the crude behavioralism of nudges with a much more precise set of incentives and commands tailored to a regimen of self-improvement.

Of course, far too many people had very little ability to avoid many of the environmental exposures they confronted in their daily lives, and many of them had only limited, uneven access to the kind of high-quality exposome data that might allow them to do so even if they had the means.⁷² In the age of big data,

72. Concerns about the exclusionary potential of big data and its disparate impacts had been voiced since the 2010s. See, e.g., Mah, *supra* note 55, at 3 ("The big data movement is anchored in a growing fascination with the 'quantified self.' . . . Thus, the ethos of big data is entangled with the consumerism and individualism that underpins neoliberal capitalism, and sits uncomfortably alongside the concerns of environmental justice activists, rooted in disadvantaged areas without

these people found themselves on the margins of environmental protection and public health and, at the same time, on the margins of debates about the future of environmental regulation. As such, they were at risk of becoming an invisible, forgotten underclass—a residual category comprised of those unable or unwilling (in the eyes of the more fortunate) to take responsibility for their own lives.⁷³

Environmental law in 2030 thus stood at a crossroads. For the first time, it had the tools to identify and make visible (and take responsibility for) the many individual lives cut short, stunted, or diminished in some way by the environmental harms it was charged with regulating. But the burdens that came with these new capabilities seemed almost overwhelming in a world of limited resources and diminished support for government. As mid-century approached, environmental law confronted a profoundly unsettling set of questions regarding the uneven, real life effects of its decisions: Did it have an obligation not only to acknowledge but to account for (that is, to be accountable for) those lives? How much closer could it hope to get—how much closer should it hope to get—to the singularity of those lives?

access to the latest sensing devices.”); Lupton, *supra* note 50, at 116 (“Self-tracking data can be mobilized as surveillant technologies in ways that further entrench the social disadvantage of marginalized groups.”); Brent David Mittelstadt & Luciano Floridi, *The Ethics of Big Data: Current and Forseeable Issues in Biomedical Contexts*, 22 SCI. & ENGINEERING ETHICS 303, 322–25 (2016) (discussing the “big data divide” and its implications for “underprivileged” and “data poor” individuals); Jonas Lerman, *Big Data and Its Exclusions*, 66 STAN. L. REV. ONLINE 55, 56 (2013) (“Big data poses risks . . . to those persons who are *not* swallowed up by it—whose information is not regularly harvested, framed, or mined. . . . Although proponents and skeptics alike tend to view this revolution as totalizing and universal, the reality is that billions of people remain on its margins because they do not routinely engage in activities that big data and advanced analytics are designed to capture.”).

73. Unable to use the resources of big data to compile and maintain their own personalized digital archives, these people were increasingly at risk of losing control of their own histories—of being erased from official forms of memory and, as a result, limited in the kinds of claims they could make against state actors and other powerful interests. See Yuk Hui, *A Contribution to the Political Economy of Personal Archives*, in COMPROMISED DATA: FROM SOCIAL MEDIA TO BIG DATA 226 (Longolis et al. eds., 2015) (arguing that in the age of big data “everyone must be their own digital archivist”); Joan M. Schwartz & Terry Cook, *Archives, Records, and Power: The Making of Modern Memory*, 2 ARCHIVAL SCI. 1, 18 (2002) (“Memory, like history, is rooted in archives. . . . Archives validate our experiences, our perceptions, our narratives, our stories. Archives are our memories. Yet what goes in the archives remains largely unknown. Users of archives (historians and others) and shapers of archives (records creators, records managers, and archivists) add layers of meaning, layers which become naturalized, internalized, and unquestioned. This lack of questioning is dangerous because it implicitly supports the archival myth of neutrality and objectivity, and thus sanctions the already strong predilection of archives and archivists to document primarily mainstream culture and powerful records creators. It further privileges the official narratives of the state over the private stories of individuals.”).