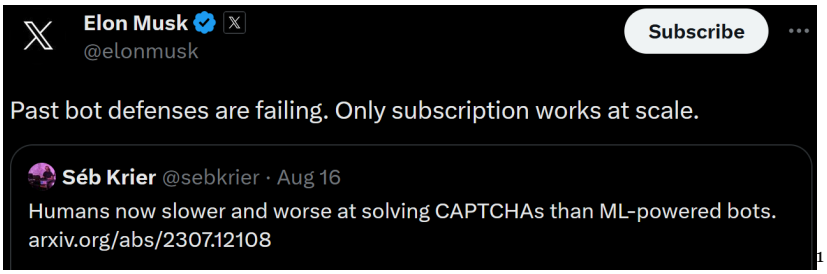
TAKING SCALE SERIOUSLY IN ROBOTICS AND A.I. LAW

Mark P. McKenna* and Woodrow Hartzog**

Dear We Robot Participants: Thank you so much for reading this very early and unfinished draft. Many more citations and additional examples are to come, so please ask permission before citing. We are thrilled to receive your feedback and are particularly interested in whether our distinction between "scale is more" and "scale is different" is coherent and useful. Comments warmly welcome.



[1]

Issues of scale—the relationship between the amount of an activity and its associated costs and benefits—permeate discussions around the law of robotics and artificial intelligence. Technologies and business models are seen as effective only if they are "scalable" and certain strategies, such as CAPTCHAs to weed out bots, are critiqued if they "don't scale." But it's not always clear how lawmakers and judges conceptualize "scale" when approaching questions around automated technologies. Intuitively, scale might just mean "more." But issues of scale can introduce new harms and benefits along different dimensions, not simply costs or efficiencies of greater magnitude. In this Article, we argue for a more sustained interrogation of the role of scale in law, one that is more sensitive to the distinction between what we describe as "scale is more" and "scale is different." When lawmakers and judges fail to properly categorize the role of scale in a particular context, they risk ignoring or misidentifying harms, misdiagnosing the causes of those

---

* Professor of Law, UCLA School of Law; Faculty Co-Director, UCLA Institute for Technology, Law & Policy; Of Counsel, Lex Lumina PLLC..

** Professor of Law, Boston University School of Law. The authors would like to thank Kabbas Azhar, Susan Hong, Benjamin Kline, and Janelle Robins for their research assistance.

[1] @elonmusk, TWITTER (Aug. 16, 2023, 8:24 PM), https://twitter.com/elonmusk/status/1691969296543711471?s=20.

harms, and potentially focusing on the wrong policy tools, and even the wrong actors, in proposing solutions.

TAKING SCALE SERIOUSLY IN AI AND ROBOTICS LAW

INTRODUCTION

Considerations of scale—the relationship between the amount of an activity and its associated costs or benefits—are everywhere in technology law and policy.[2] It's barely an exaggeration to say that scalability is *the*

[2] See, e.g., Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 CAL. L. REV. 513, 538 (2015) ("Today's robots do a variety of tasks that people could do, but don't for reasons of cost or preference. Moving more tasks into the category of automation could in and of itself cause legal issues at scale."); Julie Cohen, *How (Not) to Write a Privacy Law*, KNIGHT FIRST AMEND. INST. (Mar. 23, 2021), https://s3.amazonaws.com/kfai-documents/documents/306f33954a/3.23.2021-Cohen.pdf ("Current approaches to crafting privacy legislation are heavily influenced by the antiquated private law ideal of bottom-up governance via assertion of individual rights, and that approach, in turn, systematically undermines prospects for effective governance of networked processes that operate at scale.…[Arguments for user-governed data cooperatives] tend to ignore important qualifications that affect the ability of common-governance arrangements to scale.… Both arguments for bottom-up governance flowing from assertion of individual rights and arguments for commons-based cooperative governance of personal data

reason lawmakers are so concerned about new technologies. If tools didn't scale and only affected a few people, then lawmakers and judges would probably not deem them worthy of categorical regulatory attention. But despite the ubiquity of scale issues, lawmakers and judges usually deal with the concept intuitively—and primarily just to mean "more" of something.

Scale as "more" is often significant. More of an activity frequently produces more harm, and that can tip the balance in tests that pit the risk of harm against the costs of preventing that harm.[3] Similarly, more use of a legal tool might at some point change the incentives for affected parties and shift their behavior in ways that have offsetting costs. In the context of the Digital Millennium Copyright Act, for example, the scale of takedown notices creates strong incentives for platforms to create automated systems like Content ID, which might more efficiently deal with notices of

collection and processing overlook the structural and temporal effects of design operating at scale.... the dysfunctions of the networked information economy reflect underlying problems of networked flow and scale that are distinct from existing patterns of market domination.... To be effective at all, regimes for privacy governance need to target order of magnitude problems in ways that enable oversight and enforcement to scale up and out commensurately."), Sarah Ciston, *A Critical Field Guide For Working With Machine Learning Datasets*, KNOWING MACHINES PROJECT (2022), https://knowingmachines.org/critical-field-guide ("The speed and scale of machine learning and massive datasets make "discrimination easier, faster, and even harder to challenge.... Whether designing a dataset from scratch or using one that has been around for years, decisions made at every step will inform your project outcomes. These decisions get scaled and compounded by machine learning models.") (citing RUHA R. BENJAMIN, RACE AFTER TECHNOLOGY: ABOLITIONIST TOOLS FOR THE NEW JIM CODE (2019)); Mike Ananny & Kate Crawford, *Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability*, 20 NEW MEDIA & SOC'Y 973 ("Sometimes, the details of a system will be not only protected by corporate secrecy or indecipherable to those without technical skills, but inscrutable even to its creators because of the scale and speed of its design."). Tech companies also regularly refer to issues of scale. *See Facebook, Social Media Privacy, and the Use and Abuse of Data: Hearing Before the S. Comm. On the Judiciary & S. Comm. on Com., Sci., & Transp.*, 115th Cong. (10 Apr. 2018), www.commerce.senate.gov/2018/4/facebook-social-media-privacy-and-the-use-and-abuse-of-data (quoting Mark Zuckerburg saying "We have gotten increasingly better at finding and disabling fake accounts....This is thanks to improvements in machine learning and artificial intelligence, which can proactively identify suspicious behavior at a scale that was not possible before—without needing to look at the content itself.)

[3] The most obvious example here would be the Hand Formula, which requires courts to compare the product of the probability of harm from some conduct (P) and the magnitude of the loss produced by that conduct (L) with the burden of preventing the harm (B). When PxL>B, it is negligent not to take the precaution that would prevent the harm. U.S. v. Carroll Towing, 159 F.2d 169 (2d Cir. 1947).

infringement but not be as sensitive to legitimate uses of copyrighted content.[4]

But sometimes scale does not just mean more—it creates a qualitatively different situation. Scale in this sense does not relate simply to harms or benefits that are greater in magnitude but instead involves new harms or benefits along different dimensions. The dominance of social media platforms and "network effects" is a good example. At a certain scale, the size and popularity of a social media platform is meaningfully different, both in its perceived value to users and in the business models it enables. This is the tipping point where people are attracted to a platform not primarily because of the platform's features but because of who else is there. Importantly, scale in this context creates a new collective action problem–social media platforms are more valuable precisely because of other users' presence on the platform, making it harder to switch from one platform to another. We are seeing this in real time with users' attempts to find replacements for Twitter.

In this Article, we argue for a more sustained interrogation of the role of scale in technology law, one that is more sensitive to the distinction between what we describe as "scale is more" and "scale is different." That distinction is particularly crucial in the context of robotics and artificial intelligence. When lawmakers and judges fail to properly consider the role of scale in a particular context, they risk ignoring or misidentifying harms, misdiagnosing the causes of those harms, and focusing on the wrong policy tools, and even the wrong actors, in proposing solutions.

In our terminology, "scale as more" refers to situations in which costs and/or benefits of an activity have a more or less linear relationship with the amount of that activity.[5] If an act causes x units of harm, then the

---

[4] Katharine Trendacosta & Corynne McSherry, *What Really Does and Doesn't Work for Fair Use in the DMCA*, EFF (July 31, 2020), https://www.eff.org/deeplinks/2020/07/what-really-does-and-doesnt-work-fair-use-dmca.

[5] *See* Adrian Bridgwater, *What Is Technology Scalability?,* FORBES (19 Feb. 2020), www.forbes.com/sites/adrianbridgwater/2020/02/19/what-is-technology-scalability/?sh=9181da04f3f0 ("[W[e can probably assume that scalability in the IT platform and application sense refers to scaling upwards, to make a piece of technology bigger and more expansive.").

total harm of that activity is x multiplied by the number of instances of the activity. One privacy violation is bad. A thousand privacy violations are worse because it's an additional 999 instances of harm. Again, that kind of scale is often important, because the aggregate harm of an activity must be compared to its aggregate benefits and/or the costs of mitigating that harm.

When "scale is different," the equation might change in at least four possible ways that should cause lawmakers and judges to think of the problem differently.

First, *the population affected could change*. For example, the data collected for machine learning doesn't just affect each individual whose data is collected in the sense that each suffers an individuated harm that we can simply multiply by the number of people whose data is used. At scale, that data provides population-level insights that can be used against different people within the same category (not just the person whose data is collected) and different categories of people.[6]

Second, the scale of an activity can *create new problems* that didn't exist in small numbers. For example, if racial bias becomes encoded in all automated systems, individual instances of wrongful discrimination at scale can have the effect of shutting people out of entire career options and other important life decisions.[7] Likewise, the number of sidewalk robots in use might fundamentally change the physical landscape: those robots might be annoying in small numbers but at scale can clog up sidewalks so much the sidewalks become unusable.

Third, the scale of activity *can challenge original assumptions about the costs and benefits of an activity*. Manipulation via dark patterns might always be wrongful but the harms might seem *de minimis* when viewed from the perspective of individual users. Scale can make the nature of the problem more apparent. Using automated scraping tools to collect "publicly available" data to train machine learning systems might seem functionally equivalent to a person simply reading and writing down

---

[6] *See* Salomé Viljoen, *A Relational Theory of Data Governance*, 131 YALE L.J. 573, 577 (2021).

[7] *See* Katherine Creel & Deborah Hellman, *The Algorithmic Leviathan: Arbitrariness, Fairness, and Opportunity in Algorithmic Decision-Making Systems*, 52 CAN. J. PHIL. 26 (2022).

information that anyone could access if they were given the right link. But most of that information wouldn't be aggregated without the scraping tools because of the time and expense that would be required. The tools enable collection of information that otherwise would have been functionally obscure.[8]

Finally, scale can *affect the efficacy of solutions*, making certain institutional designs more effective and taking some legal, social, design, and market-based remedies and strategies entirely off the table.[9] When it comes to legal remedies, sometimes scale is "different" in that lawmakers cannot assume that expanding a certain type of enforcement or ratcheting up remedies would produce a proportionte increase in efficacy. For example, the automated nature of misinformation makes private lawsuits to remedy individual instances of deception seem futile. Likewise, when scale is different, private enforcement mechanisms that rely predominantly on compensating individual harms are unlikely to address the systemic or structural harms that may only emerge beyond certain thresholds.

We conclude this article with a reflection on the nature of scale in existing legal frameworks for automated technologies and a call for more regulator nuance. We emphasize that more nuanced engagement with scale is not necessarily an argument for more regulation. Sometimes thoughtful consideration of the ways scale matters will have more to do with *how* we regulate than whether we do. Other times an appreciation of the ways scale

---

[8] *See* Woodrow Hartzog & Evan Selinger, *Surveillance as Loss of Obscurity*, 72 WASH. & LEE L. REV. 1343, 1345–46 (2015) ("[W]e argue that the concept of "obscurity," which deals with the transaction costs involved in finding or understanding information, is the key to understanding and uniting modern debates about government surveillance."); Woodrow Hartzog & Evan Selinger, *Increasing the Transaction Costs of Harassment*, 95 B.U. L. REV. ANNEX 47 (2015); Evan Selinger & Woodrow Hartzog, *Obscurity and Privacy*, *in* SPACES FOR THE FUTURE: ROUTLEDGE COMPANION TO PHILOSOPHY OF TECHNOLOGY (Joseph Pitt & Ashley Shew eds., 2018), https://www.routledge.com/Spaces-for-the-Future-A-Companion-to-Philosophy-of-Technology/Pitt-Shew/p/book/9780415842969; *see also* Woodrow Hartzog & Frederic Stutzman, *The Case for Online Obscurity*, 101 CALIF. L. REV. 1, 5 (2013) ("We argue the case for obscurity for two reasons. First, we argue that obscurity is a common and natural condition of interaction, and therefore human expectation of obscurity will transfer to the domains in which we spend time, both physical and virtual. Second, we argue that obscurity is a desirable state because we are protected by an observer's inability to comprehend our actions, and therefore social practice encourages us to seek obscurity."); Woodrow Hartzog & Frederic Stutzman, *Obscurity by Design*, 88 WASH. L. REV. 385 (2013).

[9] *See generally*, Ryan Calo, *Code, Notice, or Nudge?*, 99 IOWA L. REV. 773 (2014).

is different might suggest less need for regulation. In some cases, legal intervention is needed to prevent or remedy harms that are the result of outlier behavior. When someone acts outside the norm, that party might cause unique harms that affected parties are not well situated to avoid. But when that same design or activity becomes the norm, there might be incentive for technological or legal adaptation that makes it less likely the individual harms will be visited in the same way. Those are cases where the scale of an activity changes our assessment of the harm caused by that activity because the scale produces (or is likely to produce) responsive measures that wouldn't exist at lower levels.

I. SCALE IS AN UNDERSPECIFIED CONCEPT

"Scale" is widely invoked in conversations about technology and its governance. Yet, remarkably, that concept is rarely explicitly defined or clarified when used in consequential settings. The Merriam-Webster dictionary defines scale as "something graduated especially when used as a measure or rule," "a graduated series or scheme of rank or order," and "a proportion between two sets of dimensions….a distinctive relative size, extent, or degree."[10]

Within the natural and social sciences, "scale" typically refers to "the spatial or temporal dimension of a phenomenon, and scaling is the transfer of information between scales."[11] Scientists often identify space, time, and organizational level as dimensions or kinds of scale.[12] In statistics, "scaling usually refers to a set of techniques for data reduction and detection of underlying relationships between variables."[13] Ecologists

---

[10] *Scale,* Merriam-Webster Dictionary, https://www.merriam-webster.com/dictionary/scale (last accessed Aug. 7, 2023).

[11] Jianguo Wu & Harbin Li, *Concepts of Scale and Scaling, in Scaling and Uncertainty Analysis*, *in* ECOLOGY: METHODS AND APPLICATIONS 3 (2006).

[12] *Id.* at 5 ("Space and time are the two fundamental axes of scale, whereas organizational hierarchies are usually constructed by the observer.").

[13] *Id.* at 9-10 ("In physical sciences, scaling usually refers to the study of how the structure and behavior of a system vary with its size, and this often amounts to the derivation of a power-law relationship. This notion of scaling has often been related to the concepts of similarity, fractals, or scale-invariance, all of which are associated with power laws. For example, a phenomenon or process is said to exhibit "scaling" if it does not have any characteristic length scale; that is, its behavior is independent of scale – i.e., a power law relationship."). To poorly paraphrase (and with apologies to statisticians), something is scalable where a relative change in one dimension results in a relative proportional change in the other dimension, independent of the initial aspects of those dimensions.

use scaling to predict and understand.[14] In technological circles, scalability is often conceptualized as "the capability of a system, network, or process to handle a growing amount of work, or its potential to be enlarged in order to accommodate that growth."[15] Economists often discuss "economies of scale," whereby an average cost falls as output increases.[16]

Parties in litigation commonly use the language of scale, suggesting that something is "scaling up" or, conversely, "doesn't scale." In all of these cases, courts invoke scale in very general terms, referring in some way to the magnitude of some activity. For example, the court in a case alleging fraud over a tech company's products cited affidavits asserting that the defendant "is not currently competitive on large-scale parallel systems, as Sybase's database does not scale well past four CPUs."[17] In patent litigation,

---

[14] *Id.*

[15] *Scalability*, NETWORK SECURITY, https://www.networxsecurity.org/members-area/glossary/s/scalability.html (last accessed Sept. 16, 2023) ("For example, [scalability] can refer to the capability of a system to increase its total output under an increased load when resources (typically hardware) are added. An analogous meaning is implied when the word is used in an economic context, where scalability of a company implies that the underlying business model offers the potential for economic growth within the company. Scalability, as a property of systems, is generally difficult to define and in any particular case it is necessary to define the specific requirements for scalability on those dimensions that are deemed important. It is a highly significant issue in electronics systems, databases, routers, and networking. A system whose performance improves after adding hardware, proportionally to the capacity added, is said to be a scalable system. An algorithm, design, networking protocol, program, or other system is said to scale if it is suitably efficient and practical when applied to large situations (e.g. a large input data set, a large number of outputs or users, or a large number of participating nodes in the case of a distributed system). If the design or system fails when a quantity increases, it does not scale….Scalability refers to the ability of a site to increase in size as demand warrants. The concept of scalability is desirable in technology as well as business settings. The base concept is consistent – the ability for a business or technology to accept increased volume without impacting the contribution margin (= revenue ? variable costs). For example, a given piece of equipment may have a capacity for 1–1000 users, while beyond 1000 users additional equipment is needed or performance will decline (variable costs will increase and reduce contribution margin.").

[16] *Economies of Scale*, SCIENCEDIRECT, https://www.sciencedirect.com/topics/social-sciences/economies-of-scale (last accessed Sept. 16, 2023).

[17] In re Sybase, Inc. Sec. Litig., 48 F. Supp. 2d 958, 962 (N.D. Cal. 1999); *see also*, In re Cloudera, Inc. Sec. Litig., No. 19-CV-03221-MMC, 2022 WL 14813896, at *14 (N.D. Cal. Oct. 25, 2022) ("Cloudera's offerings provided "[c]loud and on-premises deployment at scale and across hybrid cloud environments[.]"); Indiana Pub. Ret. Sys. v. Pluralsight, Inc., No. 119CV00128JNPDBP, 2021 WL 1222290, at *9 (D. Utah Mar. 31, 2021), *aff'd in part,*

a judge wrote that the "[d]efendant was arguing that while the processes were successful for 'one-off tablets', a POSA would have sought a process that could be *scaled up*. Plaintiffs d[id] not make a plausible argument that a POSA would not want to develop a *scalable process*. Plaintiffs also d[id] not make a plausible argument that a POSA would have [had] options other than modifying Bartholomaus and McGinity if they wanted to produce hardened tablets *at scale*."[18]

In securities litigation, a judge wrote that "Talis also sought to capitalize on a rapidly closing window to sell a new COVID-19 test before demand cooled due to the FDA's approval of the Pfizer and Moderna COVID-19 vaccines in December 2020, and before competing tests captured the market. Talis would need to persuade investors that its product provided fast, accurate, reliable results and could be manufactured *at scale*."[19] In a lawsuit over allegedly fraudulent statements regarding Novavax's production of the COVID-19 vaccine, the court cited an executive's statement, "'We appear to have got past (certain) supply issues and are now being able to produce *at scale*.'"[20]

Legal scholars have also invoked concepts of scale in their attempts to explain of various legal doctrines. According to Richard Epstein, "the doctrine of efficient breach does not 'scale' as the number of parties increases."[21] Jonthan Adler wrote that the Clear Air Act's core provisions

---

rev'd in part and remanded, 45 F.4th 1236 (10th Cir. 2022) ("At the time, we had about 80 quota-bearing reps and little infrastructure around our sales reps....None of that infrastructure really existed at scale.").

[18] Purdue Pharma L.P. v. Accord Healthcare Inc., No. CV 20-1362-RGA, 2023 WL 2894939, at *6 (D. Del. Apr. 11, 2023).

[19] In re Talis Biomedical Corp. Sec. Litig., No. 22-CV-00105-SI, 2022 WL 17551984, at *2 (N.D. Cal. Dec. 9, 2022) (emphasis added).

[20] Sinnathurai v. Novavax, Inc., No. CV TDC-21-2910, 2022 WL 17585715, at *8 (D. Md. Dec. 12, 2022) (emphasis added).

[21] Richard A. Epstein, *Common Ground: How Intellectual Property Unites Creators and Innovators*, 22 GEO. MASON L. REV. 805, 815 (2015); Nicolas P. Terry, *The Opioid Litigation Unicorn*, 70 S.C. L. REV. 637, 667 (2019) ("Unfortunately, litigation is a blunt instrument that--to the extent it is effective at all--is best suited to well prescribed, narrow claims between individuals or between an individual and a corporation. Litigation *does not scale well*, and it is not a good tool for remedying mass social ills. It is also extremely inefficient both in its procedural costs (including attorneys' fees and other expenses) and a lack of timely resolution that almost guarantees that any recovery will be too late to help those who are currently suffering.") (emphasis added); Benjamin Ewing, *The*

that "focus on locally concentrated pollutants and a cooperative federalism model…[do] not scale cleanly to the control of a ubiquitous and globally dispersed pollutant such as carbon dioxide."[22] Francis Fukuyama wrote about the creation of norms and values that support legal enforcement that "[s]pontaneous order does not scale well: the larger the group size, the lower the likelihood that free riders will be detected or punished."[23]

Law and tech scholars have frequently used the language of scale to describe problems related to the extent of an activity and that activity's costs or harms.[24] For example, David Post wrote regarding the growth of the Internet that turning small into big "can be a tricky proposition indeed, because scaling problems--the problems that arise solely as a consequence of increasing size or increasing numbers--can be profound, and profoundly

---

Structure of Tort Law, Revisited: The Problem of Corporate Responsibility, 8 J. TORT L. 1, 7 (2015) ("[I]t begins to look unfair that tort law *does not scale* the extent of tortfeasors' liability to their degree of culpability or to the foreseeable extent of the harm they cause. Although in negligence law defendants are generally liable only for categories of harm that were reasonably foreseeable, under the so-called "egg-shell skull rule" they are liable for the full extent of a reasonably foreseeable harm they cause, even if the extent of the harm far exceeds normal expectations because of a hidden and unusual vulnerability in the victim.") (emphasis added).

[22] Jonathan H. Adler, *The Environmental Protection Agency Turns Fifty*, 70 CASE W. RES. L. REV. 871, 876 (2020).

[23] Francis Fukuyama, *Differing Disciplinary Perspectives on the Origins of Trust*, 81 B.U. L. REV. 479, 490 (2001). Scholars have even referenced the concept of scale when criticizing The Bluebook, writing "that the core problem with The Bluebook is that it is unwieldy. It still applies a twentieth-century method in a much larger, twenty-first century world. What worked for The Bluebook with twenty-six pages in 1926 *does not scale* well to its current 511 pages and beyond." Stephen M. Darrow & Jonathan J. Darrow, *Beating the Bluebook Blues: A Response to Judge Posner*, 109 MICH. L. REV. FIRST IMPRESSIONS 92, 95 (2011) (citing Richard A. Posner, *The Bluebook Blues*, 120 YALE L.J. 850, 859 (2011) (emphasis added)).

[24] *See e.g.,* Douglas Lichtman, *Copyright as Innovation Policy: Google Book Search From a Law and Economics Perspective*, 9 INNOVATION POL'Y & ECON. 55, 72 (2008) ("In a world with a large and ever-changing list of opt-out projects, authors would be forced to invest substantial sums finding each project and notifying each about their desire to participate. The problem would be even worse if some of those opt-out programs were designed strategically to make things difficult on authors, for instance, imposing high standards of proof before acknowledging that an opt-out really came from the correct copyright holder. (Infringers have an incentive to do just that because in an opt-out system, infringers benefit if authors find it too expensive to actually engage in the mechanism of opting out.) Overall, then, the problem with an opt-out program is that it does not *scale*.") (emphasis added); Naomi Appelman & Paddy Leerssen, *On "Trusted" Flaggers*, 24 YALE J. L. & TECH. 452, 473 (2022) ("[T]rusted flagging *does not scale*. If third parties wish to influence content moderation as it is currently practiced, they must leverage its automation.") (emphasis added).

difficult to solve."[25] Regarding the regulation of professional speech, Cassandra Burke Robertson and Sharona Hoffman wrote that, "[t]he scale of modern mass communication offers a much larger threat to the viability of traditional regulatory approaches."[26] And in the context of copyright infringement and enforcement, Annemarie Bridy wrote that "[w]ith each successive iteration, P2P network architecture has become not only more scalable and efficient, but also more perfectly adapted to 'massive infringement.' The key to effective online copyright enforcement in the P2P context is identifying and implementing enforcement strategies that are commensurately scalable."[27]

Scale is a common theme in the privacy literature too. According to Daniel Solove, "[r]eading privacy notices is a task that does not scale. There are simply too many privacy notices to read--people get notice fatigue."[28] Likewise, "[m]anaging one's privacy is a vast, complex, and never-ending project that does not scale; it becomes virtually impossible to do comprehensively."[29] Even one of us has used the concept without

---

[25] DAVID POST, IN SEARCH OF JEFFERSON'S MOOSE: NOTES ON THE STATE OF CYBERSPACE 30 (2009); *see also* Jeffrey L. Vagle, *Tightening the Ooda Loop: Police Militarization, Race, and Algorithmic Surveillance*, 22 MICH. J. RACE & L. 101, 123 (2016) (noting that police departments often attempt to justify algorithmic surveillance by relying on the common trope that "an experienced and talented officer can apply their knowledge and analytical skills to attain an imperfect version of predictive policing, but that the model does not scale well.").

[26] Cassandra Burke Robertson & Sharona Hoffman, *Professional Speech at Scale*, 55 U.C. DAVIS L. REV. 2063, 2100 (2022).

[27] Annemarie Bridy, *Is Online Copyright Enforcement Scalable?*, 13 VAND. J. ENT. & TECH. L. 695, 736 (2011); *see also* Thomas C. Rubin, *Leveraging Notice and Takedown to Address Trademark Infringement Online*, 37 COLUM. J.L. & ARTS 585, 591 (2014) ("Forcing platforms to choose between uncertain but potentially enormous liability, or policing its users in a way that does not scale and that undermines the utility of the service, is no choice at all."); Doug Lichtman, *Google Book Search in the Gridlock Economy*, 53 ARIZ. L. REV. 131, 142 (2011) ("Thus, opt-out, while better than nothing, does not seem to justify a fair use finding. It simply does not scale.").

[28] Daniel J. Solove, *The Limitations of Privacy Rights*, 98 NOTRE DAME L. REV. 975, 996 (2023).

[29] Daniel J. Solove, *The Myth of the Privacy Paradox*, 89 GEO. WASH. L. REV. 1, 5 (2021) ("Second, the CCPA does not scale well. The number of organizations gathering people's data is in the thousands. Are people to make thousands of requests? Opt out thousands of times? People can make a few requests for their personal data and opt out a few times, but this will just be like trying to empty the ocean by taking out a few cups of water."); *see also* Tyler Prime & Joseph Russomanno, *The Future of FOIA: Course*

explaining it, writing that the concept of informed consent "does not scale without losing its legitimacy."[30]

Outside the legal literature, commentators describing challenges in regulating information technologies commonly focus on scale. A great example is content moderation. Journalist Mike Mansick, who runs the popular website Techdirt, wrote that it is sometimes "difficult to get across to people 'the scale' part when we talk about the impossibility of content moderation at scale. It's massive."[31] Journalist Helena Pozniak wrote, "Moderating content online is messy, arbitrary and expensive – a huge headache for lawmakers and social media companies alike. While automating such moderation is essential at scale due to the sheer volume of traffic, it remains problematic."[32] The idea is that it's simply impossible to respond individually to the frequent posts of tens of millions of users.

Sometimes this same concept of scale has been invoked as an explanation of the limits of enforcement. Obama White House cyber security policy coordinator Howard Schmidt said in an interview that "[o]n cyber crime, we've always had an issue, as we have with other types of crime, which is there is oftentimes more than we can handle as law enforcement. [Our ability to respond] just does not scale."[33]

Corrections for the Digital Age, 23 COMM. L. & POL'Y 267, 298 (2018) ("Currently, manual "sanitization" [of public records] is expensive, time-consuming, susceptible to disclosure risks and *does not scale* as the volume of data increases.") (emphasis added); George S. Geis, *Automating Contract Law*, 83 N.Y.U. L. REV. 450, 476 (2008) ("Manual tagging [of documents] also takes a lot of time and *does not scale*.") (emphasis added).

[30] Neil Richards & Woodrow Hartzog, *The Pathologies of Digital Consent*, 96 Wash. U.L. Rev. 1461, 1500 (2019); *see also* Woodrow Hartzog & Evan Selinger, *The Internet of Heirlooms and Disposable Things*, 17 N.C. J. L. & TECH. 581, 588 (2016) ("Bad defaults on IoT devices are common and most users cannot easily patch them. The process is usually complicated. What's worse is that the updating process for the IoT does not scale well.").

[31] Mike Mansick, *The Scale of Content Moderation is Unfathomable*, TECHDIRT (Nov. 2, 2021), https://www.techdirt.com/2021/11/02/scale-content-moderation-is-unfathomable/.

[32] Helen Pozniak, *Tackling the Impossible Problem of Content Moderation*, ENGINEERING AND TECHNOLOGY (April 18, 2023), https://eandt.theiet.org/content/articles/2023/04/tackling-the-impossible-problem-of-content-moderation/.

[33] 185 INTELLECTUAL PROPERTY COUNSELOR NL 10.

II. TWO MEANINGS OF SCALE

All of these uses of "scale" in some way refer to the relationship between the extent of an activity and the related costs or benefits. Often scale is invoked to help explain the difficulty of (cost-effectively) increasing some activity or precaution. In general, these invocations of scale that focus on "increases" do little to explain the nature of the relationship between extent and costs or benefits. But that relationship is important.

In this part, we focus on one important distinction in the meaning of scale, particularly in relation to technology law. Specifically, we draw out the distinction between ideas of "scale as more" and "scale as different." We use "scale is more" to refer to a dynamic where the amount of an activity and the activity's costs and/or benefits have a generally linear relationship. "Scale is different" refers to situations where qualitatively new and different issues arise beyond some amount of the activity. Below, we identify the deployment of the concept of scale in law and technology rules and jurisprudence and make the argument that scale is being oversimplified.

When "scale is more," the costs or benefits of an activity increase as some linear function of the number of instances of that activity. In the simplest example, if the amount of harm caused by each unit of activity is x, then the total harm is simply x multiplied by the number of instances of the activity. But the important point here is not the specific function by which harms or benefits increase, it's that "scale" in this sense implies primarily an increase in magnitude of the same kinds of costs or benefits, not a qualitative change in the nature of those costs or benefits. People seem to rely upon "scale is more" logic often when thinking about whether an action can be increased indefinitely at an acceptable cost or when trying to set legal or policy thresholds. The idea behind "scale is more" logic is that, at some point, enough is enough.

In other contexts, scale is not simply more: "scale is different." In those cases, increases in magnitude do not only create more of the same kinds of harms or benefits; instead, they generate new kinds of costs or benefits that only emerge beyond some amount of the activity. In this sense, the relationship between the extent of an activity and its associated costs and benefits is not linear.

We certainly don't contend scale as "more" is irrelevant. Indeed, scale as more commonly matters. Law frequently requires a comparison of costs and benefits, and particularly aggregate costs and benefits of repeated activities. In the tort context, courts have long evaluated negligence by considering the probability of harm, the magnitude of that harm, and the burden of preventing the harm. (This is commonly referred to as the "Hand formula" because it was introduced by Judge Learned Hand in the famous case of *United States v Carroll Towing Co.*)[34] Using that formula, the scale of an activity matters at least to the aggregate harm (the total loss) whenever the loss associated with the activity increases in relation to the amount of that activity. Scale in this sense can also matter to the burden of preventing the harm to the extent the burden increases with the amount of the activity. Because the Hand Formula requires a comparison, the rate at which costs and benefits increase with the amount of the activity is highly relevant to determining the point at which the burden outweighs the discounted probability of loss. That same kind of comparison is also very common in the regulatory context, despite the difficulty of quantifying all of the relevant costs and benefits in relation to most technologies.[35]

But that is not the only way costs and benefits can relate to the amount of an activity. Take, for example, vaccination rates. Public health experts have long understood that, given the efficacy rate of a particular vaccine and the infectiousness of the disease against which it inoculates, a certain percentage of the relevant population needs to be vaccinated to achieve "herd immunity." Herd immunity is the idea that, once that percentage of the population is vaccinated, the disease is effectively prevented from spreading, even though no vaccine is 100% effective for any particular recipient. "Scale is different" when it comes to vaccines because the desired effect on a population doesn't exist at the individual level. Herd immunity is not achieved incrementally—it is not 80% achieved at 80% of the necessary vaccination rate. It only appears one a certain magnitude threshold has been met.

There's a corollary to that idea that we unfortunately have seen in real time in the COVID era. The failure to achieve herd immunity means

---

[34] 159 F.2d 169 (2d. Cir. 1947).
[35] CASS R. SUNSTEIN, THE COST-BENEFIT REVOLUTION (2018).

that the disease will continue to circulate, and some percentage of people will continue to be infected even when vaccinated. Here is where "scale as different" comes in: the failure to achieve herd immunity not only means that the same strain of COVID will continue to circulate, but the extent of that circulation also creates opportunities for new mutations and therefore new and different strains to emerge (ones not covered by the existing vaccines). Put differently, low vaccination rates don't just mean that more people will continue to get sick with the known disease (scale is more); it means that new and different harms will emerge (scale is different).

Of course, we are hardly the first to observe that sometimes new dynamics emerge at a certain magnitude of activity. Social and political scientists, economists, engineers, and people from a variety of backgrounds have demonstrated this fact repeatedly, and they have often incorporated it into the general wisdom of their fields. It's not even foreign to legal scholarship. Scholars have long understood that technologies can have "network effects" – the phenomenon where the value or utility a user derives from a good or service depends on the number of other users of that good or service.[36] That concept has been a particularly powerful way of understanding the value of networked technologies. Indeed, network effects are one explanation for natural monopolies – circumstances where the value of a service depends on number of users, and the number of users necessary to achieve sufficient value can't realistically be achieved by multiple parties.[37]

But we think the legal discourse has not fully appreciated that those are examples of a broader category, and that law and technology scholars do not always sufficiently consider the variety of ways in which scale can matter. We emphasize the more general distinction between "scale is more" and "scale is different" because attention to that distinction is important to determining the appropriate policy responses. Once we've identified a

---

[36] *See* Catherine Tucker, *Network Effects and Market Power: What Have We Learned in the Last Decade?*, ANTITRUST (Spring 2018), https://sites.bu.edu/tpri/files/2018/07/tucker-network-effects-antitrust2018.pdf.

[37] *See* Christopher S. Yoo & Daniel F. Spulber, *Antitrust, the Internet, and the Economics of Networks*, *in* OXFORD HANDBOOK OF INTERNATIONAL ANTITRUST ECONOMICS (2014) ("A given production technology is said to exhibit natural monopoly characteristics if it has a subadditive cost function, i.e., a single firm can supply the entire market demand at lower cost than could two or more firms.").

problem, we tend to conceive of the solution set in reference to the original framing of that problem. If we see privacy violations as instances of individualized harm perpetrated on the particular individuals whose information has been used, the legal frameworks are likely to be designed to remedy those individualized harms, even if at "scale" in the sense that there are a lot of those individualized harms. Unless policymakers are open to the idea that scale can create new and different problems that may require different kinds of solutions, the natural tendency will be to miss the real effect of scale in some contexts.

In exploring the concept of "scale" in tech regulation, Paul Ohm has argued that "[m]ost laws either treat all regulated actors the same or assume that twice as large means only twice as powerful and twice as harmful."[38] So, for example, "penalties for causing harm often multiply the number of individuals harmed by a set dollar figure, assessing $10,000 for each victim wiretapped or around $40,000 for each child monitored without parental consent."[39]

Ohm's critique is about the tendency to treat scale simply as more. Here, if an act is harmful x 1, when it is done at scale as "more", then it is harmful x multiplied by the number of instances. One privacy violation is bad. A thousand privacy violations are worse because it's an additional 999 instances of harm. That way of thinking tends to produce responses of the same structure: if the penalty for 1 violation is x, then the penalty for 1000 violations is just 1000(x). Ohm argues persuasively that a linear approach to scale is misguided, primarily because it fails to properly account for power dynamics. "Linearly bound regulation fails to reflect how the power and harm of some digital actors increase at much more than a linear, proportional rate. In at least three important ways, a platform with one billion users is more than one hundred times more powerful and potentially harmful than a company of ten million users."[40]

In our terminology, the problem with linearly bound regulation is that it ignores the ways that scale can be different. "[A] linear model fails to offer a proper moral accounting of the way human misery scales. We

---

[38] Paul Ohm, *Regulating at Scale*, 2 GEO. L. TECH. REV. 546 (2018).

[39] *Id.* (citing 18 U.S.C. § 2520 (2002), Adjustment of Civil Monetary Penalty Amounts, 16 C.F.R. pt. 1 (increasing FTC civil penalties to account for inflation)).

[40] Ohm, *supra* note 38, at 546–47.

might feel more impelled to prevent a small harm affecting one million victims out of one billion users than we would to prevent the same harm affecting only ten victims out of ten thousand users, even though they reflect the same rate of injury with the only difference being the size of the injurer. Second, purely digital platforms expand automatically into any territory that the Internet touches, meaning platform providers need not attend to local regulators and regulations. Third, size begets power, particularly for artificial intelligence, meaning we can expect more from globe-spanning digital platforms."[41]

We agree with Ohm that "[m]assive digital platforms thus raise significant concerns of potential harm that calls for a regulatory response that accounts for effects of size. From privacy to tort to contract to consumer protection to intellectual property laws, we should better account for the power and potential harm of size."[42] What was once a salesperson's attempt to wheedle you into buying that shirt now becomes a structured and systematized user interface that simultaneously affects billions. What was once a conspiracy theory exchanged at the bar becomes amplified to billions.

However, as we argue below, we think that recognizing the ways scale can be different does even more than allow us to account for the magnitude of power accumulation. It's not just that lawmakers and judges are getting the math wrong when thinking about scale too simplistically and linearly. Sometimes when the instances of something related to information technologies significantly increase, whole assumptions about actions and consequences must be challenged.

III. A FULLER PICTURE OF THE CONSEQUENCES OF SCALE

In this part, we describe at least four ways that scale can mean "different" and not just "more." We don't mean to suggest that this is an exhaustive account of the effects of scale, nor do we argue that these are entirely distinct from each other. We describe these effects to highlight the ways that scale can be different, and to help guide policymakers toward

[41] *Id.*
[42] *Id.*

more nuance in considering the effects of scale and the corresponding range of policies regarding new technologies.

A. The Population Affected Could Change

One important way in which scale is different is that the scale of an activity might change the population that is affected by that activity. In isolation, certain practices only seem to implicate those actors that are directly involved. For example, when a company collects a person's information, we might assume that only that person's privacy and autonomy was at risk. Your browsing history probably doesn't directly reveal anything about me, so that's a "you" problem. This isn't always true, of course, even in isolated cases. For example, if your family member takes a DNA test and gives that information to a company, you are exposed because of the strong overlaps in familial DNA.[43] But generally speaking, our default frame of analysis for isolated actions focuses only on the people involved, either directly or because they have "skin in the game," by being somewhere in the supply chain or otherwise standing to gain or lose something as a result of the action.

At scale, someone's actions might implicate not just related third parties, but the interests of entire populations with shared characteristics. For example, Salome Viljoen has argued that "data-collection practices of the most powerful technology companies are aimed primarily at deriving (and producing) population-level insights regarding how data subjects relate to others, not individual insights specific to the data subject. These insights can then be applied to all individuals (not just the data subject) who share these population features."[44] According to Viljoen,

---

[43] Law enforcement officers have recently solved a number of "cold" cases using forensic genetic genealogy – matching the DNA profile of the suspect to living family members whose genetic profiles are known, often because those family members voluntarily tested with a commercial ancestry testing company like 23andMe. *See, e.g.*, *Multiple Cold Cases Solved with Assist from Attorney General's Dna Forensic Genetic Genealogy Program*, ATTORNEY GENERAL OF WASHINGTON (July 11, 2022), https://www.atg.wa.gov/news/news-releases/multiple-cold-cases-solved-assist-attorney-general-s-dna-forensic-genetic; Joe Hernandez, *Genealogy DNA is used to identify a murder victim from 1988—and her killer*, NPR (Sept. 8, 2022), https://www.npr.org/2022/09/08/1121542171/genealogy-dna-murder-stacey-lyn-chahorski-henry-frederick-wise-michigan-georgia.

[44] Viljoen, *supra* note 7, at 578.

> This population-level economic motivation matters conceptually for the legal regimes that regulate the activity of data collection and use; it requires revisiting long-held notions of why individuals have a legal interest in information about them and where such interests obtain. The status quo of data-governance law, as well as prominent proposals for its reform, approach these population-level relational effects as incidental or a byproduct of eroded individual data rights, to the extent that they recognize these effects at all. As a result, both the status quo and reform proposals suffer from a common conceptual flaw: they attempt to reduce legal interests in information to individualist claims subject to individualist remedies, which are structurally incapable of representing the interests and effects of data production's population-level aims. This in turn allows significant forms of social informational harm to go unrepresented and unaddressed in how the law governs data collection, processing, and use.[45]

Something similar can be said about AI training sets. Since the goal of training is for the system to learn patterns, especially patterns that were not visible to human observers, the size and representativeness of the training set matters enormously to the functioning of the AI system. Indeed, many of the documented problems of bias in AI systems are attributable to training sets that were not sufficiently diverse. For our purposes here, the point is that these systems aren't useful primarily because of individual bits of information they learn from specific inputs; their real value is in recognition of patterns that are only learnable when the data set is of a certain size. Those population-level insights are then frequently baked into algorithms in ways that have much more systemic effect than do the bits of information themselves.

B. Emergent Problems

One of the most obvious ways that scale can be different is that new and qualitatively different kinds of problems can emerge at certain thresholds. That is what we described with respect to the insufficient uptake of COVID vaccines: the lack of herd immunity allowed the virus to circulate at a scale that didn't just make more individuals sick with the same

---

[45] *Id.*

variant, it enabled the emergence of new variants that would affect even the vaccinated.

Kathleen Creel and Deborah Hellman have described the ways that algorithmic decision-making at scale can produce meaningfully different problems as compared to individualized decisions on the same issues.[46] Specifically, Creel and Hellman argue that arbitrary individualized decisions (hiring decisions based on irrelevant characteristics, for example) generally don't rise to the level of moral concern because there's no strong interest in any individual decision being non-arbitrary (as opposed to non-biased). But, they argue, widespread adoption of arbitrary algorithmic systems does cause harm because use of those algorithms has the effect of systematically locking people out of opportunities (jobs, credit, etc.).[47] Arbitrariness at scale creates a new and different problem that isn't just the sum of the harms of individual decisions.

Misinformation is also a good example of scale as different. Individual pieces of misinformation are, of course, potentially harmful, because they can affect the behavior of those who receive it. And that harm is surely multiplied as more misinformation circulates. That means that there's an important scale is more effect in this context. But there are also important ways in which scale of misinformation is different. One is that, beyond a certain point of circulation of any particular piece of misinformation, that information may start to be perceived as credible by more people. That might be characterized as an example of a change in the population affected: people who might be skeptical of some particular piece of misinformation might find it credible because of the degree of its circulation. Scale even plays a key factor in distinguishing the idea of misinformation from disinformation. Ryan Calo, Chris Coward, Emma Spiro, Kate Starbird, and Jevin D. West have helpfully distinguished the two concepts along the lines of intent and scale, writing:

> [Misinformation is the] erroneous or misleading information to which the public may be exposed, engage with, and share….Disinformation refers to a purposive

---

[46] Kathleen Creel & Deborah Hellman, *The Algorithmic Leviathan: Arbitrariness, Fairness, and Opportunity in Algorithmic Decision-Making Systems*, 52 CAN. J. PHIL. 26 (2022).

[47] *Id.*

> strategy to induce false belief, channel behavior, or damage trust. Misinformation is usually discrete or standalone, as when a neighbor shares a false rumor or overhears a misleading exchange. Disinformation tends to take the form of a multifaceted campaign with a predetermined financial, political, or other objective. Disinformation campaigns blend orchestrated action and organic activity, relying on the participation of willing but unwitting online audiences.[48]

Misinformation can exist in isolation, but disinformation requires scale for success.

There's another sense in which the scale of misinformation and disinformation can be different and not just more. Specifically, there's a point at which there's so much misinformation, particularly in certain places or among certain groups, that it threatens destruction of belief in the idea of truth. That is a widely recognized feature of Russian disinformation: its purpose is more than just to convince people of the specific claims in individual pieces of misinformation, it is to sow chaos and create doubt that there is any such thing as truth, particularly in official information. That "flood the zone" strategy is premised entirely on the recognition that scale is different: beyond some point, the problem isn't really the specific misinformation, it's the epistemic free-for-all.

Facial recognition is another example of how scale is different. Individual uses of facial recognition technology can cause a variety of harms. If the technology is trained on disproportionately white faces, that technology is much more likely to misidentify non-white people, causing any number of discrete harms to the people misidentified. More extensive use of such biased technology might repeat that harm over many people, increasing the aggregate harm. In that sense, scale is more. But widespread deployment of facial recognition technology across a range of settings

---

[48] Ryan Calo, Chris Coward, Emma Spiro, Kate Starbird, & Jevin D. West, *How Do You Solve a Problem Like Misinformation?*, SCIENCE ADVANCES (Dec. 8, 2021), https://www.science.org/doi/epdf/10.1126/sciadv.abn0481.

threatens a total surveillance society and a complete loss of obscurity.[49] In that sense, scale is very different.

Website scraping is a similar example. People probably expect websites unrestricted by passwords and privacy settings to be accessed by all kinds of people as part of their normal use of a computer. But when bots scrape social media websites like LinkedIn and Twitter and preserve snapshots of those same websites at scale, things get weird. Not only can massive simultaneous access crash a server, but once scraped, snapshots of information that people might have assumed would just be viewed by a person becomes sortable, cheaply stored, easily aggregated, effortlessly shared, perfectly preserved, and repurposed such that insights over time can be added up to paint pictures of human behavior that were unlikely part of people's threat modeling when they originally posted on social media. On top of that, their photos can power databases that turn your face into a tracking beacon, obliterating our collective practical anonymity and ability to hide in plain sight. Bad times at scale.

C. Challenge the Assumption of the Original Problem

Actions at scale might also cause us to challenge the original assumptions regarding incentives, implementation, and costs and benefits of an activity. In a way, this is kind of a subset of "emergent problems." But we separate it out to highlight the difference between recognizing new problems that only exist at scale ("emergent problems") and revisiting the originally perceived value or cost of an action ("challenging assumptions").

---

[49] Woodrow Hartzog & Evan Selinger, *Surveillance as Loss of Obscurity*, 72 WASH. & LEE L. REV. 1343, 1345–46 (2015) ("[W]e argue that the concept of "obscurity," which deals with the transaction costs involved in finding or understanding information, is the key to understanding and uniting modern debates about government surveillance."); Woodrow Hartzog & Evan Selinger, *Increasing the Transaction Costs of Harassment*, 95 B.U. L. REV. ANNEX 47 (2015); Evan Selinger & Woodrow Hartzog, *Obscurity and Privacy*, *in* SPACES FOR THE FUTURE: ROUTLEDGE COMPANION TO PHILOSOPHY OF TECHNOLOGY (Joseph Pitt & Ashley Shew eds., 2018), https://www.routledge.com/Spaces-for-the-Future-A-Companion-to-Philosophy-of-Technology/Pitt-Shew/p/book/9780415842969; *see also* Woodrow Hartzog & Frederic Stutzman, *The Case for Online Obscurity*, 101 CAL. L. REV. 1, 5 (2013) ("We argue the case for obscurity for two reasons. First, we argue that obscurity is a common and natural condition of interaction, and therefore human expectation of obscurity will transfer to the domains in which we spend time, both physical and virtual. Second, we argue that obscurity is a desirable state because we are protected by an observer's inability to comprehend our actions, and therefore social practice encourages us to seek obscurity."); Woodrow Hartzog & Frederic Stutzman, *Obscurity by Design*, 88 WASH. L. REV. 385 (2013).

For example, for years policymakers have considered information privacy issues as an individual's problem. Any harm they felt would likely be felt acutely as a person due to the nature of their information revealed. Public revelations of private information might cause emotional distress, chilling effects, or financial harm. Leaked health information might cause others to act differently towards you. If your credit card number gets out, a thief might wrongfully make charges to your account. The harms were individual harms visited on particular people; the costs and benefits of legal responses or of tools for avoiding those harms were understood in relation to the nature of the individual harms. But at scale, lawmakers might (and should) conceive of the risk of harm differently.

For example, Salome Viljoen has argued that "[p]rivacy and data-governance law have traditionally governed forms of private interpersonal exchange in order to secure the benefits of data-subject dignity or autonomy. Yet as data collection and use become key productive activities (i.e., economic activities that define the contemporary economy as an information economy), new kinds of information-based harm arise. There is growing evidence of the role that digital technology plays in facilitating social and economic inequality. Digital-surveillance technologies used to enhance user experience for the rich simultaneously provide methods of discipline and punishment for the poor. Algorithmic systems may reproduce or amplify sex and race discrimination. Even seemingly innocuous data collection may be used in service of domination and oppression. The pursuit of user attention and uninterrupted access to data flows amplifies forms of identitarian polarization, aggression, and even violence. Such evidence suggests that social processes of datafication not only produce violations of personal dignity or autonomy, but also enact or amplify social inequality."[50] As a result, "alongside traditional concerns over individual autonomy, the social inequalities that result from data production are also forms of informational harm."[51]

We might say something similar about the structuring of our regulatory system around the value of choice: it's not that each individual choice is hard, but once we have adopted a system that prioritizes choice, it

---

[50] Salomé Viljoen, *A Relational Theory of Data Governance*, 131 YALE L.J. 573, 580–81 (2021).

[51] *Id.* at 582.

throws consumers into a world where they're constantly inundated by it. This is the fundamental problem with notice and consent as a model of privacy regulation. There's a sense in which this might be considered a scale as more problem: each of these choices has some cost, and there's just a tipping point beyond which the costs overwhelm the benefits of choice. But we think this is an example of the ways that scale can illuminate problems with the original understanding of the costs and benefits of a choice-focused model. It's not just that the costs of each choice will mount. It's that, through the lens of scale, we can see that the model produces an environment that is not conducive to meaningful choice even in the individual instances.

Dark patterns might be another example like this: practices or design features that don't really register as harms in isolation seem different when viewed in the aggregate. What is important about these examples is that the regulatory model works outwardly from characterization of the individual instance, dismissing harms as *de minimus* or perhaps even seeing each instances as net beneficial), but only because each action is viewed individually.[52] At scale, things look very different, even in terms of how we see individual instances.

Instances at scale also raise the possibility of normalizing them, discouraging public resistance, and encouraging conformity, which might cause people to reevaluate their initial resistance to problems. In research with Evan Selinger and Johanna Gunawan, one of us has argued that the ubiquity and ultimate mundanity of *de minimus* privacy encroachments at scale can both distort and bypass our ability to critically reflect upon the danger of exposure.[53] We wrote, "Two normalization dynamics that revolve around repeated exposure, "unexceptional habituation" and "favorably disposed normalization," might also play important roles in shaping how people view surveillance. *Unexceptional habituation* occurs when people in liberal Western democracies take ubiquitously encountered surveillance systems for granted—seeing them as so commonplace and mundane they

---

[52] *See generally* Max L. Veech & Charles R. Moon, *De Minimis Non Curat Lex*, 45 MICH. L. REV. 537 (1947); Frederick G. McKean Jr., *De Minimis Non Curat Lex* , 75 U. PA. L. REV. 429 (1927).

[53] Woodrow Hartzog, Evan Selinger, and Johanna Gunawan, *Privacy Nicks: How the Law Normalizes Surveillance*, WASH. U. L. REV. (forthcoming 2023), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4384541.

are not worth thinking about critically….The psychological dynamic of *favorably disposed normalization*, whereby the routine experience of being surveilled inclines people to view surveillance as acceptable, if not desirable, might significantly influence what people believe is appropriate privacy policy."[54]

Scale can also change incentives about the problem. Consider the idea of "too big to fail" as an issue of scale. Tort law justifies the economic loss doctrine which denies plaintiffs liability for purely monetary harms by

---

[54] *Id.* (citing Clare Southerton & Emmeline Taylor, *Habitual Disclosure: Routine, Affordance, and the Ethics of Young Peoples Social Media Data Surveillance*, 6 SOC. MEDIA & SOC'Y (2020); Evan Selinger & Judy Rhee, *Normalizing Surveillance*, 22 N. EUR. J. PHIL. 49 (2021)). We expanded upon this idea, writing:

> One plausible psychological basis for favorably disposed normalization is the impact of believing something is normal. Thinking something is normal does not necessarily entail a commitment to deeming that thing ethical. Nevertheless, normality judgments often are accompanied by positive affective experiences. For example, imagine someone believes using Facebook is ethically problematic but normal. That person might feel less badly about using Facebook than someone who believes the practice is ethically problematic and abnormal. The difference in how people feel has implications for governance. The person with a stronger felt sense of discomfort might have a greater incentive to quit the platform. After all, people frequently complain about ethical violations. But taking the next step of committed action can require more than intellectual awareness that change is needed. Given the practical value of heightened moral motivation for rectifying injustice, in some circumstances, "beliefs about normality might be more important than moral beliefs." But how do people develop the belief something is normal? According to experiments conducted by philosophy and cognitive science professor Joshua Knobe and psychology professor Adam Bear, both prescriptive and descriptive information matter if people know how good something is perceived and how prevalent it is. Nevertheless, simply "increasing the frequency of something occurring," such as surveillance more becoming more prevalent, can lead people to perceive it as "more normal," not just increasingly widespread. Supporting evidence for this thesis exists in the experimental literature on environmental messaging. Alternatively, one might explain the dynamic of favorably disposed normalization through the psychological process of rationalization. From this perspective, people generally are motivated to see themselves positively, as moral, intelligent, and in control of their lives. To maintain this narrative and minimize inconsistency when making decisions that seem unethical, stupid, or unfree, they often subconsciously turn to rationalization. Put otherwise, being aware of a gap between how we would like to act and how we actually behave can be stressful because it creates cognitive dissonance. Rationalization is ameliorative because it can minimize or dispel cognitive dissonance. Rationalization provides people with a means to convince themselves they should see their situation differently—that seemingly troubling behavior is justifiable, tolerable, and in some cases, even laudable.

Id. (citing Nathanael J. Fast & Arthur S. Jago, *Privacy Matters…or Does It? Algorithms, Rationalization, and the Erosion of Concern for Privacy*, 31 CURRENT OP. PSYCH. 44 (2020)).

saying that, in most cases, the scale of economic liability would be so big as to make recovery an apocalyptic event for defendants and the court system itself. If actions cause too much harm or if actors are just too big, you just might be off the hook. An anticipation of scale changes how we might think about the problem, which leads us to the final way scale can be different and not just more.

D. The Solution Set Can Change

For lawmakers, the most important consequence of scale might be that it can affect the set of solutions they might turn to for a given problem. For example, in Ryan Calo's paper exploring the virtues of using "visceral notice" to solve technology-related problems, he wrote about how once electric cars began to be produced and adopted at scale "[r]egulators in the United States and Europe became concerned that electric or hybrid vehicles do not emit an engine noise. There is evidence that the absence of such noise leads to more pedestrian collisions. Rather than blanketing the sidewalks with signs warning pedestrians that some cars are now silent, these regulators investigated another expedient: requiring fake engine noises that change depending on the distance of the car as a natural warning embedded in the pedestrian's experience."[55] The scale of engines affected human behavior to the point where a design solution that took advantage of the societal expectation that the way to tell if a car is coming is to listen for the sound of an engine.

These solutions would only exist if the relevant technologies existed at scale because a few cars driving around here and there are less likely to shape our collective expectations. Scale also presents a collective action problem—if all cars lost the engine noise, over time people would stop subconsciously listening for it. This presents a potential threshold legal intervention question for lawmakers. At what point should the law intervene or stop caring? Scale can help us understand not just when we need more regulation, but also when we need less.

A greater interrogation into actions at scale might also help us identify regulatory choke points or the futility of enforcement. For example, filing copyright lawsuits against those who use file sharing software to

---

[55] M. Ryan Calo, *Against Notice Skepticism in Privacy (and Elsewhere)*, 87 NOTRE DAME L. REV. 1027, 1036 (2022).

download files in violation of copyright law might seem like a futile game of whack a mole, which is why secondary liability targeted the means and instrumentalities, which were fewer and more manageable at scale. In other instances, it might lead lawmakers to try and determine whether most instances of an activity at scale, say a prominent botnet operator state-sponsored purveyor of disinformation was responsible for a lion's share of misconduct. Even if there are billions of instances of an activity, lawmakers' solution set should depend upon whether there are 100 or 1 million bad actors.

IV. HOW TO TAKE SCALE SERIOUSLY IN LAW AND POLICY

A more developed concept of scale would have significant implications for technology law and policy. The most fundamental change might be to the way scholars and policymakers reason through problems involving data, algorithms, sensors, and actuators. Some of the most common ways we reason in tech law is through metaphors and threat modeling modeling, and to properly do that we must account for the effect of scale.[56] To that end, we join scholars like Ryan Calo who have called for law and technology to adopt a more sophisticated approach to technology and its relationship to humans and human goals by drawing from science and technology studies (STS) and related disciplines.[57]

STS scholars have explored how human behavior can change how a technology works at scale for decades. In his exploration of the relationships between STS and law and technology, Calo highlights what could have been gained if legal scholars had more explicitly embraced STS earlier, including more nuanced metaphors, more case studies, and fewer redundancies. Calo cites two downsides that arise from law and technology overlooking STS. First, failing to deeply engage with STS denies the field of law and tech wisdom and nuance. Additionally, law and tech scholarship often falls into some of the very traps STS grew up to avoid, such as too strong a sense of technological determinism and the misguided idea that technology will shape behavior in one single way and no other. This wisdom

---

[56] See Ryan Calo, *Modeling Through*, 71 DUKE L. J. 1391 (2022); Ryan Calo, *Robots as Legal Metaphors*, 30 HARV. J. L. TECH. 209 (2017).

[57] Ryan Calo, *The Scale and the Reactor* (Apr. 9, 2022), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4079851.

can also help lawmakers better project how and in what situations scale might be different, and not just more.

We recommend a simple rule of thumb for all policymakers and scholars approaching law and technology issues: *start with scale*. People studying and working in law and technology often seem to think about these technologies by starting with individual or atomized instances of technological deployments and working outward only later, if at all. Privacy is a great example. Over the past fifty years, it seems that lawmakers have based most of privacy law around giving people control over their personal information. Control is a laudable goal in theory and in isolation. It serves our interests in autonomy, one of the most foundational values in nearly all Western legal frameworks. But informational self-determination fails at scale. We think basing privacy law and policy on concepts like consent and individual data subject rights is the wrong starting point because it ignores how these approaches work, change, and ultimately fail at scale. Consent models start with the efficacy of an individual choice and then work outward.

But if lawmakers were to assume that scale is inevitable for all issues implicating the use of technology and the course of events could take several different paths at scale, our rules would likely adhere closer to the original purported goals of legal frameworks and our rulemaking process would likely be more efficient. If lawmakers had started with scale for privacy law, they might have embraced more structural, social, and relational approaches that focused on mitigating abuses of power instead of prioritizing control. They might have better recognized that consent is easily extracted through manipulative design at scale, and that exercising any meaningful control is overwhelming in the aggregate, and that our perceived agency is typically illusory in mediated environments.[58] They also might have recognized that the collective wisdom from trillions of individual self-motivated decisions might not reflect or account for collective and societal concerns.

---

[58] WOODROW HARTZOG, PRIVACY'S BLUEPRINT: THE BATTLE TO CONTROL THE DESIGN OF NEW TECHNOLOGIES (2018); Woorow Hartzog, *The Case against Idealising Control*, 4 EUR. DATA PROT. L. REV. 423 (2018); Neil Richards & Woodrow Hartzog, *Pathologies of Digital Consent*, 96 WASH. U. L. REV. 1461 (2019); Evan Selinger & Woodrow Hartzog, *The Inconsentability of Facial Surveillance*, 66 LOYALA L. REV. 101 (2019).

Beyond changing the starting point for analysis of law and technology problems, we think a more developed conceptualization of scale would have three important implications. First, lawmakers should assume that regulator approaches should be continually (or at least periodically) reassessed to confront how the popular adoption of new tools changes costs and benefits. Additionally, we argue that a better conception of scale supports a greater adoption of the precautionary principle. Finally, we argue that scale could shape how legal institutions are designed and the choice of remedies in law and technology disputes.

When it comes to technology, it's been clear for some time that rules should be periodically revisited. Technologies work within society to change practices and people's perceptions. Expectations and laws that were based on technological practices that existed in 1985 (and business models that leverage those technologies) no longer make sense in 2023. But technological development isn't the only reason our rules related to technology need to be continually updated.

Sometimes it's not clear how scale is different until it manifests. Even when it is clear how law and technology will interact if everyone adopted them, policymakers often do not feel motivated to act upon speculation. But the reality of scale can be compelling, as we've seen with the plague of misinformation and disinformation on social media. Mass deception was always possible with social media, but lawmakers didn't take it seriously until it was widespread enough to be a serious threat to undermining elections at scale (and they arguably have yet to meaningfully respond). Acting upon scale concerns would be a way to interrupt regulatory inertia by requiring a periodical reassessment of the costs and benefits of both rules and tools. It's a way to build policy responses to anticipate that the changes of scale will happen.

Scale being different can also justify a precautionary approach to new technologies. For so long advocates of innovation have criticized early legal intervention where technology is involved because they claim it could hinder the development of new and useful tools. But there's a real danger to waiting so long to fully understand the social impacts of technologies that when clarity finally arrives these tools and systems are too entrenched to resist. In STS scholarship this is referred to as the "Collinridge dilemma," and it gives more nuance to what some law and tech scholars describe as

the "avocado ripeness" problem. (Not yet…not yet…not yet……..too late.). Calo has called this dynamic a kind of "constant state of watchful paralysis."[59] The precautionary principle is even further justified if lawmakers were to periodically revisit rules, injecting a regular democratic deliberation into lawmaking to fight harmful creep.[60]

When scale is different, its consequences can serve as a good reason to keep problems from getting to the tipping point. Once the cat is out of the bag, lawmakers now have new problems to solve, changing the regulatory cost. Facial recognition is a great example of this dynamic. And it's a cautionary tale for how lawmaker are currently treating generative AI. Lawmakers around the world are hesitant to regulate facial recognition without being able to a specific individual harm such as emotional distress, financial loss, a diminished reputation, or significant denial of autonomy and dignity through lack of consent.

Sometimes facial recognition leads to these kinds of harms. But other times, the real cost of these surveillance systems is social, involves the creation of a power imbalance and eventual exploitation of that power, and is hard to see at the individual level.[61] Meanwhile the most dangerous surveillance tool ever created is becoming entrenched in the digital systems that run our lives and is being normalized with every Face ID scan, Snapchat filter, airline check-in, and IoT doorbell.[62] We are in a brief window where it the cost of substantive prohibitions on these tools would be acceptable, but the more we come to rely upon them, the greater the cost. At some point, we will have no choice but to tolerate tools that have irrevocably exposed us and permanently diminished our privacy with virtually no democratic accountability.

Finally, scale compels important questions about institutional design and legal remedies. Specifically, should policymakers address a problem through regulation, or is it better revolved through private

---

[59] Calo, *supra* note 57.

[60] *See* BRETT FRISCHMANN AND EVAN SELINGER, RE-ENGINEERING HUMANITY (2018); Hartzog, Selinger, & Gunawan, *supra* note 53..

[61] Hartzog, Selinger, & Gunawan, *supra* note 53.

[62] *Id.; see also* Daniel Wroclawski, *Facial Recognition Is Coming to Your Neighborhood Through Home Security Cameras and Video Doorbells*, CONSUMER REPORTS (May 2, 2023), https://www.consumerreports.org/electronics/privacy/facial-recognition-and-home-security-cameras-video-doorbells-a9500287020/.

litigation? Additionally, is the best approach to a problem that involves scale to seek monetary relief, or would an injunction address the problems of scale better? Lastly, what about other strategies to enact policy that go beyond liability rules, such as taxation, property interests, and human rights law?

The answers to these questions will vary, and scores of scholars have volumes to say on when and why certain strategies are desirable over others. Our point here is simply to emphasize that the different consequences of scale should be a part of this calculus. For example, if scale changes the population affected by a set of actions to include third parties otherwise unrelated to the relevant actors, then litigation alone might not be the best response because people besides the plaintiffs and the defendants will be affected. This is true even if class action relief is possible. Class actions respond to "scale is more." They simply aggregate the harm of all the class members. There is no obligation in class action lawsuits to address externalities or accommodate unrelated (but incidentally affected) third parties. Scale also might affect the remedies sought in litigation, counseling an injunction that affects everyone potentially affected in the future instead of monetary relief which only directly benefits the plaintiffs.

Issues of scale might also affect the structure and grant of authority to regulatory agencies. If problems only emerge (or appear to emerge) at scale, it's possible that federal agencies might need rulemaking power that doesn't hinge upon a showing of individualized harm. They might also need better information disclosure rules to achieve more transparency, a superstructure to encourage collaboration with researchers to improve expertise, since issues of scale might not be apparent through individualized case studies, past litigation, and anecdotes. Problems that emerge at scale might also cut across various domains like health, finance, the environment, and consumer protection, necessitating rules to encourage harmony and collaboration, or possibly even a new regulatory agency designed to collect information, provide expertise, and assist with enforcement efforts.[63]

---

[63] *See* Ryan Calo, *The Case for a Federal Robotics Commission*, BROOKINGS (2014), https://www.brookings.edu/wp-

Scale might also direct lawmakers to go beyond the standard suite of regulatory liability rules and embolden property rights to better enable market dynamics (though we remain skeptical of property rights in information as a way to protect people's privacy).[64] Or lawmakers might consider a human rights approach that is less likely to wilt as part of a cost/benefit analysis or political compromise.[65]

Because some problems only manifest at scale, lawmakers might craft legislation that only kicks in at scale. We're already seeing examples of this at the federal and state levels. Senators Elizabeth Warren and Lindsey Graham have targeted "dominant platforms" in legislation that imposes, among other things, robust duties of loyalty, care, confidentiality, and mitigation upon only those businesses that among other things, have more than 50 million US-based monthly active users, 1 billion users worldwide, or an annual revenue of more than $550 billion.[66] If enacted, this law would only affect those operating at the largest scale. A California senator has just proposed a legislative framework that would regulate only those "frontier" AI systems that operated at the largest and most robust scale, capturing those problems that exist at the most extreme edges of artificial intelligence and overlooking those that operate at the smallest or more modest scales.[67]

Or perhaps concerns over scale might encourage lawmakers to look to grants, taxation, and other fiscal approaches to better capture negative

---

content/uploads/2014/09/RoboticsCommissionR2_Calo.pdf; Woodrow Hartzog, *Unfair and Deceptive Robots*, 74 MD. L. REV. 785 (2015); *see also* Chris J. Hoofnagle, Woodrow Hartzog, & Daniel J. Solove, *The FTC can rise to the privacy challenge, but not without help from Congress*, BROOKINGS (Aug. 8, 2019), https://www.brookings.edu/articles/the-ftc-can-rise-to-the-privacy-challenge-but-not-without-help-from-congress/.

[64] *See* Pamela Samuelson, *Privacy As Intellectual Property?*, 52 STANFORD L. REV. 1125 (2000); Ignacio Cofone, *Beyond Data Ownership*, 43 CARDOZO L. REV. 501 (2021).

[65] *See* Guido Calebrasi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV 1089 (1972).

[66] Digital Consumer Protection Commission Act of 2023, S. _, 118th Cong. (2023), https://www.warren.senate.gov/imo/media/doc/Tech%20Bill_Full%20Text.pdf.

[67] Billy Perrigo, *Exclusive: California Bill Proposes Regulating AI at State Level,* TIME (Sep. 13, 2023), https://time.com/6313588/california-ai-regulation-bill/ ("It proposes that systems that require above a certain quantity of computing power to train—a threshold not specified by the bill—be subject to transparency requirements. It proposes establishing legal liability for "those who fail to take appropriate precautions" to prevent unintended consequences and malicious uses of advanced AI systems.").

externalities of a practice or particular design.[68] Grants, deductions, taxable items, and more all reflect policy preferences that can and should be sensitive to issues of scale. Lawmakers could make it more expensive to use a technology as scale increases or create rules that don't activate until a particular size or different scale threshold is met.

V. CONCLUSION

Discussions of scale abound in law and policy discussions related to automated technologies. But the concept feels underspecified in ways that might matter. Intuitively, scale means simply "more." But in this essay we've argued scale can also mean "different." More or different communities might be implicated when people deploy technology at scale. New problems might arise at scale, or we might some assumptions we had once held about the nature of the deployment. Finally, when technologies exist at scale, some legal, social, market-driven, or design-based solutions might become available or be taken off the table. Lawmakers should take scale more seriously and, in doing so, could better respond to the challenges of automated tools.

---

[68] *See, e.g.*, Salome Viljoen & Amanda Parsons, *Valuing Social Data*, COLUMB. L. REV. (forthcoming 2023), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4513235. For examples in encouraging innovation, *see, e.g.*, W. Nicholson Price II, *Grants*, 34 BERKELEY TECH. L.J. 1 (2019); Arti K. Rai, Rachel Sachs & W. Nicholson Price II, *Cryptic Patent Reform Through the Inflation Reduction Act*, HARV. J.L. & TECH. (forthcoming 2023), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4402378.