Tennessee Law Review

Volume 86
Issue 4 *Summer 2019*

Article 6

2019

BIG DATA IS NOT BIG OIL: THE ROLE OF ANALOGY IN THE LAW OF NEW TECHNOLOGIES

Lauren Henry Scholz

BIG DATA IS NOT BIG OIL: THE ROLE OF ANALOGY IN THE LAW OF NEW TECHNOLOGIES

LAUREN HENRY SCHOLZ[*]

INTRODUCTION

The idea of data as the petroleum oil of the information economy might not be familiar to every reader, but the analogy's underlying influence is ubiquitous. The analogy is palpable in the way the general public discusses the modern landscape. We speak of data being "mined." We talk of data "brokers," which entirely sidesteps the question of how these entities acquire the underlining resource they sell. We value companies based on the information they have, with little or no inquiry into how they acquire it. The modern data economy presumes that data does not come from anywhere, or that wherever it comes from is insignificant.

Superficial similarities exist between oil and data. Oil powers the industrial economy.[1] Data powers the information economy. Companies use oil to produce physical goods. Companies use big

---

* Assistant Professor, Florida State University College of Law. The author is grateful to Mark Lemley, Greg Rosston, Eric Goldman, and participants in the Stanford-Penn-Northwestern Junior Faculty Forum for Law and STEM for comments on earlier versions of this Article.

1. This Article will use oil as the emblematic analogy of natural resources to data. Commentators have compared data to other natural resources such as coal or gold, but oil is the typical comparator. In any case, the comparison is at a level of generality that each of these natural resources serves the same purpose.

data[2]—or, put more simply, extensive, complex databases—to make customized predictions.[3] In an economy with both industrial and information elements, data is valuable and so is oil. However, although oil may be central to the industrial economy, and data is central to the information economy, centrality alone does not make them comparable as a matter of law or policy. Just because a resource is central to the functioning of the economy does not mean that even the most basic goals of society—for example, peace, safety, and order—can be achieved by similar social, economic, and legal structures.[4]

Data differs from oil in meaningful ways.[5] Oil is the byproduct of

---

2. Big data itself is a contested term with multiple definitions. *See The Big Data Conundrum: How to Define It?*, MIT TECH. REV. (Oct. 3, 2013), https://www.technologyreview.com/s/519851/the-big-data-conundrum-how-to-define-it/ (describing a survey of incompatible industry definitions of big data); *see also* Margaret Hu, *Small Data Surveillance v. Big Data Cybersurveillance*, 42 PEPP. L. REV. 773, 794 (2015) (discussing a definition of big data).

3. For example, many email providers use proprietary algorithms to provide targeted advertisements to their users based on their emailing activity. *See* Daisuke Wakabayashi, *Google Will No Longer Scan Gmail for Ad Targeting*, N.Y. TIMES (June 23, 2017), https://www.nytimes.com/2017/06/23/technology/gmail-ads.html (describing Google's former email scanning and ad targeting practices, which remain the norm for email providers without Alphabet's (Google's parent company) business cloud conflict of interest). More targeted ads are not necessarily more effective. *See* Gian M. Fulgoni, *Are You Targeting Too Much? Effective Marketing Strategies for Brands*, 58 J. ADVERT. RES. 8, 9 (2018). Targeted ads also have complex relationships to consumers, who can feel judged and uncomfortable by targeted ads. *See* Sammi Caramela, *Are You Creeping Out Your Customers? Here's How to Stop*, BUS. NEWS DAILY (Mar. 11, 2018), https://www.businessnewsdaily.com/7908-creepy-online-ads.html; Leslie K. John, Tami Kim & Kate Barasz, *Ads that Don't Overstep*, HARV. BUS. REV., https://hbr.org/2018/01/ads-that-dont-overstep (last visited Feb. 15, 2020).

4. After all, land was central to the pre-industrial economy and has never ceased to be valuable. However, real property is governed in a different way from fossil fuels and indeed, all other forms of movable personal property.

5. In this Article, I will directly compare the properties of data to oil, limiting the use of the "big" modifier for each. First, this avoids definitional ambiguity, because the precise point at which "small data" becomes big data in unclear. *See* Hu, *supra* note 2, at 798. Further, for many commentators, the phrase relates more to how the data is used rather than the size of the data set. *See* Woodrow Hartzog & Evan Selinger, *Big Data in Small Hands*, 66 STAN. L. REV. ONLINE 81, 81 (2013) ("'Big data' can be defined as a problem-solving philosophy that leverages massive datasets and algorithmic analysis to extract 'hidden information and surprising correlations.'"). Second, and more fundamentally, the question this paper examines is whether oil and data have similar enough properties in themselves to justify similar regulatory measures. For this purpose, including bigness of each industry in the analysis conceals more than it reveals, as most major areas of commerce have certain characteristics in common, such as complexity and the potential for economic and political industry

animal corpses and plant matter from several millennia ago.[6] By contrast, human and corporate actors produce data about living people for particular purposes.[7] The data-as-oil analogy side-steps evaluation of any misappropriation or exploitation that might arise from data use and processing by adopting an analogy that presumes the history of the data prior to collection is irrelevant. The analogy, in effect, puts the humans about whom the data touches and concerns in the position of long-dead dinosaurs and plants. The analogy gives the data source neutrality and obscurity that cannot be justified by technical facts or legal precedent.

This Article has two theses: (1) data as the oil of the information economy is a bad analogy, as a matter of logic; and (2) data as oil is a misleading and dangerous analogy as applied to law and policy because it does not acknowledge key features of the underlying resource, data, and its function in the economy. Part I traces the origins and characteristics of the data-as-oil analogy. Part II describes analogy's function in argument and persuasion and outlines its particular role in legal argument regarding novel technologies and business models. Part III analyzes the factual limitations of the analogy. Part IV evaluates the particular law and policy implications of the data-as-oil analogy. The final Part discusses possible alternative analogies to help make sense of data's role in the information economy.

This Article contributes to the small but growing literature on interrogating the limits of analogy in the law of new technologies.[8]

---

influence. Insofar as regulatory justifications for data and oil stem from common "bigness," they may well have little to do with the unique characteristics of these industries. My approach factors out generic considerations against or in favor of "big business."

6. *See Petroleum*, NAT'L GEOGRAPHIC, https://www.nationalgeographic.org/encyclopedia/petroleum/ (last visited Feb. 15, 2020).

7. *See, e.g.*, Susel Góngora Alonso et al., *A Systematic Review of Techniques and Sources of Big Data in the Healthcare Sector*, 41 J. MED. SYS. 182, 183 (2017) (discussing sources of big data in healthcare); M.I. Pramanik et al., *Big Data Analytics for Security and Criminal Investigations*, 7 WIREs: DATA MINING & KNOWLEDGE DISCOVERY, July–Aug. 2017, at 2–7 (discussing sources of big data in security); *see also* FED. TRADE COMM'N, DATA BROKERS: A CALL FOR TRANSPARENCY AND ACCOUNTABILITY 11 (2014), https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf (discussing sources of general consumer data via data brokers).

8. *See* Heather M. Whitney & Robert Mark Simpson, Search Engines and Free Speech Coverage, *in* FREE SPEECH IN THE DIGITAL AGE 33, 37–42 (Susan J. Brison & Katharine Gelber eds., 2019) (analyzing the limits of analogical reasoning in

Data, like oil, is valuable and it powers the modern economy. But the analogy, as a matter of law and policy, is misleading and dangerously rash for our developing information society.

I. A HISTORY OF THE DATA-AS-OIL ANALOGY

Data scientist Clive Humby coined the phrase "Big Data is the New Oil" in the year 2006.[9] On a panel at Northwestern University, Humby remarked that data was like oil because it had to be "broken down" to be useful.[10] In 2012, venture capitalist Ann Winblad used the analogy in response to the general question "what's the next big thing?"[11] In 2013, International Business Machines Corporation (IBM) Chief Executive Officer Virginia (Ginni) Rometty used the analogy in remarks to the public about her firm's latest moves to convey how data has to be processed to be useful.[12] Neil Lawrence, a former Director of Machine Learning at Amazon who now teaches at Cambridge, once compared data to coal at the start of the industrial age because of the value and ready availability of coal then and data now.[13] What these and other statements by thought leaders have in

---

information-age discussions of search engines and free speech); Kailash Awati & Simon Buckingham Shum, *Big Data Metaphors We Live by*, TOWARDS DATA SCI. (May 14, 2015), https://towardsdatascience.com/big-data-metaphors-we-live-by-98d3fa44ebf8 (analyzing the impacts and concerns of big data as a metaphor); John Perry Barlow, *The Economy of Ideas*, WIRED (Mar. 1, 1994, 12:00 PM), https://www.wired.com/1994/03/economy-ideas/ (discussing the role of analogy in the tech field); Tim Hwang & Karen Levy, *'The Cloud' and Other Dangerous Metaphors*, ATLANTIC (Jan. 20, 2015), https://www.theatlantic.com/technology/archive/2015/01/the-cloud-and-other-dangerous-metaphors/384518/ (discussing the danger of analogies regarding the Cloud); *see also* Eben Moglen, *The Invisible Barbeque*, 97 COLUM. L. REV. 945 (1997) (discussing the danger of analogies in tech law).

9. Humby used the phrase on a panel at a 2006 marketing conference held at Northwestern's Kellogg School of Management. Michael Palmer, *Data Is the New Oil*, ANA MARKETING MAESTROS (Nov. 3, 2006, 5:43 AM), https://ana.blogs.com/maestros/2006/11/data_is_the_new.html. The talk is listed on Humby's LinkedIn page, where Humby also claims credit for coining the phrase "Data is the New Oil." Clive Humby OBE, LINKEDIN, https://www.linkedin.com/in/clivehumby/ (last visited Jan. 6, 2020).

10. Palmer, *supra* note 9.

11. Perry Rotella, *Is Data the New Oil?*, FORBES (Apr. 2, 2012, 11:09 AM), https://www.forbes.com/sites/perryrotella/2012/04/02/is-data-the-new-oil/#4f2665fa7db3.

12. *See* Maria Deutscher, *IBM's CEO Says Big Data Is Like Oil, Enterprises Need Help Extracting the Value*, SILICONANGLE (Mar. 11, 2013), https://siliconangle.com/2013/03/11/ibms-ceo-says-big-data-is-like-oil-enterprises-need-help-extracting-the-value/.

13. *See* Neil Lawrence, *Deep Learning, Pachinko, and James Watt: Efficiency Is*

common is that they do not explain in detail what they mean by the analogy or how far the analogy extends. Their goal was likely to evoke the value, excitement, and potential in data for companies to a generalist audience in an environment where data's value was not yet obvious.[14]

Comparing data to oil is a way to highlight data's importance in broad business discussions, gesturing that data is the next big thing in terms of investment and value.[15] The role of press releases, for example, is to build and maintain trust, especially when businesses are pursuing strategies they know will be unfamiliar but believe will be beneficial to consumers.[16] With the history of the analogy's most notable uses revealed, it is fair to say that it did not arise from careful reference to active business practice or scholarly study.[17]

Nonetheless, regulators realized what these corporate leaders were saying and began using the analogy as a starting point for understanding the burgeoning use of data in the economy. For example, in 2011, the World Economic Forum discussed personal data as the "new oil—a valuable resource of the 21st century."[18] The former European Commissioner, Neelie Kroes, responsible for the Digital Agenda,[19] proclaimed that "[j]ust as oil was likened to black gold, data takes on a new importance and value in the digital age."[20]

---

the Driver of Uncertainty, INVERSEPROBABILITY.COM (Mar. 4, 2016), http://inverseprobability.com/2016/03/04/deep-learning-and-uncertainty.

14. *See generally* Brian Bergstein, *How to Think Like a Futurist*, MIT TECH. REV. (Dec. 28, 2016), https://www.technologyreview.com/s/603207/how-to-think-like-a-futurist/ (discussing the implications of future technology).

15. *See* Rotella, *supra* note 11 (Ann Winbald stated that data was the new oil in response to the general question "what's the next big thing?").

16. *Cf.* Elaine Henry, *Are Investors Influenced by How Earnings Press Releases Are Written?*, 45 INT'L J. BUS. COMM. 363, 363 (2008) (discussing the alternative use of press releases at the end of each financial quarter to highlight financial performance).

17. *See* Glyn Moody, *Going with the Flow: The Global Battle for Your Personal Data*, ARSTECHNICA (Nov. 21, 2016, 2:51 AM), http://arstechnica.co.uk/tech-policy/2016/11/eu-us-personal-data-flows-explainer/ (portraying this theme as having been taken up over and over again, so much so that "data is the new oil" has been labeled a "cliché," dating back to at least 2006).

18. *See generally* WORLD ECON. FORUM, PERSONAL DATA: THE EMERGENCE OF A NEW ASSET CLASS 5 (2011), http://www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf.

19. *See* Shellie Karabell, *Neelie Kroes on Women, Startups, and the Digital Agenda*, FORBES (Aug. 30, 2016, 6:01 PM), https://www.forbes.com/sites/shelliekarabell/2016/08/30/nee-lie-kroes-on-women-startups-and-the-digital-agenda/#49c5e04f7a1b.

20. Neelie Kroes, Vice-President, European Comm'n, Opening Remarks, Press Conference on Open Data Strategy, Data is the New Gold (Dec. 12, 2011).

In response, in the 2010s journalists wrote articles discussing the analogy generally,[21] as well as opinion pieces in support of[22] and against[23] the analogy. For the first time, the statement was being analyzed—albeit within the constraints of short-form pieces—rather than merely viewed as an evocative aphorism. The analogy's use in speeches by leaders of prominent companies[24] gave the analogy initial credibility among the public and by extension among some policymakers, stakeholders, and academics outside of the field.[25] At first, it was taken for granted that merely because the analogy was used often it made sense.[26] For example, the Code of Federal Regulations' Medicaid fraud rules define and employ the term "data mining" to describe certain ways of discovering fraud.[27] Additionally,

---

21. *See, e.g.*, Charles Arthur, *Tech Giants May Be Huge, but Nothing Matches Big Data*, GUARDIAN (Aug. 23, 2013, 3:21 PM), http://www.theguardian.com/technology/2013/aug/23/tech-giants-data; Rotella, *supra* note 11.

22. *See, e.g.*, Alex Kantrowitz, *Why Data Is the New 'Oil'*, ADAGE (Oct. 28, 2014), http://adage.com/article/dataworks/data-oil/295565/; Joris Toonders, *Data Is the New Oil of the Digital Economy*, WIRED, http://www.wired.com/2014/07/data-new-oil-digital-economy/ (last visited Jan. 10, 2020); Jonathan Vanian, *Why Data Is the New Oil*, FORTUNE (July 11, 2016), http://fortune.com/2016/07/11/data-oil-brainstorm-tech.

23. *See, e.g.*, Antonio García Martínez, *No, Data Is Not the New Oil*, WIRED (Feb. 26, 2019, 7:00 AM), https://www.wired.com/story/no-data-is-not-the-new-oil/; Simon Moss, *Big Data: New Oil or Snake Oil?*, WIRED, https://www.wired.com/insights/2014/10/big-data-new-oil-or-snake-oil/ (last visited Jan. 10, 2020) (discussing the value of big data in the context of this analogy); Jer Thorp, *Big Data Is Not the New Oil*, HARV. BUS. REV. (Nov. 30, 2012), https://hbr.org/2012/11/data-humans-and-the-new-oil.

24. *See, e.g.*, Susie Gharib, *Intel CEO Says Data Is the New Oil*, FORTUNE (June 7, 2018), https://fortune.com/2018/06/07/intel-ceo-brian-krzanich-data/; David Reid, *Mastercard's Boss Just Told a Saudi Audience that 'Data Is the New Oil'*, CNBC (Oct. 24, 2017, 6:58 AM), https://www.cnbc.com/2017/10/24/mastercard-boss-just-said-data-is-the-new-oil.html.

25. *See, e.g.*, WORLD ECON. FORUM, *supra* note 18, at 5; Uthayasankar Sivarajah et al., *Critical Analysis of Big Data Challenges and Analytical Methods*, 70 J. BUS. RES. 263, 264 (2017) (stating that big data "has been regarded as today's Digital Oil").

26. The mere-exposure effect is a *psychological* phenomenon by which people tend to develop a preference for *things* merely because they are familiar with them. In social *psychology*, this effect is sometimes called the familiarity principle. *See* R.B. Zajonc, *Mere Exposure: A Gateway to the Subliminal*, 10 CURRENT DIRECTIONS IN PSYCHOL. SCI. 224, 225 (2001). This is the case even if the repeated information is an outright lie. *See* Lisa K. Fazio et al., *Knowledge Does Not Protect Against Illusory Truth*, 144 J. EXPERIMENTAL PSYCHOL. 993, 993 (2015); Tom Stafford, *How Liars Create the 'Illusion of Truth'*, BBC FUTURE (Oct. 26, 2016), http://www.bbc.com/future/story/20161026-how-liars-create-the-illusion-of-truth.

27. *See* 42 C.F.R. § 1007.1 (2019).

the practice of using data collected from the internet is framed by universities' institutional review boards—entities that evaluate the propriety of use of subjects in research—as data mining.[28]

Beyond the occasional broad gesturing at its use in the industry, law professors and legislators have ignored the analogy. Only two articles have discussed the implications of the data-as-oil analogy at length. In a *Maine Law Review* article, Dennis Hirsch argues that emphasizing the negative aspects of the data-as-oil metaphor could have salutary effects on how we think about data regulation.[29] In an article about free speech in the algorithmic economy, Jack Balkin also characterizes data as the new oil but is less rosy about the implications of that analogy.[30] He argues that the processing of

---

28. *See* Lauren B. Solberg, *Data Mining on Facebook: A Free Space for Researchers or an IRB Nightmare?*, 2010 U. ILL. J.L. TECH. & POL'Y 311, 317–20.

29. Dennis D. Hirsch, *The Glass House Effect: Big Data, the New Oil, and the Power of Analogy*, 66 ME. L. REV. 373, 377 (2014) ("Big Data and data analytics will create many important benefits for society. The positive side of the 'Big Data is the new oil' analogy is, in many ways, an accurate comparison. However, it is vital also to appreciate the negative dimension of the analogy—the comparison between Big Data's privacy impacts and oil pollution. As explained above, it is only by doing so that society will be able to unlock Big Data's great potential. The positive and negative dimensions of the analogy are linked: in order to have the first, one must also explore and address the second."). The article was written in a political environment where the Environmental Protection Agency's policies and very existence were heavily in question; therefore, one wonders if this is really an analogy that would serve to create accountability from the negative externalities arising from data usage. *See generally* JAMES E. MCCARTHY & CLAUDIA COPELAND, CONG. RESEARCH SERV., R41561, EPA REGULATIONS: TOO MUCH, TOO LITTLE, OR ON TRACK? 1 (2016) (discussing EPA regulations and the resulting controversy). On a deeper level, the environmental regulation system has always been highly contested from the left and right. *See, e.g.*, *id.* A highly contested and controversial regulatory system may not be useful in figuring out how to build a new regulatory structure. Hirsch focuses on the analogy between pollution from fossil fuels and "pollution" from data business practices. *See* Hirsch, *supra*, at 375–76. The risk to the public from both, argues Hirsch, justifies government intervention. *See id.* at 376. So, in this way, Hirsch argues that the data-as-oil analogy is actually a positive for those who would contend that data flows require regulation to protect the interests of individuals and the public interest at large. *See id.* at 376–77. He has continued the analogy in a later work that calls for and describes big data sustainability. *See* Dennis D. Hirsch & Jonathan H. King, *Big Data Sustainability: An Environmental Management Systems Analogy*, 72 WASH. & LEE L. REV. ONLINE 406, 409 (2016).

30. Jack M. Balkin, *Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation*, 51 U.C. DAVIS L. REV. 1149, 1155 (2018) (arguing that data is the new oil because it makes the major industries of the economy run, and that it is a source of power). Katherine Britton is another commentator who accepts the data-as-oil analogy but is concerned about its impact on

personal information by companies "is a central method of governance and control over large populations of people, determining their opportunities and their fates."[31] Balkin outlines the concept of algorithmic nuisance as a way of describing the effect big data users have on individuals in society who fuel the big data machine.[32] Neither Hirsch nor Balkin, however, evaluate the analogy on its logical merits, as this Article does in Part III. Rather, they accept the analogy as conventional wisdom that is shaping the popular discussion of big data and move on to analyzing the implications of that wisdom, each jumping effectively to where this Article goes in Part IV, discussing the potential legal implications of the analogy.

No cases or regulations as of this writing have employed the analogy explicitly. However, in the mere twelve years since its first use, the indirect use of the data-as-oil analogy has deeply permeated American regulation and organizational practice through the recognition of "data mining."[33] Data mining can be defined as "nontrivial process of identifying valid, novel, potentially useful and ultimately understandable patterns in data."[34] The term "data mining" has been popularized as a common term in a way its parent analogy, data-as-oil, has not.[35] This robs citizens and policymakers introduced to the term "data miner" of the ability to evaluate its

---

individuals, stating (putting her own spin on European Commission leader Margrethe Vestager's comments), "[i]f data are the new oil, then like other long-established natural resources, they must be safeguarded, because '[w]hile big data may be the new oil, our personal data are more like weapons-grade plutonium—dangerous, long lasting, and once they are leaked, there's no getting them back.'" Katherine Britton, *Handling Privacy and Security in the Internet of Things*, 19 J. INTERNET L. 3, 6 (2016).

31. Balkin, *supra* note 30, at 1157.

32. *See id.* at 1164–65 (stating that algorithmic nuisance "stems from the fact that companies collect data about people from multiple sources and use algorithms to make decisions about people.").

33. *See* Newton N. Minow & Fred H. Cate, *Government Data Mining*, *in* MCGRAW-HILL HANDBOOK OF HOMELAND SECURITY 1063, 1065–66 (David G. Kamien ed., 2005). As of February 19, 2020, a Westlaw search indicates that the term "data mining" has been cited in 272 cases, 35 regulations, and 68 federal and state statutes.

34. Tal Z. Zarsky, *Governmental Data Mining and Its Alternatives*, 116 PENN ST. L. REV. 285, 291 (2011).

35. In the computer science field, the term "data mining" predates the popularization of the data-as-oil analogy. *See* Ray Li, *History of Data Mining*, KDNUGGETS, https://www.kdnuggets.com/2016/06/rayli-history-data-mining.html (last visited Feb. 15, 2020). The average citizen and policymaker is unlikely to be familiar with the usage of the term as a term of art and will associate it with a broad comparison between oil and data, rather than seeing it as a term of art.

intuitive soundness by explicitly considering the analogy from which it stems.

II. ANALOGY'S FUNCTION IN REASON IN GENERAL VERSUS IN LAW AS A DISCIPLINE

There is comprehensive literature on the relationship between legal and moral reasoning.[36] The main difference between analogy in argument and analogy in legal analysis is that when applied, the latter may ultimately wield the coercive power of the state.[37] Lawyers use applied analogies to understand and address specific problems in the law. The goal of an analogy in other contexts, by contrast, is to illustrate a concept or open up the audience to an alternative point of view. "The law," as Justice Oliver Wendell Holmes wrote, "is not the place for the artist or the poet."[38]

The use of an analogy is a time-honored method of persuasion. An analogy enables a speaker to convey a concept in a way that allows an audience to perceive as they might not have before.[39] For example, in a recent study, two controlled sets of individuals each read one of two stories describing crime in a city: one story analogized crime to a disease and the other story analogized crime to a beast.[40] The group that just read the disease version was relatively more receptive to holistic curative solutions to crime management.[41] By contrast, the group that just read the beast version were relatively more receptive to punitive solutions.[42] The results of the study emphasize the power an analogy has when shaping perception. The study is a continuation

---

36. *See generally* LON L. FULLER, THE MORALITY OF LAW (1964) (discussing the relationship between legal and moral reasoning); H. L.A. Hart, *Positivism and Separation of Law and Morals*, 71 HARV. L. REV. 593 (1958) (discussing positivism as related to the relationship between law and morals).

37. *See* FELIX FRANKFURTER, *When Judge Cardozo Writes*, *in* LAW AND POLITICS 103, 104 (Archibald Mac Leish & E. F. Prichard, Jr. eds., 1939) ("Literature is not the goal of lawyers though they occasionally attain it.").

38. Oliver Wendell Holmes, *The Profession of the Law*, *in* COLLECTED LEGAL PAPERS 29, 29 (1921).

39. *See Analogy and Analogical Reasoning*, STAN. ENCYCLOPEDIA PHIL. (Jan. 25, 2019), https://plato.stanford.edu/entries/reasoning-analogy/.

40. Paul H. Thibodeau & Lera Boroditsky, *Metaphors We Think with: The Role of Metaphor in Reasoning*, PLOS ONE, Feb. 23, 2011, at 1, 2–3.

41. *See id.* at 5.

42. *See id.*

of longstanding linguistics literature on the power of language to shape perception.[43]

An analogy creates a deep impression of the mood and emotion the writer wishes to convey. A strong analogy resonates deeply with the reader and becomes memorable while convincing the audience to conceive the matter in the way the speaker wished.[44] The role of analogies in law is narrower, however, and both more and less ambitious.[45] It is a less ambitious role in the sense that analogies at law do not seek to deeply resonate or even persuade every reader. Analogies in law are crafted by and for specialist interpretation.[46] Analogies in law are more ambitious than analogies in pure reason, in so far as legal analogies are intended to ultimately do real work in governance. The role of analogy in law is to solve problems, not merely to describe problems.[47] Legal analogies must consider both the nature of the two things being compared and the workability of policy. What makes for a good legal comparison is its effectiveness in providing sensible and governable legal rules. Legal analogies are relatively less concerned with promoting pure understanding and more concerned with coming up with workable rules.[48]

---

43. *See generally* GEORGE LAKOFF & MARK JOHNSON, METAPHORS WE LIVE BY (1980) (examining the impact of metaphors).

44. *See Analogy and Analogical Reasoning*, *supra* note 39.

45. *See Precedent and Analogy in Legal Reasoning*, STAN. ENCYCLOPEDIA PHIL. (June 20, 2006), https://plato.stanford.edu/entries/legal-reas-prec/.

46. *See* HANOCH DAGAN, RECONSTRUCTING AMERICAN LEGAL REALISM & RETHINKING PRIVATE LAW THEORY 50–58 (2013) (describing the legal profession as bounded and determined by craft, including institutional values and the characteristic lawyerly situation sense).

47. *See* Katharina Stevens, *Reasoning by Precedent—Between Rules and Analogies*, 24 LEGAL THEORY 216, 238–39 (2018). Mindful use of analogical reasoning is helpful in the courtroom and on the bench, not merely in public discourse and academic policy discussions. *See generally* Krista C. McCormack, *Ethos, Pathos, and Logos: The Benefits of Aristotelian Rhetoric in the Courtroom*, 7 WASH. U. JURIS. REV. 131 (2014) (arguing for the use of Aristotelian rhetoric in trial).

48. Gordley on usage of the concept of "will" or choice by modern law and economists and rights-based theorists:

> What has changed, I believe, is that few of them are really willing to take seriously the intellectual premises that their philosophical predecessors recognize were crucial to the enterprise in which they were engaged. The founders of utilitarianism believed there were units of 'utility' or pleasure or satisfaction to be maximized. Kant and Hegel believed in freedom as a capacity to choose without regard to one's own inclination or purposes. For the most part, the modern jurists

Analogical reasoning has been employed to justify and suggest new developments in diverse sets of legal contexts, from international treaty law[49] to obesity law.[50] It mitigates against the paralysis that comes from focusing on the newness of a phenomenon. In any case, the impact a rule will have on a society cannot be fully proven or falsified without experiencing that rule in society. However, an analogy to legal rules in existence in society (or even comparable rules in a similar society) can provide approximations of the rule's impact.

Several often-cited scholarly articles have justified analogical reason's role in legal reasoning.[51] Professor Cass Sunstein argued that "analogical reasoning has important advantages over general theories, because those who use analogies are especially attuned to the diverse and plural goods that are at stake in legal and ethical decisions."[52] Scott Brewer provided a philosophical explanation for the logic and rational force of analogical reasoning as a sequence of constrained reasoning steps.[53] Professor Emily Sherwin argued that

---

> who are their heirs do not regard those concepts as central and perhaps not even as defensible.

JAMES GORDLEY, FOUNDATIONS OF PRIVATE LAW: PROPERTY, TORT, CONTRACT, UNJUST ENRICHMENT 18 (2006).

49. *See* Anthea Roberts, *Clash of Paradigms: Actors and Analogies Shaping the Investment Treaty System*, 107 AM. J. INT'L L. 45, 49 (2013).

50. *See* Brooke Courtney, *Is Obesity Really the Next Tobacco? Lessons Learned from Tobacco for Obesity Litigation*, 15 ANNALS HEALTH L. 61, 63 (2006).

51. *See* Scott Brewer, *Exemplary Reasoning: Semantics, Pragmatics, and the Rational Force of Legal Argument by Analogy*, 109 HARV. L. REV. 923, 925–26 (1996); Frederick Schauer & Barbara A. Spellman, *Analogy, Expertise, and Experience*, 84 U. CHI. L. REV. 249, 267–68 (2017); Emily Sherwin, *A Defense of Analogical Reasoning in Law*, 66 U. CHI. L. REV. 1179, 1186–87 (1999).

52. Cass R. Sunstein, *On Analogical Reasoning*, 106 HARV. L. REV. 741, 743 (1993).

53. *See* Brewer, *supra* note 51, at 928. Brewer states:

> This Article . . . focus[es] on two features of analogical argument: its *logical form*—the relation between the truth of the argument's premises and that of its conclusion (this relation is part of its 'semantics,' as I shall explain)—and its *pragmatics*—those features of the context in which the argument occurs that shape and constrain the proper interpretive reconstruction of its logical form. Closely related to this explication of the semantics and pragmatics of analogical argument is an inquiry into its *rational force*. 'Rational force' will be defined in a narrow sense: the rational force of an argument is the degree to which the *form* of the argument yields a reliable judgment about the truth of its conclusion based on the assumed truth of its premises. Although this notion of rational force is narrow, its use as a measure of

analogical reasoning keeps law "on a stable path" and tends to improve the quality of judicial decision making.[54] Analogical reasoning remains a preeminent and characteristic form of legal analysis. The practice of law is grounded in the use of analogies by lawyers and judges in both oral and written arguments, which in turn shapes the law as it is implemented. This suggests analogies can be of particular significance in evaluating the law of new technologies.

III. THE DATA-AS-OIL ANALOGY FAILS AS A MATTER OF LOGIC

Data differs from oil in many substantial ways. The differences include scarcity, rivalrousness, fungibility, relationship to individual humans, and traceability to other particular businesses and organizations.

Data is not scarce, while oil is a scarce resource.[55] An increasing amount of data is produced every year.[56] The limit to the amount of

---

> analogical arguments will provide valuable insights into the rational process that is surely the *élan vital* of legal reasoning.

Id.

54. Sherwin, *supra* note 51, at 1197.

> General rules are . . . justified if they prevent more errors than they bring about. If I am right about the benefits of the practice of analogical reasoning, then this practice can function in much the same way as a good rule. By displacing open-ended reasoning, it can correct errors and keep law on a stable path. Of course a practice such as analogical reasoning is quite different from a rule: just how judges are to draw comparisons among cases is not something that can be captured in canonical form. Nevertheless, a practice of analogical reasoning, ingrained by training and tradition, can work indirectly—in the manner of a rule—to improve the quality of judicial decisionmaking. . . . The method in which judges are trained and habituated should be the one that will serve best in the mass of ordinary decisions by ordinary judges. In any event, a practice such as analogical reasoning may have little effect in extraordinary circumstances. Because it is simply a practice, with nothing standing behind it but habit, general acceptance, and the sense that it has worked well over time, it will sometimes be trumped by glaring error or urgent need.

Id.

55. *See generally* James J. MacKenzie, *Oil as a Finite Resource*, 7 NONRENEWABLE RESOURCES 97 (1998) (providing an overview of oil as a resource).

56. *See* Andy Patrizio, *IDC: Expect 175 Zettabytes of Data Worldwide by 2025*, NETWORKWORLD (Dec. 3, 2018, 2:30 AM), https://www.networkworld.com/article/3325397/idc-expect-175-zettabytes-of-data-worldwide-by-2025.html.

data is the storage capacity of computers, which is continually expanding at a rapid rate.[57] By contrast, there is a finite amount of oil.[58] Oil takes millennia to produce and there is currently no way for modern human society to produce any more of it.[59]

Further, data is a non-rivalrous resource, in sharp contrast to oil. Multiple actors can use the same database of information at the same time.[60] This is because data can be copied without changing the characteristics of the data. Two actors cannot process the same portion of crude oil at the same time.[61] Oil also cannot be copied, it can only be processed once, and oil processing is not reversible; plastic generally cannot be converted back to crude oil.

Finally, data is also non-fungible, whereas oil is generally more fungible.[62] For example, 100 terabytes of data on Medicare patients in Tampa is not the same as 100 terabytes of data on rain and sunlight trends on a farm outside Des Moines. While crude oil may differ in some senses,[63] by and large, one bank of 100 tons of oil is the same as another bank of 100 tons of oil.[64]

---

57. *See* Andy Patrizio, *IDC: Expect 175 Zettabytes of Data Worldwide by 2025*, NETWORKWORLD (Dec. 3, 2018, 2:30 AM), https://www.networkworld.com/article/3325397/idc-expect-175-zettabytes-of-data-worldwide-by-2025.html. "Moore's Law, the observation that the number of transistors—tiny on/off switches that churn out electrical signals that get represented as 1s and 0s—on a chip can be doubled in a short period of time," is a common shorthand for this phenomenon of consistently expanding data storage capabilities. *See* Michael Kannellos, *FAQ: Forty Years of Moore's Law*, CNET (Apr. 1, 2005, 4:00 AM), https://archive.is/20130102082556/http://news.com.com/FAQ+Forty+years+of+Moores+Law/2100-1006_3-5647824.html?tag=nefd.lede#selection-971.1-971.32.

58. *See* Mackenzie, *supra* note 55, at 97.

59. *See id.*

60. *See, e.g.*, Larry Myler, *Data Sharing Can Be a Catalyst for B2B Innovation*, FORBES (Sept. 11, 017, 12:03 PM), https://www.forbes.com/sites/larrymyler/2017/09/11/data-sharing-can-be-a-catalyst-for-b2b-innovation/#237591ff27d (advocating for data sharing and providing an example of data sharing).

61. *See generally Petroleum*, *supra* note 6 (discussing oil processing).

62. *See* Sergey Mityakov et al., *International Politics and Import Diversification*, 56 J.L. & ECON. 1091 n.2 (2013) (stating that "many economists maintain that the world oil market is one great pool, because crude oil is fungible in an integrated oil exchange market").

63. *See* Wendy Lyons Sunshine, *The Basis of Crude Oil Classification*, BALANCE (Nov. 6, 2019) https://www.thebalance.com/the-basics-of-crude-oil-classification-1182570 (highlighting the differences among oil classifications).

64. *See* Mityakov et al., *supra* note 62, at n.2. Assuming that the oil is the same grade and from the same location.

Data can relate to or be about particular people, while oil does not. Most valuable data comes from people.[65] That is, it is produced by them and/or is about them.[66] Data has a relationship to people because it describes them or their behaviors. For example, in a large database of medical vitals of diabetic men ages 45–65, each of the men listed in the database has some relationship to the database because they are listed in it. This is not a problem that has any comparison in the area of oil because oil is not directly made out of information relevant to any person. Data is directly traceable to the actions of other specific businesses and actors in society, whereas oil itself is a natural resource. In order to build the massive databases for data, companies must use data that government or other companies have collected or synthesized.[67] In most cases, the data analyzed is the direct output of an upstream company, organization, or individual—far from being merely "found" with no one having previous rights in it.[68] For example, if a people-tracking company adds information from public records about residents of New York to their database, their data is made of, in part, information that was collected by the state of New York for some specific public purpose. Through the use of metadata, direct analysis of the data itself, or tracking transactions for data purchase, it is possible to identify with reasonable certainty from where data originated.[69] By contrast, when oil is tapped from the

---

65. *See* Kim Hart, *The Most Valuable Data Tech Companies Collect About You*, AXIOS (Sept. 27, 2017), https://www.axios.com/the-most-valuable-data-tech-companies-collect-about-you-1513305802-32c36827-8d74-4dca-b5fd-530c02d88cf3.html (the most valuable data is "intimate data").

66. *Id.*

67. *See Big Data: What It Is and Why It Matters*, SAS, https://www.sas.com/en_us/insights/big-data/what-is-big-data.html (last visited Jan. 15, 2020) (discussing the sources of big data); Christo Petrov, *Big Data Statistics 2020*, TECHJURY (Mar. 22, 2019), https://techjury.net/stats-about/big-data-statistics/ (discussing the amount of data that companies are collecting).

68. *See Where Does Big Data Come from?*, IBM BIG DATA & ANALYTICS HUB, https://www.ibmbigdatahub.com/infographic/where-does-big-data-come (last visited Jan. 15, 2020); *see also* FED. TRADE COMM'N, BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION? 3–5 (Jan. 2016), https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf (discussing the compilation process for big data).

69. The rise of blockchain-mediated transactions in the digital age only underscores and renders more accessible the traceable nature of online transactions. Blockchain is a trustless ledger that records transactions; for example, one can trace every Bitcoin back to the moment it was produced. *Protect Your Privacy*, BITCOIN, https://bitcoin.org/en/protect-your-privacy (last visited Jan. 15, 2020) ("Bitcoin works with an unprecedented level of transparency that most people are not used to dealing

ground, there is no sense in which another legally cognizable actor could be said to be responsible for creating the oil in the first instance.[70]

Oil and data, then, have different characteristics. These substantial differences shape the business models employed in both sectors. Since oil requires expensive infrastructure to mine and process, only a specialized sector of the economy directly participates in mining and processing crude oil.[71] This small group of actors provide the processed versions of oil to producers.[72] Governments can regulate oil by circumscribing the behavior of this small group of actors.[73] By contrast, data is non-rivalrous, cheap to access, and can be inexpensive to process.[74] Put another way, the structural economic barriers to entry that exist in oil processing do not exist in data processing. Virtually every company, big and small, is doing some form of data collection and processing, although their methods may not necessarily be sophisticated at this point.[75]

The different characteristics of oil and data shape social attitudes about the two. Oil itself, as a substance, does not relate to any particular person. The people impacted directly by oil mining and processing are those living near oil mining sites or in states impacted

---

with. All Bitcoin transactions are public, traceable, and permanently stored in the Bitcoin network."). People willing to contribute electricity and computer processing power are rewarded with bitcoin units to create the incentive for the system to stay online. *See* SATOSHI NAKAMOTO, BITCOIN: A PEER-TO-PEER ELECTRONIC CASH SYSTEM 4, https://bitcoin.org/bitcoin.pdf ("By convention, the first transaction in a block is a special transaction that starts a new coin owned by the creator of the block. This adds an incentive for nodes to support the network, and provides a way to initially distribute coins into circulation, since there is no central authority to issue them.").

70. *See* Hobart M. King, *Mineral Rights: Basic Information About Mineral, Surface, Oil, and Gas Rights*, GEOLOGY.COM, https://geology.com/articles/mineral-rights.shtml (last visited Jan. 15, 2020).

71. *See* U.S. GEN. ACCT. OFF., GAO-04-96, EFFECTS OF MERGERS AND MARKET CONCENTRATION IN THE U.S. PETROLEUM INDUSTRY 15–30 (2004), https://www.gao.gov/assets/160/157537.pdf.

72. *See id.* at 16–22.

73. *See id.*

74. *See* D. Daniel Sokol & Roisin Comerford, *Antitrust and Regulating Big Data*, 23 GEO. MASON L. REV. 1129, 1136–38 (2016).

75. *See* Louis Columbus, *53% of Companies Are Adopting Big Data Analytics*, FORBES (Dec. 24, 2017, 7:29 PM), https://www.forbes.com/sites/louiscolumbus/2017/12/24/53-of-companies-are-adopting-big-data-analytics/#5e54943c39a1; Dina Gerdeman, *Companies Love Big Data but Lack the Strategy to Use It Effectively*, HARV. BUS. SCH. (Aug. 21, 2017), https://hbswk.hbs.edu/item/companies-love-big-data-but-lack-strategy-to-use-it-effectively.

by pollution.[76] Debates about the benefits and drawbacks of oil, then, focus cleanly on externalities[77] because there is only a limited set of people who can make the argument that they are personally impacted by the mining processing of the oil.[78] By contrast, much of the most valuable data is intimate data about individual people.[79] As a result, all people in information economies are directly affected because their names, addresses, and many other data points appear discretely in databases held by many actors in society. The public has a natural, direct interest in data because information about each individual is being processed.[80]

Data scientist Jer Thorp bemoans the data-as-oil analogy as "vapid."[81] The shallowness of the analogy is clear once we tease out the factual differences between oil and data. Ultimately, the only meaningful similarity between the two is that they are valuable subjects of commerce. But merely being important for the economy does not imply any cultural, social, or moral equivalence.

IV. DATA-AS-OIL FAILS AS A LEGAL ANALOGY

This Part will discuss the legal significance of the logical failure of the analogy between oil and data. First, it will discuss, with examples, why the law tends to make different choices when it comes to regulating items that differ with respect to scarcity, non-rivalrousness, traceability, and items that animate concerns about personal interests and social norms. This will show that the ways in which oil and data differ make it unreasonable to take lessons from the regulation of the one and apply it to the other. Second, this Part makes the broader point that the differences between data and oil are a strong argument that we should not have the same goals in data

---

76. *See, e.g.*, Massachusetts v. Envtl. Prot. Agency, 549 U.S. 497, 521 (2007) (holding that states had standing to sue about EPA greenhouse gas regulation for motor vehicles because of the impacts as a pollutant).

77. *See, e.g.*, E. Allison & B. Mandler, *Air Quality Impacts of Oil and Gas*, AGI (June 1, 2018), https://www.americangeosciences.org/geoscience-currents/air-quality-impacts-oil-and-gas.

78. *See, e.g.*, Juliana v. United States, 947 F.3d 1159, 1175 (9th Cir. 2020) (finding that the plaintiffs lacked standing).

79. *See* Hart, *supra* note 65 (the most valuable data is "intimate data").

80. *See, e.g.*, ANNA H. TURNER, PUBLIC INTEREST IN DATA SURVEILLANCE AND DATA LEAKS BEFORE AND AFTER SNOWDEN—GOOGLE BIG DATA IN CROSS-NATIONAL PERSPECTIVE 23–27 (2017) (ebook) (discussing the public interest in data surveillance).

81. Thorp, *supra* note 23.

regulation as oil regulation. The analogy, in its broadest scope, invites the conclusion that the goals of regulation should be similar, without earning that conclusion by making an affirmative argument.

A. Legal Implications of the Differing Properties Between Data and Oil

The ways in which data differs from oil in other areas of law have led, and should continue to lead, to different forms of regulation. As described in Part III, oil and data are different. This suggests that it might not be wise to model a regulatory program for data after the regulatory structure in place for the management of oil and other natural resources.

First, oil is definitely a good—that is, a physical, movable resource as opposed to a service.[82] By contrast, industry observers disagree whether "big data" is an intangible asset or a service.[83] This is partly because data's value varies based on the data it is combined with and the algorithms used to process it.[84] Additionally, federal intellectual property law does not cover data sets,[85] so it cannot be said with certainty that data is property. This potential for definitional ambiguity must be dealt with. The common law surrounding the Uniform Commercial Code has procedures for separating transactions that are based on goods (subject to the Uniform Commercial Code) and transactions that are based on services (subject to common law

---

82. *See Goods*, BLACK'S LAW DICTIONARY (11th ed. 2019). One definition of service is "[l]abor performed in the interest or under the direction of others." *Service*, BLACK'S LAW DICTIONARY (11th ed. 2019).

83. This question is often debated, either explicitly or implicitly, in the context of financial technology. *See, e.g.*, William Magnuson, *Regulating Fintech*, 71 VAND. L. REV. 1167, 1171–72 (2018) (discussing the need for a reconceptualization of financial technology regulation); Nizan Geslevich Packin & Yafit Lev-Aretz, *Big Data and Social Netbanks: Are You Ready to Replace Your Bank?*, 53 HOUS. L. REV. 1211, 1216 (2016) (stating that big data companies are providing "bank-like services"); Rory Van Loo, *Making Innovation More Competitive: The Case of Fintech*, 65 UCLA L. REV. 232, 238 (2018) (classifying Fintech as a service).

84. *See* Andy Patrizio, *Why Your Big Data Needs Good Algorithms*, DATAMATION (Nov. 25, 2015), https://www.datamation.com/data-center/why-your-big-data-needs-good-algorithms.html.

85. *See, e.g.*, Feist Publ'ns, Inc., v. Rural Tel. Serv. Co., 499 U.S. 340, 361 (1991) (holding that the names and phone numbers in a phone book were facts and thus not protectable intellectual property). Data sets may be protected as trade secrets at state law. *See, e.g.*, *In re* TXCO Res., Inc., 475 B.R. 781, 804–05 (Bankr. W.D. Tex. 2012) (stating that "a compilation of data may constitute a trade secret").

contract law).[86] Therefore, how data is defined matters in regard to what law applies. Additionally, an important point in any policy discussion on data is when it can be said to belong to someone, which involves many definitional concerns that have no analogue in discussions of oil.[87]

Second, oil is a scarce resource—data is not.[88] Resources that are more scarce are governed differently than less scarce resources.[89] In societies where land is a less scarce resource, that is, societies where the amount of useable land far outstrips the population, the right to exclude is less absolute than where land is scarce.[90] For example, consider that in arid American states, there is a right of use for the first user of water, whereas in wet American states there are, or traditionally were, riparian rights.[91] If every owner adjacent to a stream could freely use that stream, that would have little effect in a wet society, but could be devastating in a dry one. Therefore, oil and data are likely to be governed differently because oil is scarce and data is not.

---

86. *See, e.g.*, Princess Cruises, Inc. v. Gen. Elec. Co., 143 F.3d 828, 832–33 (4th Cir. 1998) (discussing the determination of whether the UCC or common law applies).

87. Articles in law reviews and legal journalism discussing data ownership include the following: Barbara J. Evans, *Much Ado About Data Ownership*, 25 HARV. J.L. & TECH. 69, 75 (2011) (discussing data ownership in the context of health privacy and arguing that ownership alone will not solve many major issues in the field); Mauricio Paez & Mike La Marca, *The Internet of Things: Emerging Legal Issues for Businesses*, 43 N. KY. L. REV. 29, 35 (2016) (subsection on data ownership); Jeffrey Ritter & Anna Mayer, *Regulating Data as Property: A New Construct for Moving Forward*, 16 DUKE L. & TECH. REV. 220, 226–27 (2018).

88. Despite the huge amount of data storage now, many extremely high-density data storage methods, such as DNA storage, are possible, but have not yet been implemented. *See* Yaniv Erlich & Dina Zielinski, *DNA Fountain Enables a Robust and Efficient Storage Architecture*, 355 SCIENCE 950, 950 (2017), http://science.sciencemag.org/content/355/6328/950. The storage of large quantities of data presents theoretical physical limitations. Such limitations are so high and remote enough that literature has focused on the human limitations of processing it. *See, e.g.*, Sofia Grafanaki, *Drowning in Big Data: Abundance of Choice, Scarcity of Attention and the Personalization Trap, A Case for Regulation*, 24 RICH. J.L. & TECH. 1, 4–5 (2017); Tongkui Yu & Shu-Heng Chen, *Big Data Scarce Attention and Decision-Making Quality*, COMPUTATIONAL ECON. 1, 1 (2018).

89. JAMES GORDLEY, FOUNDATIONS OF PRIVATE LAW: PROPERTY, TORT, CONTRACT, UNJUST ENRICHMENT 61–65 (2006).

90. *Id.*

91. *See The Water Rights Process*, CAL. WATER BOARDS, http://www.waterboards.ca.gov/waterrights/board_info/water_rights_process.shtml (last updated Sept. 5, 2018) (discussing differences in water rights across the United States).

Oil is a rivalrous good.[92] Meaning, only one actor can process a given portion of oil at one time.[93] Data, on the other hand, is non-rivalrous.[94] Provided access is available, any number of actors could access and use data. This is a significant difference that accounts for increased complexity and centrality of intellectual property interests versus real property interests.

Traceability is a matter of some importance for both data and oil. The question is far more complex when it comes to data than it is for oil. Oil is found in a particular spot. Some countries make the claim that when oil or other valuable resources are found within the country's borders, the whole body politic has a claim to the profits from the resource.[95] This is different from the sovereign prerogative of fair distribution that some nations have asserted.[96]

In some countries, it is a question of policy whether the country that hosts the oil field should directly benefit from the money made from the sale of that oil.[97] Because the oil is located within one nation, most nations hold that there is a national right to profit from the resource.[98] For this reason, ninety percent of the world's oil is controlled by state-run oil companies, including the national oil corporations of Saudi Arabia, Venezuela, and Nigeria.[99] The nature of the claims of others, besides the database holder, to data is more complicated. On the one hand, data is less national because pure territorial claims are harder to make given the location of servers and the distributed network of people about whom the data concerns. However, unlike oil, a lot of valuable data is about individual people. This personal relationship creates at least a plausible claim to ownership of the data or claims—notwithstanding lack of an ownership interest in data[100]—to be able to put limitations on its use.

---

92. *See* Joshua New, *Why Do People Still Think Data Is the New Oil?*, CTR. FOR DATA INNOVATION (Jan. 16, 2018), https://www.datainnovation.org/2018/01/why-do-people-still-think-data-is-the-new-oil/.

93. *See id.*

94. *See id.*

95. Paasha Mahdavi, *Why Do Leaders Nationalize the Oil Industry? The Politics of Resource Expropriation*, 75 ENERGY POL'Y 228, 234–35 (2014).

96. *See* King, *supra* note 70.

97. *See* Mahdavi, *supra* note 95, at 234–36.

98. *See id.*

99. *See* VALÉRIE MARCEL, OIL TITANS: NATIONAL OIL COMPANIES IN THE MIDDLE EAST 1–2, 6 (2006); *see also Oil's Dark Secret*, ECONOMIST (Aug. 12, 2006), https://www.economist.com/special-report/2006/08/10/oils-dark-secret.

100. Due to adverse case law, data is not considered property in the United States. *See infra* notes 114, 128 and accompanying text. Whether it should be considered property is a live debate, with most commentators concluding that a notion of data

In areas of commerce that involve taking land rights from others, those deontological values must be weighed on top of the mere question of who owns the underlying thing.[101] There is a question here that goes beyond the efficient and fair spread of money, as in the oil question. It is more of a distributive justice question—that is, a question about who should have the right to benefit from a resource where two or more parties have a broad moral argument in their favor.[102] The question of who owns oil does not provide us with enough nuance to guide the conversation about data.

Taken together, data has different characteristics from oil such that it will require different regulatory tools, based on existing legal frameworks, to determine a legal approach. A comparison to oil sheds no light on the critical question when it comes to data regulation—what to do about the relationship between the current data processor and upstream data processors, and ultimately, between the data processor and the individuals about whom the data concerns.

B. Fallout from the Cultural Impact of the Analogy

Even if we did have the same attitudes toward and goals for oil as data when it comes to regulation, the differences between the two means that an approach to one would not be transferable to achieve similar results with the other. This section goes a step further and argues that the underlying differences discussed above are reasons why the law should not have the same goals for oil regulation as data regulation. The permeation of the data-as-oil analogy in society, particularly in the form of the term "data mining," is thus a direct threat to reasonable regulation of data use.

---

ownership is not well suited to promoting either human dignity of consumers or innovation. *See* Evans, *supra* note 87, at 76–82. While Europe recognizes a property interest in data there are limitations—known as data protection rules—on what a company that owns data, but is agnostic about who owns the data, can do with personal information. *See* Gianclaudio Malgieri, *"Ownership" of Customer (Big) Data in the European Union: Quasi-Property as Comparative Solution?*, 20 J. INTERNET L. 1, 2, 6 (2016); Ritter & Mayer, *supra* note 87, at 220, 229, 246.

101. Consider takings clause cases where the effect of eminent domain is to transfer private property from one person to a corporation for the purpose of a speculative redevelopment plan. *See* Kelo v. City of New London, 545 U.S. 469, 484–90 (2005).

102. *See Distributive Justice*, STAN. ENCYCLOPEDIA OF PHIL. (Sept. 26, 2017), https://plato.stanford.edu/entries/justice-distributive/.

The most common forms of regulation today is through legislative and administrative rulemaking.[103] An important sphere for analogy, then, is not analogies as judges might apply them, but how stakeholders in industry, consumers, and legislators might encounter the analogy in rulemaking.[104] As Oxford researcher Josh Cowls put it: "How big data is symbolised affects how it is perceived by readers and writers alike, which in turn influences expectations about the impact of big data on our collective quality of life."[105]

In regard to ownership, outside the American tradition there is controversy about who exactly owns oil, that is to say, whether the government and people of the country where the oil was mined have a claim to it.[106] However, in the United States, it is settled law that the mining company that owns the land is entitled to own the oil they mine.[107] So, the analogy creates the feel, to an American ear, that when a company "mines" data, there is no legal or moral problem with them having complete property rights in that data. It invites a policy conclusion without making the argument that it is a good policy.

One might think it would be good policy to allow companies ownership over the data they "mine."[108] However, the oil analogy, taken seriously, should not move toward that conclusion. The analogy provides a guide for the questions that an advocate of liability-free good-title "data mining" must answer to show that their position would make for good law. For example, if data is cheap to copy and effectively unlimited, why must we allocate a property right to create

---

103. *See* MAEVE P. CAREY, CONG. RESEARCH SERV., IF10003, AN OVERVIEW OF FEDERAL REGULATIONS AND THE RULEMAKING PROCESS 1–2 (2019), https://fas.org/sgp/crs/misc/IF10003.pdf.

104. *See supra* Part II.

105. Josh Cowls, *Big Data: The New Water or the New Oil?*, OXFORD INTERNET INST. (Dec. 18, 2014), https://www.oii.ox.ac.uk/blog/big-data-the-new-water-or-the-new-oil/.

106. *See supra* Section IV.A for a discussion of the national policy decision between nationalized and privatized oil.

107. *See* Michael P. Joy & Sashe D. Dimitroff, *Oil and Gas Regulation in the United States: Overview*, THOMSON REUTERS PRAC. L. (June 1, 2016), https://content.next.westlaw.com/Document/I466099551c9011e38578f7ccc38dcbee/View/FullText.html?transitionType=Default&contextData=(sc.Default)&firstPage=true&bhcp=1.

108. This Article brackets the important question of competing ownership claims to data sets as between different firms and people about whom the data concerns. *See generally* Pamela Samuelson, *Privacy as Intellectual Property?*, 52 STAN. L. REV. 1125 (2000) (evaluating and rejecting the idea of privacy as property); Peter K. Yu, *Data Producer's Right and the Protection of Machine-Generated Data*, 93 TUL. L. REV. 859 (2019) (evaluating and rejecting corporate ownership rights in databases).

an incentive to exploit it in the first place? Given the lack of scarcity, why wouldn't free flowing data lead to a bigger pie? What data does a company with a data database own? How and why would the law protect a non-unique interest that cannot be identified and valued with reasonable certainty? Why should these upstream creators not be compensated for uses they did not agree to? What is the effect on the entire database if another actor has a better claim to one or several entries? These questions must be answered in order to determine data's place in the modern economy and social landscape.

The data-as-oil analogy, despite its conclusory veneer, actually poses more questions than answers. Before accepting "data mining" as a description of the economy, stakeholders, legislators, and academics must consider these questions head-on.

V. ALTERNATIVES TO THE DATA-AS-OIL ANALOGY

This Part outlines alternative analogies for understanding data's role in society and the economy.[109] The three forms considered are: (1) data as intellectual property; (2) data as personhood; and (3) data as salvage.[110] I will show that while each analogy offers insights unavailable in the data-as-oil analogy, each also has limitations. This suggests that no analogy can stand alone as an all-purpose analogy for understanding data's role in the information economy. Perhaps

---

109. Some commentators have made other possible analogies for understanding big data. One other analogy is data as infrastructure. *See* W. Nicholson Price II, *Risk and Resilience in Health Data Infrastructure*, 16 COLO. TECH. L.J. 65, 67 (2017). While the idea of data as infrastructure has some substantive appeal, the contested definition of infrastructure itself limits its value as an analogy from a legal perspective. *See, e.g.*, *What Is Infrastructure? Definition and Examples*, MKT. BUS. NEWS, https://marketbusinessnews.com/financial-glossary/infrastructure-definition-means/ (last visited Jan. 15, 2020). Another is data as a periodic table. *See* Michael McGuiness, *Big Data Isn't the New Big Oil—It's the New Periodic Table*, HR POL'Y ASS'N (Nov. 16, 2018), http://www.hrpolicy.org/news/story/big-data-isnt-the-new-big-oil%E2%80%94its-the-new-periodic-table-16180.

110. Comparing data to oil is analogizing data to personal property. This analogy and its implications were discussed in detail in the previous two Parts. *See supra* Parts III, IV. Comparing data to personal property—be it oil or any other type of physical good—is a weak analogy, as discussed in the previous sections, due to differences with respect to rivalrousness and scarcity. *See supra* Parts III, IV. Joshua Fairfield and Christoph Engel argue that privacy should be treated as a public good and policy decisions should be made accordingly. *See* Joshua A.T. Fairfield & Christoph Engel, *Privacy as a Public Good*, 65 DUKE L.J. 385, 385–86 (2015). So, if data is to be usefully compared to any type of good, public goods (e.g., lighthouses and military defense) may be the way forward due to the shared key characteristic of non-rivalrousness with big data.

there is no resource that fully serves as an analogy to data. That does not mean that analogies are useless in policy analysis of data. However, it does mean analogies must be carefully tailored and contextualized to prevent misleading results.

A. Data as Intellectual Property

The analogy between data and intellectual property holds more promise than the analogy between personal property and data. This is because personal property and data are both intangible assets. Like data, intellectual property is non-rivalrous, non-fungible, non-scarce, and very valuable in the information economy.[111] What's more, some intellectual property is about particular people (the right of publicity)[112] and some intellectual property even holds particular personal significance to individuals that cannot be disclaimed (moral rights, which are particularly strong as applied to artistic works).[113] In response to these similarities, some European jurisdictions already consider databases to be a form of intellectual property, based on the theory that databases are business resources that have value.[114]

In the United States, intellectual property is understood to be analogous to, but not the same as, real and personal property.[115] How

---

111. *See* Brian M. Hoffstadt, *Dispossession, Intellectual Property, and the Sin of Theoretical Homogenity*, 80 S. CAL. L. REV. 909, 914–15 (2007).

112. The right of publicity is "a form of intellectual property right that protects against the misappropriation of a person's name, likeness and perhaps other indicia of personal identity for commercial benefit." *What Is Right of Publicity?*, INT'L TRADEMARK ASS'N, https://www.inta.org/Advocacy/Pages/RightofPublicity.aspx (last visited Jan. 15, 2020).

113. Since the United States ratified the Berne Convention, if not before, the space for moral rights in intellectual property might suggest individuals should be able to assert some control over information they created but no longer control. *See, e.g.*, Benjamin S. Hayes, *Integrating Moral Rights into U.S. Law and the Problem of the Works for Hire Doctrine*, 61 OHIO ST. L.J. 1013, 1013–14 (2000).

114. *See generally* F.W. Grosheide, *Database Protection—The European Way*, 8 WASH. U. J.L. & POL'Y 39 (2002) (discussing how Europe protects electronic and non-electronic databases). However, this outcome is precluded by adverse case law in the United States that states that intellectual property must involve some creativity. Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 363 (1991) (information alone, without a minimum of original creativity, cannot be protected by copyright).

115. *See* Julie E. Cohen, *Property as Institutions for Resources: Lessons from and for IP*, 94 TEX. L. REV. 1, 32–56 (2015) (distinguishing characteristics of intellectual property from other categories of property); Mark A. Lemley, *Ex Ante Versus Ex Post Justifications for Intellectual Property*, 71 U. CHI. L. REV. 129, 141–48 (2004). *But see* Frank H. Easterbrook, *Intellectual Property Is Still Property*, 13 HARV. J.L. & PUB. POL'Y 108, 112–18 (1990) (drawing parallels between intellectual property and

exactly to analogize intellectual property to personal property, though, is a question that still plagues courts and commentators.[116] Intellectual property law is still property law—that is, its function is to strictly protect the interests and incentives of the intellectual property owner, despite the reality of corporate collective innovation in the modern economy.[117]

Even though the intellectual property analogy holds more promise than the personal property analogy, the objectives the law has for regulating intellectual property are a weak fit for the realities of the data economy. Michael Mattioli has characterized this as "the data-pooling problem" in the context of data and intellectual property.[118] The modern data processing economy relies upon the frequent transfer, collection, and sharing of data between companies.[119] Mattioli argues that intellectual property law encourages individual investments but does not facilitate, and at times hinders, cooperation among intellectual property holders.[120] Working from the case study of private medical databases, Mattioli proposes a regulatory model that differs substantially from the existing American intellectual property apparatus.[121] Mattioli persuasively contends that the intellectual property framework is at odds with the reality of the needs of the market for data.[122]

An argument against applying intellectual property rules to data is the argument that it promotes monopolies. Monopolies often have negative impacts for consumers.[123] Industry consultant Nathan

---

tangible property); David Fagundes, *Property Rhetoric and the Public Domain*, 94 MINN. L. REV. 652, 670–702 (2010) (advocating the comparison of intellectual and real property).

116. *See, e.g.*, Shyamkrishna Balganesh, *Copyright as Market Prospect*, 166 U. PA. L. REV. 443, 444–45 (2018).

117. *See, e.g.*, Christopher A. Cotropia, *The Individual Inventor Motif in the Age of the Patent Troll*, 12 YALE J.L. & TECH. 52, 54–55 (2009); Shubha Ghosh, *Enlightening Identity and Copyright*, 49 BUFF. L. REV. 1315, 1317 (2001) (book review) ("Copyright law is premised on the assumption of the 'romantic author'—the lone genius that creates valuable expression. Many scholars have criticized this assumption without providing a substitute.").

118. Michael Mattioli, *The Data-Pooling Problem*, 32 BERKELEY TECH. L.J. 179, 179 (2017).

119. *See id.* at 205–36 (discussing the need for data sharing in the health research context).

120. *Id.*

121. *Id.* at 222–36.

122. *See id.* at 235–36.

123. Emily Stewart, *America's Monopoly Problem, in One Chart*, VOX (Nov. 26, 2018, 12:00 PM), https://www.vox.com/2018/11/26/18112651/monopoly-open-markets-institute-report-concentration.

Newman argues that granting companies intellectual property-like rights in their databases will grant them monopolies giving insurmountable market power to current major market players.[124] Newman argues that major market players effectively behaving as though they have full property rights in their data has led to "the failure of markets in online sectors."[125] These concerns have been echoed by other practitioners of antitrust as a looming issue for the data economy to resolve.[126] Of course, intellectual property deliberately gives owners monopolies over the intellectual property they create.[127] There is an intuition, immortalized in American law in the seminal Supreme Court case, *Feist Publications, Inc. v. Rural Telephone Service Co.*, that this monopoly should only be granted when the inventor has produced something creative.[128] The right to the monopoly in the American tradition is based on giving the inventor their due for their creativity. It may, then, be less just to give such a powerful monopoly to companies that just happened to be early information age collectors of valuable databases.

B. Data as Personhood

Data produced by and about humans can be considered an aspect of personhood that is hurt by certain types of processing, with or without permission. Personal data is, in this analogy, as much a part of a person as their arms and legs. People are entitled to their bodily integrity and have similar rights in things that the law finds to be an extension of their person. The law occasionally recognizes this type of

---

124. Nathan Newman, *Search, Antitrust, and the Economics of the Control of User Data*, 31 YALE J. ON REG. 401, 435, 450–54 (2014) [hereinafter *Control of User Data*]. Nathan Newman's credentials can be found on LinkedIn. Nathan Newman, LINKEDIN, https://www.linkedin.com/in/nathan-newman-3206724/ (last visited Jan. 15, 2020).

125. *Control of User Data*, *supra* note 124, at 450.

126. *See, e.g.*, Darren S. Tucker & Hill B. Wellford, *Big Mistakes Regarding Big Data*, ANTITRUST SOURCE, Dec. 2014, at 1, 1.

127. *See* Aaron Xavier Fellmeth, *Copyright Misuse and the Limits of the Intellectual Property Monopoly*, 6 J. INTELL. PROP. L. 1, 2 (1998) (stating that "[t]he 'exclusive right' conferred on intellectual property thus results in a limited, federal government-granted monopoly on the subject of the patent, trademark, or copyright").

128. Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 346 (1991) ("'[W]hile the word *writings* may be liberally construed, as it has been, to include original designs for engraving, prints, [sic], it is only such as are *original*, and are founded in the creative powers of the mind. The writings which are to be protected are *the fruits of intellectual labor*, embodied in the form of books, prints, engravings, and the like.'" (quoting Trade–Mark Cases, 100 U.S. 82, 94 (1879))).

interest in personal property.[129] For example, in tort law, one can commit battery by touching or harming an object in direct proximity to the plaintiff.[130] Similarly, moral rights in intellectual property are based on the notion that, particularly with respect to works of art, distortions and misrepresentations of intellectual property made by the author are harms to the person because the art is an extension of them as a person.[131]

Several scholars also made the argument that privacy is linked to personhood and human dignity, well before the dawn of the information age. In 1964, Edward Bloustein argued that the "gist of the wrong in the intrusion cases" is "a blow to human dignity, an assault on human personality."[132] Richard Turkington, in a retrospective on the privacy right, grounded a theory of privacy in human dignity and respect for persons.[133]

Considering privacy as an essential aspect of personhood would have implications for information policy. Since each person would have a non-disclaimable personal interest in data that is about them, regulation would have to take the form of limitations on what an actor can do to data in general. This would, at least in substantial part, free data regulation from notions of consent. After all, the notion of bodily integrity can allow tort actions to go forward even when there has been some consent.[134] In a groundbreaking 2012 book, Anita Allen argued for mandatory privacy protections that could be foisted on people who do not want them, writing "[t]he protection of privacy rights . . . should not be thought of as something that can be waived

---

129. *E.g.*, *In re* B.L., 239 Cal. App. 4th 1491, 1496 (2015) ("'[C]ase law from other jurisdictions supports the proposition that there need not be direct touching of the victim's person in order for a battery to occur. Touching something intimately connected with the victim's body is sufficient.'" (citations omitted)).

130. *Id.*

131. *See* Irma Sirvinskaite, *Toward Copyright "Europeanification": European Union Moral Rights*, 3 J. INT'L MEDIA & ENT. L. 263, 264 (2010) ("Moral rights are rights that are personal to the author which go beyond the author's desire for pecuniary gain.").

132. Edward J. Bloustein, *Privacy as an Aspect of Human Dignity: An Answer to Dean Prosser*, 39 N.Y.U. L. REV. 962, 974 (1964).

133. *See* Richard C. Turkington, *Legacy of the Warren and Brandeis Article: The Emerging Unencumbered Constitutional Right to Informational Privacy*, 10 N. ILL. U. L. REV. 479, 484 (1990).

134. *E.g.*, Allen v. Dover Co-Recreational Softball League, 807 A.2d 1274, 1288 (N.H. 2002) (owners and insurers of amateur softball league not liable in negligence for injuries sustained by a player during play).

by intended beneficiaries at will."[135] Allen defines the scope of the mandatory privacy right, as she sees it, as follows:

> The law justly identifies some of us as beneficiaries and targets of legal duties to hide, duties to embrace even unpopular privacy. What must we hide then? We must hide what is necessary to preserve our common dignity and separate virtues. We must hide what is necessary to keep ourselves safe from harm. We must hide what our roles and responsibilities and professions dictate that we hide as matters of efficacy, beneficence, or contract. And we must hide, notwithstanding all of technology's attractions, what good relationships and reputations—now and in our distant and uncertain futures—render it prudent to hide.[136]

As a result of what American law deems fundamental to personhood, a person can neither voluntarily choose to sell themselves into slavery[137] nor sell their services as a prostitute.[138] The question, then, is whether privacy is the type of interest that the government must make mandatory.

Many European jurisdictions consider privacy to be a fundamental aspect of personhood.[139] By contrast, many American privacy scholars have critiqued the notion of personhood, specifically related to data privacy. For example, privacy scholar Julie Cohen has critiqued the intuitiveness of this formulation, writing "[p]ersonhood theory . . . seems an odd way of talking about my control over data that others already possess."[140] Therefore, whether privacy is considered to be a component of personhood may differ among countries, impacting how data privacy is treated in each setting.

The data as personhood analogy does face one challenge. The fundamental difference between considering data as personhood and other situations where the law has implied that an object or interest is an aspect of personhood is the relative lack of intimacy of each

---

135. ANITA L. ALLEN, UNPOPULAR PRIVACY: WHAT MUST WE HIDE?, at xii (2011).

136. *Id.* at 197.

137. U.S. CONST. amend. XIII.

138. Prostitution is illegal as a matter of state law in every state but Nevada. *Prostitution*, JUSTIA, https://www.justia.com/criminal/offenses/sex-crimes/prostitution/ (last visited Jan. 15, 2020).

139. *See* James Q. Whitman, *The Two Western Cultures of Privacy: Dignity Versus Liberty*, 113 YALE L.J. 1151, 1155–56 (2004).

140. Julie E. Cohen, *Examined Lives: Informational Privacy and the Subject as Object*, 52 STAN. L. REV. 1373, 1382 (2000).

instance of data transfer. Each instance of battery by, say, cutting off a person's shirt is viscerally all on its own violative of personhood. By contrast, violations to personhood by data transfer, processing, and collection are not always apparent in each transaction, but rather the threat to personhood is created by a whole environment in which personal information is used to evaluate and curate each person's world.[141] This difference potentially weakens the analogy between personhood and data. At any rate, it is a factor that any framework approaching data from a personhood perspective should contemplate.

C. Data as Salvage

Finally, one could think of data as analogous to salvage.[142] The purpose of a salvage award policy is to promote humanitarian rescue of life and property and maritime commerce by preserving property from destruction.[143] This is a policy that creates an incentive for professionals to provide rescue services, which is beneficial for commerce because it allows property that would otherwise be in disuse to reenter the market.[144]

Justice Story described the measurement of salvage as follows:

> In cases of salvage, the measure of reward has never been adjusted by a mere estimate of the labor and services performed by the salvors. These, to be sure, are very important ingredients; and are greatly enhanced in value, when they have been accompanied by personal peril and gallantry, by prompt and hardy enterprise, and by severe and

---

141. *See* FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION 3, 5, 19, 21 (2015).

142. Black's Law dictionary defines salvage in general as "the rescue of imperiled property." *Salvage*, BLACK'S LAW DICTIONARY (11th ed. 2019). A notable treatise provides further detail:

> A salvage award, or reward, is the compensation allowed to the volunteer whose services on navigable waters have aided distressed property in whole or in part. The award is not regarded merely as pay on the principle of quantum meruit or as remuneration pro opera et labore, but as a reward to persons participating and the owners of salving property, voluntarily rendering their services and to encourage others to similarly undertake the saving of life and property.

MARTIN J. NORRIS, THE LAW OF SALVAGE § 3 (1958).

143. Ocean Servs. Towing & Salvage, Inc. v. Brown, 810 F. Supp. 1258, 1262 (S.D. Fla. 1993).

144. *See id.*

> long-continued exposure to the inclemencies of the winds and waves. But an enlarged policy, looking to the safety and interest of the commercial world, decrees a liberal recompense, with a view to stimplate [sic] ambition, by holding out what may be deemed an honorable reward.[145]

To many readers, salvage may sound as exotic and remote as talking of pirates, but salvage is actually quite a lively area of modern practice.[146] The use of this analogy would address the intuition that the data processor should be economically remunerated for the work they did to produce the outputs, but also gives voice to the reality that the resource has a traceable claim from another person. Importantly, those who salvage resources are not necessarily the ultimate owners of those resources.[147] Nonetheless they have a claim to substantial value.[148]

The analogy of salvage does justice to the complexity of the landscape and variety of interests involved in the data industry. Much like cyberspace, the sea is ungainly, unpredictable, and complicated. The complexity and value of the service done by both salvagers and companies that harvest and process data are valuable. Nonetheless, in both cases, the question of who ultimately owns the resource and how much value the salvager receives depends on the particular case.

---

145. Rowe v. Brig, 20 F. Cas. 1281, 1283 (C.C.D. Mass. 1818).

146. *See* Lewis v. Atchison (The Alamo), 75 F. 602, 607–08 (5th Cir. 1896) ("An iron steamship, heavily laden, ashore anywhere on the Atlantic coast, is in very great peril, and she is probably in greater peril on the coral formation of Florida than elsewhere. . . . On the Florida reefs the danger is that the stranded vessel will rapidly break to pieces, and the master who fails to recognize the imminent danger of the vessel when aground from 24 to 48 hours is guilty of temerity not belonging to a prudent and reasonable navigator."); *see, e.g.*, Odyssey Marine Expl., Inc. v. Shipwrecked & Abandoned SS Mantola, 333 F. Supp. 3d 292, 302–03 (S.D.N.Y. 2018); *Ocean Servs. Towing & Salvage, Inc.*, 810 F. Supp. at 1262; *see also* Cathryn Henn, *The Trouble with Treasure: Historic Shipwrecks Discovered in International Waters*, 19 U. MIAMI INT'L & COMP. L. REV. 141, 142–45 (2012) (describing how technological advances have increased the number of salvages which in turn creates a larger amount of complex legal claims). *See generally* FLA. BAR, FLORIDA MARITIME LAW & PRACTICE (6th ed. 2019) (serving as a current practice guide for Florida attorneys who encounter salvage law issues).

147. Cory T. Stuart, *The Wake of Discovery—A Primer on Legal, Historical, and Practical Shipwreck Salvage Dynamics*, 9 LOY. MAR. L.J. 45, 47 (2011) (distinguishing salvage from "the law of finds"—the latter implies ownership while the former does not).

148. *See id.* (specifically in the form of "remuneration for efforts involved in assisting others in hazardous conditions" and "an award to be given by the owner.").

This is because the value of the services of the salvager must be considered alongside the interest of the person or company that caused the resource to be there in the first place.[149] Ultimately, salvage is a better analogy to data than oil because salvaged items, like data, come from people and companies. In data, as in salvage, there may be personal or corporate interests associated with data that must be untangled. But despite this potential complexity, there remains an active salvage industry, which may be of some comfort to those who worry about preserving the beneficial aspect of the data economy. However, this analogy also reveals why we may not be willing to promote and facilitate the industry's growth at all costs. The companies, governments, and people who created the data also have an interest in the data in question because it is made by or about them. Any analogy we use to understand data should acknowledge these interests, regardless of what choices we make from there.

From a socio-political perspective, the salvage analogy serves a useful function. It is a reminder that although the data "miner" might have done work to acquire its data, there may well be other actors who need to be compensated for the work, even if the "miner" cannot be said to have acted wrongfully and is even serving a useful function in society.

CONCLUSION

It is the particular competence of lawyers and judges to evaluate metaphors as applied to guide the dispensation of justice. Silicon Valley has much to teach us about human and machine capabilities. However, the legal profession should not uncritically defer to industry thought leaders on the question of the appropriateness of an analogy as applied to the legal context.[150]

What makes the data-as-oil analogy so misguided is that it presumes no link between data and individuals whose personal information are in the databases or data and the entity that rendered the data in machine-readable form. We need analogies that give voice to the set of issues data presents, not analogies that paper over them.

Data can also be analogized to intellectual property, personhood, or naval salvage. These analogies take into account the contributions of upstream data producers, including the people about whom data

---

149. *See* Craig J. S. Forrest, *Has the Application of Salvage Law to Underwater Cultural Heritage Become a Thing of the Past?*, 34 J. MAR. L. & COM. 309, 319 (2003).

150. *See, e.g.*, Whitney & Simpson, *supra* note 8, at 37–42 (analyzing the limits of analogical reasoning in information-age discussions of search engines and free speech).

may be about, while acknowledging the value of what downstream data processors contribute. But ultimately, no single analogy is a perfect match for the unique opportunity and challenge presented by data. The thoughtful lawyer and policymaker must understand the limits of analogy when understanding new technology, and the need to tailor analogy to the particular context or problem to be solved. The future is at stake.