SMU | Dedman School of Law

SMU Law Review

Volume 75 | Issue 3 Article 2

2022

Algorithmic Destruction

Tiffany C. Li
University of New Hampshire, Franklin Pierce School of Law

Author(s) ORCID Identifier:

https://orcid.org/0000-0001-5015-0726

ALGORITHMIC DESTRUCTION

*Tiffany C. Li**

ABSTRACT

Contemporary privacy law does not go far enough to protect our privacy interests, particularly where artificial intelligence and machine learning are concerned. While many have written on problems of algorithmic bias and data deletion, this Article introduces the novel concept of the "algorithmic shadow" and explains the new privacy remedy of "algorithmic destruction," also known as algorithmic disgorgement or machine unlearning.

The algorithmic shadow describes the persistent imprint of training data that has been fed into a machine learning model and used to refine that machine learning system. This shadow persists even if data is deleted from the initial training data set, which means privacy rights like data deletion do not solve for the new class of privacy harms that arise from algorithmic shadows. Algorithmic destruction (deletion of models or algorithms trained on misbegotten data) has emerged as an alternative, or perhaps supplementary remedy and regulatory enforcement tool to address these new harms.

This Article introduces two concepts to legal scholarship—the algorithmic shadow and algorithmic destruction. First, the Article defines the concept of the algorithmic shadow, a novel concept that has so far evaded significant legal scholarly discussion, despite its importance in changing understandings of privacy risks and harms. Second, the Article argues that data deletion does not solve for algorithmic shadow harms and advocates for the development of new privacy remedies to address these new harms. Finally, the Article introduces algorithmic destruction as a potential right and remedy, explaining its theoretical and practical applications, as well as potential drawbacks and concerns.

TABLE OF CONTENTS

https://doi.org/10.25172/smulr.75.3.2.

* Assistant Professor of Law, University of New Hampshire Franklin Pierce School of Law; Fellow, Yale Law School Information Society Project; Affiliate, University of North Carolina Center for Information, Technology, and Public Life. The author thanks Brenda Dvoskin, Mailyn Fidler, Asaf Lubin, Jon Penney, Blake Reid, Alan Rozenshtein, Andrew Sellars, Ari Ezra Waldman, Laurin Weissinger, and Andrew K. Woods for their helpful feedback and support.

I. INTRODUCTION

WHEN you are arrested in America, you are awarded certain rights—among them is the right to a fair and speedy trial.[1] You may expect certain things from your justice system, and, broadly speaking, in many cases, you would be right to expect them. For example, you might expect that there will be a judge (preferably learned, preferably wise, hopefully fair) who will guide the proceedings and who will, should you be found guilty, determine a fair and proportionate punishment for your misdoings.

You might not expect the artificial intelligence (AI). That is, you might not expect that, throughout the criminal justice process, some key analysis influencing your ultimate legal consequences will be performed, not by humans but by automated decision-making systems.[2] You might not expect the AI, but having heard of it, you might, in fact, *suspect* the AI. You would be right to do so.

In 2016, the investigative reporting outlet ProPublica broke the then-shocking news that many criminal sentencing decisions (which we would

1. *See* U.S. Const. amend. VI.

2. *See* Julia Angwin, Jeff Larson, Surya Mattu & Lauren Kirchner, *Machine Bias*, ProPublica (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing [https://perma.cc/TFL5-JS78] (criticizing automated "risk assessments . . . used to inform decisions about who can be set free at every stage of the criminal justice system").

like to believe are fair and just) were made using an automated system that employed a form of AI known as machine learning.[3] The AI helped determine appropriate sentences based on an analysis of factors that supposedly predicted the likelihood for any particular defendant to re-offend.[4] ProPublica's investigation found that the system, Correctional Offender Management Profiling for Alternative Sanctions, or COMPAS, produced consistently discriminatory results.[5] Using COMPAS for sentencing resulted in judgments awarding heavier sentences to Black defendants and lighter sentences to White defendants, regardless of the severity of the crime or other relevant factors.[6]

The problems of algorithmic bias in the criminal justice system are not new. Countless well-respected scholars and advocates have dedicated great portions of their careers to raising awareness about, and potentially solving issues related to, algorithmic bias and algorithmic fairness in the criminal justice system and across all sectors of human activity.[7] This Article will not rehash decades of debate on the subject. Nor will it provide an exhaustive overview of the legal and philosophical dimensions of deletion in privacy conceptualization.[8]

---

3. *See id.*

4. One of the "most widely used assessment tools in the country . . . gave ProPublica the basics of its future-crime formula—which includes factors such as education levels, and whether a defendant has a job." *Id.*

5. *Id.* But for another perspective on COMPAS's potential bias, see Sam Corbett-Davies, Emma Pierson, Avi Feller & Sharad Goel, *A Computer Program Used for Bail and Sentencing Decisions Was Labeled Biased Against Blacks. It's Actually Not That Clear*, WASH. POST (Oct. 17, 2016), https://www.washingtonpost.comnews/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas [https://perma.cc/J8S8-SHMR].

6. *See* Angwin, Larson, Mattu & Kirchner, *supra* note 2.

7. *See, e.g.*, Joy Buolamwini & Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, 81 PROC. MACH. LEARNING RSCH. 1 (2018), https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf [https://perma.cc/5YY7-RG28]; SAFIYA UMOJA NOBLE, ALGORITHMS OF OPPRESSION: HOW SEARCH ENGINES REINFORCE RACISM (2018); VIRGINIA EUBANKS, AUTOMATING INEQUALITY: HOW HIGH-TECH TOOLS PROFILE, POLICE, AND PUNISH THE POOR (2018); Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671 (2016); Ifeoma Ajunwa, *The Paradox of Automation as Anti-Bias Intervention*, 41 CARDOZO L. REV. 1671 (2020); Joshua A. Kroll, Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson & Harlan Yu, *Accountable Algorithms*, 165 U. PA. L. REV. 633 (2017); Latanya Sweeney, *Discrimination in Online Ad Delivery: Google Ads, Black Names and White Names, Racial Discrimination, and Click Advertising*, ACM QUEUE 1 (Apr. 2, 2013), https://dl.acm.org/doi/pdf/10.1145/2460276.2460278 [https://perma.cc/U9HR-QL9W]; Kate Crawford & Ryan Calo, Opinion, *There Is a Blind Spot in AI Research*, 538 NATURE 311 (2016), https://www.nature.com/articles/538311a.pdf [https://perma.cc/VX9D-YQGY]; Pauline T. Kim, *Data-Driven Discrimination at Work*, 58 WM. & MARY L. REV. 857 (2017); Crystal S. Yang & Will Dobbie, *Equal Protection Under Algorithms: A New Statistical and Legal Framework*, 119 MICH. L. REV. 291 (2020). The past few decades have witnessed an explosion of research into problems of machine bias and algorithmic fairness, including multiple conferences dedicated to the subject. *See, e.g.*, *ACM Conference on Fairness, Accountability, and Transparency*, ACM FACCT CONFERENCE, https://facctconference.org/index.html [https://perma.cc/2YP9-PQ7F]; *ACM Conferences*, ACM, https://dl.acm.org/conferences [https://perma.cc/Q7W4-KAYZ].

8. For an in-depth discussion of the legal and philosophical dimensions of deletion, see VIKTOR MAYER-SCHÖNBERGER, DELETE: THE VIRTUE OF FORGETTING IN THE DIGITAL AGE (2009); MEG LETA JONES, CTRL + Z: THE RIGHT TO BE FORGOTTEN (2016).

This Article introduces the novel concept of the "algorithmic shadow"[9] and uses the algorithmic shadow as a lens through which to view the failures of current privacy and AI laws in dealing with the realities of AI and machine learning. In particular, this Article critiques data deletion. It is also the first scholarly work to substantively critique the novel privacy remedy of algorithmic disgorgement, also known as algorithmic destruction, machine unlearning, or machine learning model deletion.

Data deletion is not a meaningful right or remedy in the advent of machine learning systems. Data deletion as a remedy does not comport with the spirit of privacy laws that seek to make victims whole, and it does not solve the harms that privacy laws seek to prevent.[10] The failure of data deletion to resolve the privacy losses caused by algorithmic shadows highlights the ineffectiveness of data deletion as a right and a remedy.

Various proposals for algorithmic destruction have emerged as an alternative, or perhaps a supplement, to data deletion.[11] Algorithmic disgorgement has so far been used as an enforcement tool requiring organizations to delete machine learning models and algorithms developed with misbegotten data.[12] While algorithmic disgorgement may resolve some of the failures of data deletion—namely, the failure to address the harms of the algorithmic shadow—this remedy and potential right is not without its own drawbacks.[13]

What is the algorithmic shadow? Simply put, when you enter a set of specific data to train a machine learning model, that data creates an impact on the model.[14] Even if you later delete data from the training data set, the already-trained model still contains a persistent "shadow" of the deleted data. In other words, the act of deleting data from the initial training data set has no bearing on the machine learning model that has already been trained on that data. The algorithmic shadow thus describes the persistent imprint of the data that has been fed into a machine learning model and used to refine that machine learning system.

This Article has three goals. First, it will introduce and define the concept of the algorithmic shadow, a novel concept that has so far evaded significant legal scholarly discussion despite its importance in future discussions of AI and privacy law.[15] Second, the Article explains why the algorithmic shadow exposes and exacerbates existing problems with data deletion as a privacy right and remedy.[16] Finally, the Article examines

9. With thanks to Andy Sellars, who suggested this phrasing the first time I explained the kernel of an idea that eventually became this Article way back in 2019.

10. *See infra* Section IV.B.

11. *See* Kate Kaye, *The FTC's New Enforcement Weapon Spells Death for Algorithms*, PROTOCOL (Mar. 14, 2022), https://www.protocol.com/policy/ftc-algorithm-destroy-data-privacy [https://perma.cc/QG2P-Y2NJ].

12. *See id.*

13. *See infra* Section V.D.

14. *See* Rachel Wolff, *What Is Training Data in Machine Learning?*, MONKEYLEARN (Nov. 2, 2020), https://monkeylearn.com/blog/training-data [https://perma.cc/FV83-XW4C].

15. *Infra* Part III.

16. *Infra* Part IV.

algorithmic destruction as a potential right and remedy and compares it with data deletion in relieving algorithmic shadow harms.[17]

II. BACKGROUND

This Article builds upon past work on privacy and AI, critiquing the failure of privacy laws to address the technical realities of computing and algorithmic harms. In a prior co-authored article, *Humans Forget, Machines Remember: Artificial Intelligence and the Right to Be Forgotten*, my co-authors and I wrote on the fundamental failure of "the right to be forgotten" and the general concept of a right to erasure, specifically in the context of AI and machine learning.[18] We argued that deletion is a flawed privacy remedy due to the technical limitations of deleting information in machine learning systems.[19]

It is important to understand and critique the right to deletion, or deletion, as a privacy remedy, if for no other reason than its prevalence in a multitude of privacy laws and proposals for privacy reform.[20] Many modern privacy laws like the European Union's General Data Protection Regulation (GDPR) include a right to deletion, right to erasure, right to be forgotten, or similar rights.[21] While there is yet no federal omnibus privacy law in the United States,[22] state laws like California's Consumer Privacy Act include rights to deletion as privacy remedies.[23]

Deletion rights appear in some form in many privacy laws, and the intent behind these laws is well-meaning. However, deletion fails as a remedy for privacy harms due to technical limitations in machine learning. To fully explore this, we must start by understanding how deletion works in machine learning systems. In *Humans Forget, Machines Remember*, my co-authors and I urged academics to pursue more interdisciplinary research and policymakers to work with researchers in order to craft laws that better account for the technical realities of machine learning.[24]

This Article attempts to further that discussion by casting light on one underdeveloped notion in machine learning discussions—the algorithmic shadow—as an example of how the law and legal scholarly discussions have misconstrued AI and machine learning. Data deletion fails as a remedy due to problems like the algorithmic shadow's persisting harms, but the fact that issues like algorithmic shadows are ignored also exposes the

17. *Infra* Part V.

18. Eduard Fosch Villaronga, Peter Kieseberg & Tiffany Li, *Humans Forget, Machines Remember: Artificial Intelligence and the Right to Be Forgotten*, 34 COMPUT. L. & SEC. REV. 304 (2018).

19. *Id.* at 305, 308–10.

20. *See id.* at 305–07, 310–12.

21. *See* Council Regulation 2016/679, art. 17, 2016 O.J. (L 119) 1 (EU) [hereinafter GDPR].

22. *See* Nuala O'Connor, *Reforming the U.S. Approach to Data Protection and Privacy*, COUNCIL ON FOREIGN RELS. (Jan. 30, 2018), https://www.cfr.org/report/reforming-us-approach-data-protection [https://perma.cc/83LX-6HYW].

23. California Consumer Privacy Act of 2018, CAL. CIV. CODE § 1798.105 (West 2018).

24. *See* Villaronga, Kieseberg & Li, *supra* note 18, at 305, 311–13.

underlying problems of vagueness, uncertainty, and lack of technical specificity in legal discussions surrounding AI.

A. Definitions

In an article that discusses problems stemming from a lack of technical specificity and misapplication of law to technology, it is perhaps most fitting to begin with definitions. This Article is not a dictionary, and the author recognizes that many of the following defined terms have multiple debatable definitions that are relevant in different fields and subfields. However, for the purposes of this Article, the following definitions are useful in providing a foundational common ground for readers seeking to understand this Article and, more broadly, the legal academic discourse surrounding AI and machine learning.

1. Artificial Intelligence

AI is a popular buzzword,[25] and like most buzzwords, it can mean everything and nothing at the same time. However, definitions are important, particularly as the law around AI develops. It is important that future legal proposals not conflate terms like AI, machine learning, and algorithms.[26]

At our current stage of technological development, AI is much less fantastical than the phrase may sound at first; we are not talking about superintelligence, which is merely theoretical today.[27] Rather, when we refer to AI in law and policy discussions, we often focus on relatively simple decision-making or prediction systems.[28]

AI refers to any form of intelligence that is man-made or artificial, generally relating to the idea of a constructed machine intelligence that could potentially equal the intelligence of a human being.[29] This Article primarily discusses the form of AI known as machine learning.

---

25. Bernard Marr, *What Is the Difference Between Artificial Intelligence and Machine Learning?*, Forbes (Dec. 6, 2016, 2:24 AM), https://www.forbes.com/sites/bernardmarr/2016/12/06/what-is-the-difference-between-artificial-intelligence-and-machine-learning/?sh=7aadcd342742 [https://perma.cc/B7NE-FMGR].

26. This Article attempts to use these terms carefully. However, the author acknowledges that it may contain overlooked missteps in usage of such terms, which, to be fair, is something that happens often in our laws and legal scholarly discussions surrounding artificial intelligence. Perhaps counterintuitively, there is an argument for scholarship focusing on the technical details of artificial intelligence, written to be intelligible to computer scientists, to coincide with the typical approach to scholarship: writing couched in the broad, highly abstracted language that policymakers will undoubtedly end up using in creating actual laws, and that judges will use in crafting actual decisions on those laws.

27. *See generally* Nick Bostrom, Superintelligence: Paths, Dangers, Strategies (2014).

28. *See, e.g.*, Angwin, Larson, Mattu & Kirchner, *supra* note 2.

29. *See* Michael Cheng-Tek Tai, *The Impact of Artificial Intelligence on Human Society and Bioethics*, 32 Tzu Chi Med. J. 339, 339 (2020).

2. Algorithms

An algorithm is a set of instructions.[30] In the computing context, an algorithm is a set of instructions or rules for a computer to do certain things in order to complete a task or solve a problem.[31] For machine learning, algorithms are sets of instructions or rules for a computer follows to discover patterns in data and make predictions or find solutions.[32]

3. Data

Data is information, including numbers, text, and facts.[33] In computing, data describes pieces of information that can be used, processed, and stored by computers.[34] In machine learning, training data is the initial set of data that humans prepare and input into a computer to create a machine learning model.[35] Testing data is the data used to test the accuracy or aptness of a machine learning model that has been trained on a separate set of training data.[36]

4. Machine Learning Models

A machine learning model is created by tasking a computer with running an algorithm on a given set of data until the computer "learns" a mathematical model to use to achieve whatever goal the human programmers set for the machine learning system.[37] For example, if programmers want to design a machine learning system to identify which ice cream flavors are the most delicious, one way to do this could be to train a machine learning system on training data related to ice cream flavors and deliciousness ratings. Programmers could task a computer to analyze that training data and discover the underlying patterns that relate to flavors and deliciousness. The output from a computer running an algorithm on that training data could be a machine learning model that finds patterns

---

30. *Algorithm*, CAMBRIDGE DICTIONARY, https://dictionary.cambridge.org/us/dictionary/english/algorithm [https://perma.cc/63PT-6CFP] (defining *algorithm* as "a set of mathematical instructions or rules that, especially if given to a computer, will help to calculate an answer to a problem").

31. *Understanding Algorithms in Computer Science*, INT'L UNIV. GENEVA, https://www.iun.ch/en-en/blog/computer-science/algorithm-computer-science-definition-and-understanding [https://perma.cc/34ZZ-UNSE].

32. *See Machine Learning*, IBM CLOUD (July 15, 2020), https://www.ibm.com/cloud/learn/machine-learning [https://perma.cc/FRH6-8WSX].

33. *Data*, CAMBRIDGE DICTIONARY, https://dictionary.cambridge.org/us/dictionary/english/data [https://perma.cc/H3JL-CLYV]. This is a broad generalization and excludes other definitions of data, including Data, the Soong-type synthetic intelligence android who served as second officer aboard the USS *Enterprise-D* and later the USS *Enterprise-E*, in the science fiction television franchise *Star Trek*. *See Data (Star Trek)*, WIKIPEDIA, https://en.wikipedia.org/wiki/Data_(Star_Trek) [https://perma.cc/RS5G-ZJZK].

34. *Data*, *supra* note 33.

35. *See* Wolff, *supra* note 14 ("The features, tags, and relevancy of your training data will be the 'textbooks' from which your model will learn.").

36. *See id.*

37. *What Is a Machine Learning Model?*, MICROSOFT (Dec. 30, 2021), https://docs.microsoft.com/en-us/windows/ai/windows-ml/what-is-a-machine-learning-model [https://perma.cc/TJ2A-W8GU].

between flavors and the deliciousness and uses those patterns to make predictions on deliciousness of future ice cream flavors.

B. How Machine Learning Works

One of the most popular forms of AI today is machine learning.[38] Machine learning is a process for drawing conclusions or making predictions by feeding new data into a machine learning model, an algorithm that is first trained on an initial data set.[39]

As noted by David Lehr and Paul Ohm, among others, when legal scholars speak about machine learning, the result is often muddled[40]—a consequence perhaps of the fact that legal scholars sometimes comment on matters of law and technology despite having no formal training in the technology at hand. Lehr and Ohm argue that legal scholars overly abstract the concept of machine learning, analyzing legal issues related to it without fully paying attention to the technical realities of machine learning.[41] Fully explaining the technical details of machine learning is beyond the scope of this Paper. However, it is necessary to first build a foundation of understanding for key machine learning concepts in order to proceed with a discussion of the legal issues, especially given that this Article introduces a technical concept—the algorithmic shadow.

In broad strokes, here is how machine learning works. A human or group of humans ("the programmer") begins with a goal.[42] The programmer decides to create a machine learning system to achieve a particular goal, like generating predictions or drawing insights on existing data.[43] The programmer then collects or creates an initial set of data to use as training data;[44] best practices include setting aside a set of data to test the model.[45] The programmer uses statistical methods to ready the training data for use, including cleaning, coding, and categorizing the data.[46] The programmer next designs an algorithm (a set of instructions) for the computer to use.[47] This algorithm will instruct the computer on how to process the data and how to achieve whatever goal the programmer initially

38. For example, a 2020 study from the consulting firm Deloitte showed that 67% of surveyed companies claimed to be using machine learning at the time of the survey, and 97% planned to use it within the next year. Deloitte AI Inst., State of AI in the Enterprise: Thriving in the Era of Pervasive AI 6 (3d ed. 2020), https://www2.deloitte.com/content/dam/Deloitte/cn/Documents/about-deloitte/deloitte-cn-dtt-thriving-in-the-era-of-persuasive-ai-en-200819.pdf [https://perma.cc/2UFN-XTMC].

39. *See* Meredith Broussard, Artificial Unintelligence: How Computers Misunderstand the World 93–97, 101–14 (2018).

40. *See* David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, 51 U.C. Davis L. Rev. 653, 655–56 (2017).

41. *See id.* at 655–57, 661–64.

42. *See id.* at 672–74.

43. *See id.* at 672–74, 717.

44. *See id.* at 665, 677.

45. *See id.* at 684–86.

46. *See id.* at 681, 683.

47. *See id.* at 688–89.

designed the machine learning system to achieve.[48] The programmer then uses the initial set of data (training data) to train the machine learning model by asking the computer to process the data using the algorithm, find the underlying patterns in the data, and use those patterns to create a model that can then achieve the goal set out by the programmer.[49]

There are different ways to design a machine learning system. The two primary methods are supervised learning and unsupervised learning.[50] With supervised learning, the programmer knows what they are looking for from the outset.[51] The programmer identifies the type of outcome they seek and classifies, or labels, the data for that purpose.[52] For example, a supervised machine learning system may be designed to predict the likelihood that a student will graduate high school based on parental income. When designing this system, the programmer knows what they are looking for: the relationship between parental income and student graduation. So, they classify these two variables in the data and ask the machine to predict student graduation based on parental income given an analysis of the labeled data set.

With unsupervised learning, the programmer does not necessarily know what they are looking for.[53] Rather, they use unsupervised learning processes to determine the underlying relationships or patterns within the data.[54] For example, maybe the programmer obtains another data set on student graduation rates, but now, in addition to data on parental income, they also have data on student age, race, gender, height, favorite subject, favorite color, least favorite ice cream flavor, and mother's maiden name. The programmer is not sure exactly what pattern to look for here, but the programmer can instruct an unsupervised machine learning system to identify clusters in the data or relationships between different variables.[55] This can help the programmer identify some useful conclusions from the data.

Once the computer has generated a trained machine learning model, the programmer can use the machine learning system with new data.[56] The programmer can input new data, and the machine learning system will give them new outputs based on that new data by using the algorithms in the machine learning model.[57] The computer can then refine the machine learning model with each new piece of data, updating the algorithms it uses as new data confirms or disputes the patterns it previously

48. *See id.* at 671.
49. *See id.* at 671, 695–96.
50. *See* Julianna Delua, *Supervised vs. Unsupervised Learning: What's the Difference?*, IBM Cloud Blog (Mar. 12, 2021), https://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning [https://perma.cc/2444-GRZS].
51. *See id.*
52. *See id.*
53. *See id.*
54. *See id.*
55. *See id.*
56. *See* Lehr & Ohm, *supra* note 40, at 701–02.
57. *See id.*

identified.[58] In this way, the machine learning system will continue to "learn" and refine its accuracy, thereby improving its "intelligence."

C. Applications of AI and Machine Learning

In our present time, most popular uses of AI rely on machine learning.[59] Machine learning is currently used in criminal justice settings, as can be seen in the COMPAS algorithm example[60] and in predictive policing (e.g., using algorithms to determine which neighborhoods require heavier police presence).[61] These machine learning systems show up not only in law enforcement and criminal justice but in many areas of life, including education,[62] employment,[63] housing,[64] and more.

While many modern uses of machine learning are relatively unseen by the public, there are a few areas where individuals can visibly witness the direct impact of algorithmic decision-making systems. Perhaps the most familiar arena that people understand to utilize automated decision-making systems and machine learning is finance; credit scores, loan determinations, and trading strategies are all based on sophisticated algorithms.[65]

There are many sectors of society that now use AI and machine learning systems, and it is critical that we evaluate each use to ensure that technologies are deployed in fair and just ways. It is particularly important that we evaluate the legal protections for AI and machine learning systems making determinations that can curtail or protect fundamental civil liberties and human rights. To protect these values in society, we must invest in studying, auditing, and improving these systems, in addition to strongly considering the limits beyond which an algorithmic decision-making system should not be implemented. We must consider how we can create legal protections for abuse of such systems, as the consequences of such abuse will can be significant, both for individuals and society as a whole. As such, this Article focuses primarily on AI applications that implicate these important issues.

---

58. *See id.* at 702.

59. Richard M. Re & Alicia Solow-Niederman, *Developing Artificially Intelligent Justice*, 22 Stan. Tech. L. Rev. 242, 244–45 (2019).

60. *See* Angwin, Larson, Mattu & Kirchner, *supra* note 2.

61. *See* Lehr & Ohm, *supra* note 40, at 655, 658. This Article does not discuss the "robot judge," who does not exist and may not merit a "who" at all. For further discussion on the topic of robot judges, see Ryan Calo, *Robots as Legal Metaphors*, 30 Harv. J.L. & Tech. 209, 217–19 (2016); Re & Solow-Niederman, *supra* note 59.

62. *See, e.g.*, Lindsey Barrett, *Rejecting Test Surveillance in Higher Education*, 1 Mich. St. L. Rev. (forthcoming 2023).

63. *See, e.g.*, Ifeoma Ajunwa, *Algorithms at Work: Productivity Monitoring Applications and Wearable Technology as the New Data-Centric Research Agenda for Employment and Labor Law*, 63 St. Louis U. L.J. 21, 23 (2018).

64. *See, e.g.*, Eubanks, *supra* note 7, at 10–13.

65. *Machine Learning (in Finance)*, Corp. Fin. Inst., https://corporatefinanceinstitute.com/resources/knowledge/other/machine-learning-in-finance [https://perma.cc/7V6R-6XQ3].

D. AI AND THE LAW

Recently, an increasing number of laws dealing specifically with AI have developed in different regions and states. In 2021 alone, seventeen states introduced bills or resolutions regarding regulation of AI.[66] Additionally, in 2021, the European Commission released the Artificial Intelligence Act, the European Union's first legal framework for regulating AI.[67]

Some of the laws relevant to regulation of AI are general privacy laws, such as the General Data Protection Regulation, which includes provisions on rights to deletion and other privacy rights that relate to, and are complicated by, AI.[68] In the absence of a federal privacy law, U.S. states continue to pass new state privacy laws.[69]

Ari Ezra Waldman theorized the recent wave of privacy laws and proposals on the state and federal level as the "second wave of privacy law," which is focused more on corporate compliance and personal control over data, rather than on regulating population-level harms caused by data-driven industries, power imbalances in the data ecosystem, and systemic inequities.[70] This second wave stands in contrast to the first wave of privacy law, which Waldman identifies as the early notice-and-choice and privacy policy requirement regime.[71] In analyzing these waves of privacy law, Waldman ultimately urges critical privacy theorists to continue expanding the conversation around privacy law in order to better protect privacy rights and values of justice and equity.[72]

Some of the failures of contemporary privacy laws in addressing the realities of AI and machine learning could be ameliorated by a "third wave" privacy law approach that evaluates data's larger infrastructure and its effects.[73] Laws that focus on enforcing individuals' rights or placing limitations on companies ignore the problems associated with the larger data ecosystem, including the downstream harms suffered by individuals whose data is collected, aggregated, and used to develop machine learning systems.[74] The persistence of the algorithmic shadow is one of these systemic issues that modern privacy laws do not solve for, perhaps

66. *Legislation Related to Artificial Intelligence*, NAT'L CONF. STATE LEGISLATURES (Jan. 5, 2022), https://www.ncsl.org/research/telecommunications-and-information-technology/2020-legislation-related-to-artificial-intelligence.aspx [https://perma.cc/WX8Z-CHBC].

67. Eve Gaumond, *Artificial Intelligence Act: What Is the European Approach for AI?*, LAWFARE (June 4, 2021, 11:50 AM), https://www.lawfareblog.com/artificial-intelligence-act-what-european-approach-ai [https://perma.cc/ZD39-V3DR].

68. Ben Wolford, *What Is GDPR, the EU's New Data Protection Law?*, GDPR.EU, https://gdpr.eu/what-is-gdpr [https://perma.cc/4DJC-BFB7].

69. *See* Thorin Klosowski, *The State of Consumer Data Privacy Laws in the US (And Why It Matters)*, N.Y. TIMES: WIRECUTTER (Sept. 6, 2021), https://www.nytimes.com/wirecutter/blog/state-of-privacy-laws-in-us [https://perma.cc/TUN7-B4ZT].

70. Ari Ezra Waldman, *The New Privacy Law*, 55 U.C. DAVIS L. REV. ONLINE 19, 21–22, 37–39 (2021).

71. *See id.* at 19, 22–24.

72. *See id.* at 40–41.

73. *See id.* at 38–41.

74. *See id.* at 38–39.

due to the lack of critical, systems-based thinking that Waldman conceptualizes.[75]

III. ALGORITHMIC SHADOWS

Deletion is a flawed remedy for privacy harms in machine learning systems due to the presence of algorithmic shadows that persist even after data is deleted. The shadow of one's data on a machine learning model can still cause harm to an individual and to groups that relate to that individual. These harms can be privacy harms, algorithmic harms, and discrimination harms. Thus, it is important that we understand what causes algorithmic shadow harms and how to minimize or protect against them.

A. The Persistence of Algorithmic Shadows

The persistence of the algorithmic shadow is a relatively simple concept that is underdeveloped in the legal discourse surrounding machine learning and privacy. This concept relates to a problem with data deletion in machine learning systems. Even after data is deleted from the data set a programmer used to train a machine learning model, that act of deletion has no impact on the already-trained model—the data used to train the model already influenced the development of the model. Deleting the data afterward does not remove the persistent shadow of the data in the algorithm. In other words, the algorithmic shadow is what remains even after data deletion has been completed. It is a problem that highlights the flaws with deletion as a privacy remedy and privacy right, as well as some of the problems with current legal thinking around privacy and AI generally.

Consider a machine learning system trained to identify whether an animal in a photo is a cat. If you train the image recognition algorithm to recognize photos of cats, and you only input photos of orange cats, the resulting algorithm trained on this data might be more likely to recognize orange cats as cats and may be less able to predict when a grey-furred animal is a cat. Let's say you decide later that the algorithm you have designed and trained is now too "biased" in favor of identifying orange animals as cats. You want to fix this. You might choose then to delete all photos of orange cats from the training data set. However, this would have no impact on the algorithm that has already been trained on the data. The algorithm will still over-predict orange animals to be cats. Even if the data is deleted from the original data set, the "shadow" of the data remains on the algorithm.

To further explain the persistence of algorithmic shadows, consider the following disgusting metaphor[76]: A conceptual artist who moonlights as a serial killer harvests one hundred human ears and uses them to create a

75. *See id. passim*.

76. Thank you to the participants of the Law and Technology Workshop for inspiring and listening to this metaphor described in detail in person.

sculpture. He places each of the ears on a surface of wet clay, which dries into a terrible blob covered in human ears. One of the serial killer's victims miraculously survives and now requests his ear back. Detectives and art gallery personnel are able to remove the ear from the surface of the sculpture, but the imprint of the ear persists. While the victim may have the ear back to use at his disposal, he can never erase the imprint of his ear on the resulting sculpture. The victim still suffers harm—both the harm of the physical violence and the psychological and emotional harm of seeing the imprint of his ear on the sculpture. The earless man is analogous to the data subject of a privacy violation, whose data (ear) has been stolen and misused by a privacy violator (serial killer) to create a machine learning model (clay and ear sculpture). While the data subject can request that their data be deleted (that the ear be pulled from the sculpture), the subject cannot remove the persistent algorithmic shadow—the imprint of their data on the resulting machine learning model.

Data deletion does not fix the psychological and emotional harms of knowing that your data has been used and misused to train a machine learning model, particularly if that model is offensive or harmful to you.[77] For example, the use of facial recognition in law enforcement surveillance is controversial.[78] Many people do not support the practice and likely would not consent to the inclusion of photos of their faces into databases for use in developing facial recognition models for surveillance.[79] Even if a surveillance company is required to delete individuals' photos, the harm of having one's photo included in a facial recognition surveillance machine learning model remains because the algorithmic shadow still persists as an imprint on the machine learning model itself and on the predictions it makes for surveillance.

B. The Harms of Algorithmic Shadows

This Article addresses two categories of harms related to the persistence of algorithmic shadows: first, privacy and algorithmic harms that are worsened by the presence and lack of acknowledgement of algorithmic shadows; and second, unique harms caused by algorithmic shadows.

1. Existing Harms Made Worse

The lingering presence of algorithmic shadows worsens some existing privacy harms that are already mentioned in privacy laws, including

77. *See* Danielle Keats Citron & Daniel J. Solove, *Privacy Harms*, 102 B.U. L. Rev. 793, 796–97, 841–44, 855–56 (2022); Daniel J. Solove, *A Taxonomy of Privacy*, 154 U. Pa. L. Rev. 477, 486 (2006). *But see* Ann Bartow, Response, *A Feeling of Unease About Privacy Law*, 155 U. Pa. L. Rev. PENNumbra 52, 56–57, 60–62 (2006).

78. *See* Nicol Turner Lee & Caitlin Chin, *Police Surveillance and Facial Recognition: Why Data Privacy Is Imperative for Communities of Color*, Brookings (Apr. 12, 2022), https://www.brookings.edu/research/police-surveillance-and-facial-recognition-why-data-privacy-is-an-imperative-for-communities-of-color [https://perma.cc/WBS2-XTJA].

79. *See id.*

harms related to use of automated decision-making systems. Algorithmic bias and discrimination are known problems with use of automated decision-making systems—bias can creep in a number of ways, and many algorithmic systems can be implemented in ways that lead to discrimination.[80] While some laws are beginning to target the discriminatory impact of some of these systems,[81] few have specifically noted the impact of algorithmic shadows and the difficulties with data deletion as a remedy or as protection from harm.

Furthermore, due to lack of understanding regarding complications like lasting algorithmic shadows, many privacy laws rely on protections like rights to deletion, even though those protections are insufficient.[82] Thus, harms related to privacy loss and surveillance are generally exacerbated due to false reliance on solutions that do not actually work. Because we do not accept algorithmic shadow harms, we continue to use deletion as a flawed remedy, further enforcing the justification for surveillance everywhere and failing to protect vulnerable populations.

2. New Privacy Harms

In addition to the types of privacy harms that current privacy laws seek to remedy, there are also new harms that have arisen due to the increased use of algorithmic systems. With the increase in mass surveillance and algorithmic decision-making systems, more attention has been paid to group privacy harms.[83] These are harms suffered by groups and communities due to privacy violations, as opposed to exclusively individual harms.[84] These group privacy harms are particularly important in light of algorithmic shadows. While data deletion and contesting automated decisions may be viable solutions for individuals who are upset about the inclusion of their data in certain algorithmic systems, these solutions do not provide a remedy for the privacy infringements suffered by entire groups—communities and groups of individuals who may be similar to the data subject, and who thus may find themselves at the mercy of algorithmic systems predisposed towards them in ways that are harmful.

These sorts of group privacy harms can include "discrimination by association" or discrimination against people who are classified or profiled as part of a protected class regardless of whether they belong to it.[85] Additionally, individuals may suffer inferential privacy harms. This is a general category of group privacy harms that stem from the idea that

80. *See supra* notes 2–6 and accompanying text.

81. *See, e.g.*, Kate Kaye, *This Senate Bill Would Force Companies to Audit AI Used for Housing and Loans*, PROTOCOL (Feb. 8, 2022), https://www.protocol.com/enterprise/revised-algorithmic-accountability-bill-ai [https://perma.cc/DN2C-BHDN].

82. *See infra* Section IV.B.

83. *See* Luciano Floridi, *Group Privacy: A Defence and an Interpretation*, *in* GROUP PRIVACY: NEW CHALLENGES OF DATA TECHNOLOGIES 83 (Linnet Taylor, Luciano Floridi & Bart van der Sloot eds., 2017).

84. *See* Citron & Solove, *supra* note 77, at 818–19, 855.

85. Sandra Wachter, *Affinity Profiling and Discrimination by Association in Online Behavioral Advertising*, 35 BERKELEY TECH. L.J. 367, 373 (2020).

algorithmic systems can analyze data to make assumptions—inferences—about a person; these inferences may even relate to data not present in the initial training set.[86] In fact, these inferences may relate to individuals whose data are not included in the training set at all.[87]

In a similar vein, the concept of the algorithmic shadow also sheds light on the importance of third-party privacy harms or distributed privacy harms. Distributed privacy harms are the privacy harms suffered by individuals who are not themselves the primary data subjects in the act of data collection or processing, perhaps best exemplified by the harms of genetic privacy violations.[88] For example, if an individual submits their DNA sample to a DNA processing company like 23andMe, not only do they give up some of their own genetic privacy to the corporation, but they also give up the privacy of everyone in their genetic line.[89] Thus, the initial data subject's biological family members also suffered a privacy loss, though they were not involved at any stage of the actual data collection or processing. Privacy laws currently do not provide a remedy or protection for individuals who suffer third-party, proxy privacy harms.[90] The increase in algorithmic systems and the persistence of algorithmic shadows means there are new and growing classes of people whose privacy rights are at stake, through no fault or action of their own.

IV. DATA DELETION

Many privacy laws provide a right to request data deletion, and enforcement agencies mandate deletion in privacy law enforcement.[91] However, data deletion fails as a privacy remedy for two reasons. First, privacy laws that require data deletion fail to define data deletion, leaving significant room for confusion, error, and many instances in which following the letter of the law does not reflect the spirit of that law at all. Second, data deletion fails as a remedy for the misuse of data in machine learning due to the technical realities of machine learning itself.

A. Data Deletion as Right and Remedy

Discussion around contemporary privacy laws often advocates for

86. *See* Sandra Wachter & Brent Mittelstadt, *A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI*, 2019 Colum. Bus. L. Rev. 494, 497–98 (2019).

87. *See* Alicia Solow-Niederman, *Information Privacy and the Inference Economy*, 117 Nw. Univ. L. Rev. 357 (2022).

88. *See* Terry Wong, Note, *Characterizing the Harms of Compromised Genetic Information for Article III Standing in Data Breach Litigation*, 53 Colum. J.L. & Soc. Probs. 461, 505–07 (2019).

89. *See* Richard Acello, *The New Frontier of Health Care Is Here, But Will DNA Privacy be Lost?*, A.B.A. J. (June 1, 2021, 1:30 AM), https://www.abajournal.com/magazine/article/23-and-you-dna-privacy [https://perma.cc/6HNN-RGHR].

90. *See* Wong, *supra* note 88, at 476.

91. *See* Alan McQuinn & Daniel Castro, *The Costs of an Unnecessarily Stringent Federal Data Privacy Law*, Info. Tech. & Innovation Found. 12, 14 (Aug. 5, 2019), https://itif.org/sites/default/files/2019-cost-data-privacy-law.pdf [https://perma.cc/GBQ3-GPVQ].

rights to data deletion or erasure.[92] The European Union's General Data Protection Regulation includes the "right to be forgotten," which gives data subjects the "right to obtain from the controller the erasure of personal data concerning him or her without undue delay."[93] "[T]he controller shall have the obligation to erase personal data without undue delay" as long as a number of conditions are met.[94] The GDPR does not define deletion or what qualifies as full deletion or erasure of a data subject's data.[95] The statute does note that data controllers must, in "taking account of available technology and the cost of implementation, . . . take reasonable steps, including technical measures, to inform [other] controllers" processing data included in the subject's invocation of their right to erasure "of any links to, or copy or replication of, those personal data."[96] This provision recognizes that data, once transferred or made public, is difficult to fully delete from the long memory of the internet. However, this is the extent of the technical discussion in the Regulation itself.

Under the GDPR's right to be forgotten, individuals can request that their data be deleted and that links or copies of such data be deleted as well.[97] This often results in requests to publications, data depositories, and search engines, as seen in the 2014 case, *Google Spain SL v. Agencia Española de Protección de Datos*.[98] This case dealt with the General Data Protection Directive (the "Directive"), the forebear of the GDPR.[99] The claimant, in this case, Mr. Costeja González, requested that Google delete links to news articles that included outdated negative information about him.[100] After various appeals, the Court of Justice of the European Union ruled that Google was responsible as a data controller for processing the personal data of Costeja González and that Google and Google Spain were within the territorial scope of the Directive.[101] Employing a balancing test to weigh Costeja González's privacy rights against the interests of Google along with those of the public, the court ultimately ruled in favor of Costeja González.[102] Specifically, the court held that Google must "remove from the list of results displayed following a search . . . of a person's name links to web pages, published by third parties and containing information relating to that person."[103] The decision noted that Costeja González; the Spanish, Italian, and Polish Gov-

92. *See, e.g.*, David L. Hudson, Jr., *Right to be Forgotten*, FIRST AMEND. ENCYCLOPEDIA, FREE SPEECH CTR. (2017), https://mtsu.edu/first-amendment/article/1562/right-to-be-forgotten [https://perma.cc/F6Y7-77JL].

93. GDPR, *supra* note 21, art. 17(1).

94. *Id.*

95. *See id.* art. 4.

96. *Id.* art. 17, at 2.

97. *See id.*

98. Case C-131/12, Google Spain SL v. Agencia Española de Protección de Datos (AEPD), ECLI:EU:C:2014:317, ¶¶ 14–15 (May 13, 2014).

99. *See id.* ¶ 1; Council Directive 95/46, 1995 O.J. (L 281) 31 (EC).

100. *Google Spain SL*, Case C-131/12, ¶¶ 14–15.

101. *Id.* ¶¶ 42–60.

102. *Id.* ¶¶ 74, 80–81, 97–99.

103. *Id.* ¶ 88.

ernments; and the European Commission all called for Google "to withdraw from its indexes and intermediate memory information containing personal data that has been published by third parties."[104] However, the court did not pass judgment specifically on whether removal from indexes and intermediate memory would be necessary or sufficient for the removal ultimately required of Google.[105]

The *Google Spain* case was influential in determining the territoriality of European data privacy law, the boundaries of data processing and data controlling, and the right to be forgotten. However, the case also illustrates the failure of deletion as a privacy right or remedy. Here, the court called for data erasure but did not specify what erasure would entail.[106] In providing guidance on the implementation of the *Google Spain* decision, the Article 29 Working Party advised that search engine operators must, if requested, remove links to articles from public indexes but can still retain copies of those links in internal data storage and can still use those links for internal purposes.[107] The Working Party Guidelines also state that the right to deletion "does not require deletion of the link from the indexes of the search engine altogether. That is, the original information will still be accessible using other search terms, or by direct access to the publisher's original source."[108] In fact, the guidance specifically suggests that "complete deletion of the page from the indexes of the search engine" is unnecessary.[109] Thus, the right to erasure or the right to be forgotten is rather limited in scope due to the lack of clarity surrounding the legal clauses requiring deletion.

While the limitations on the right to be forgotten may seem inconsequential, this limitation could lead to individuals being unable to access the privacy protection the law attempts to provide. For example, Google's search algorithm uses a variety of factors to determine the ranking of search results that show up when a user enters a query.[110] The data from the removed links may still be used as part of Google's search algorithm, even after the links have been removed. This could influence future search results, which could negatively harm the individual. Furthermore, this limited right to deletion may not protect against privacy harms associated with the inclusion of links to negative news stories in one or more of Google's databases. Thus, it is unclear if this level of "deletion" is enough to truly protect an individual's privacy rights.

---

104. *Id.* ¶ 65.

105. *See id.* ¶ 88.

106. *See id.*

107. Article 29 Data Protection Working Party, *Guidelines on the Implementation of the Court of Justice of the European Union Judgment on "Google Spain and Inc v. Agencia Española de Protección de Datos (AEPD) and Mario Costeja González" C-131/12*, at 2, 14/EN WP 225 (Nov. 26, 2014).

108. *Id.*

109. *Id.* at 9.

110. *How Search Algorithms Work*, GOOGLE, https://www.google.com/search/howsearchworks/algorithms [https://perma.cc/GH75-UBHY].

Limitations on the effectiveness of rights to data deletion and record erasure may stem from confusion over how to quantify privacy harms (and how to remedy them) and privacy rights (and how to protect them). Nonetheless, it is important to understand the benefit and limits of data deletion as privacy remedy, as it is so commonly used today.

B. Against Data Deletion

As described above, many privacy laws now include a right to deletion or right to erasure, allowing users to advocate for their own rights by requesting deletion of their data.[111] However, this approach is flawed in multiple ways. First, on a legal basis, privacy laws that enshrine a right to deletion often fail to specifically define the steps required for a data-controlling entity to successfully delete data as required by law.[112] This creates confusion and complicates the use of data deletion as a remedy or right; it is difficult to know if or when a wrong has been righted or an individual has been able to enforce their rights.

Further, individuals often lack full understanding of their rights to deletion and are thus unable to utilize them fully. In trying to understand these data rights, most individuals who visit a website or use an app—aside from privacy lawyers or legal scholars—have only a privacy policy and perhaps a set of privacy settings to reference. Consumers not understanding privacy rights, even when—or especially when—provided to them in privacy notices, is not a new problem unique to the right to deletion.[113] Privacy notices and options, including options to request deletion, can be malformed, resulting in confusion for users regardless of the data controller's good intent (and sometimes as a result of deliberate

111. *See* California Consumer Privacy Act of 2018, Cal. Civ. Code § 1798.105 (West 2018); GDPR, *supra* note 21, art. 17. Other countries, like Germany have caselaw around privacy that suggests a right to be forgotten. *See* Bundesverfassungsgericht [BVerfG] [Federal Constitutional Court] June 3, 1980, 54 Entscheidungen des Bundesverfassungsgerichs [BVerfGE] 148, 155 (Ger.) (holding an individual had a constitutional right to informational self-determination—the right to set parameters for personal information conveyed to others).

112. As quoted above, Article seventeen of the GDPR grants individuals "the right to obtain from the controller the erasure of personal data concerning him or her without undue delay." GDPR, *supra* note 21, art. 17(1). However, the regulation does not provide much guidance to controllers who must perform that erasure. Section two of Article seventeen directs controllers to "take reasonable steps, including technical measures" to retrieve the relevant data from third parties. *Id.* art. 17(2).

113. *See generally* Benjamin Fabian, Tatiana Ermakova & Tino Lentz, *Large-Scale Readability Analysis of Privacy Policies*, Proc. Int'l Conf. on Web Intel. 18 (Aug. 2017) (analyzing the readability of nearly 50,000 website privacy policies); Pedro Giovanni Leon, Justin Cranshaw, Lorrie Faith Cranor, Jim Graves, Manoj Hastak, Blase Ur & Guzi Xu, *What Do Online Behavioral Advertising Privacy Disclosures Communicate to Users?*, Proc. 2012 ACM Workshop on Priv. Elec. Soc'y 19 (Oct. 2012) (finding that users do not read or remember disclosures about privacy in Online Behavioral Advertising); Hana Habib, Yixin Zou, Aditi Jannu, Neha Sridhar, Chelse Swoopes, Alessandro Acquisti, Lorrie Faith Cranor, Norman Sadeh & Florian Schaub, *An Empirical Analysis of Data Deletion and Opt-Out Choices on 150 Websites*, Proc. Fifteenth USENIX Symp. on Usable Priv. & Sec. 387 (Aug. 2019), https://www.usenix.org/system/files/soups2019-habib.pdf [https://perma.cc/L7LS-N7CR] (identifying issues with websites' opt out and data deletion options).

obfuscation).[114]

Specifically, in regard to the right to deletion, empirical research has shown mixed results when giving individuals information about rights to deletion. Users have expectations about data deletion that are not borne out due to the technical realities of the process.[115] For example, users may believe that requesting data deletion from a platform also means the company will delete the data from its private storage, which may not be the case.[116] Deleting content from platforms can even backfire for individuals. Public figures may find that their choices to delete certain posts may actually attract more attention to said posts (a sort of deletion-related variant of the Streisand Effect),[117] particularly as a number of accounts and mechanisms have popped up specifically to track content deletion.[118]

Furthermore, deletion might not even be a right that individuals want. Some research has shown that individuals do not want their content to fade over time but rather find value in being able to access the content they posted, even content that is old and forgotten.[119] Thus, there may be a mismatch between what the right to deletion seeks to accomplish (protection of privacy by removing data) and what individuals actually want (protection of privacy while keeping data).

Even when consumers are able to access and understand privacy notices and options like deletion requests, the use of deletion as a mechanism to protect privacy can be flawed due to the technical realities of machine learning systems, including the unacknowledged persistence of the algorithmic shadow. The practical implications of the data deletion remedy stated in the law might not be clear to the people implementing the law on the ground or to users relying on the right to data deletion.

---

114. Ari Ezra Waldman, *Cognitive Biases, Dark Patterns, and the 'Privacy Paradox*,' 31 CURRENT OP. PSYCH. 105, 106–08 (2020).

115. *See* Mohsen Minaei, Mainack Mondal & Aniket Kate, *Empirical Understanding of Deletion Privacy: Experiences, Expectations, and Measures*, 31ST USENIX SEC. SYMP. (Aug. 2022), https://www.usenix.org/system/files/sec22summer_minaei.pdf [https://perma.cc/7ZAH-HLCQ].

116. *See* Ambar Murillo, Andreas Kramm, Sebastian Schnorf & Alexander De Luca, *"If I Press Delete, It's Gone"—User Understanding of Online Data Deletion and Expiration*, PROC. FOURTEENTH USENIX SYMP. ON USABLE PRIV. & SEC. 329, 334 (Aug. 2018), https://www.usenix.org/system/files/conference/soups2018/soups2018-murillo.pdf [https://perma.cc/NB3R-329S].

117. *See* Mario Cacciottolo, *The Streisand Effect: When Censorship Backfires*, BBC NEWS (June 15, 2012), https://www.bbc.com/news/uk-18458567 [https://perma.cc/UV94-226B].

118. *See* Mohsen Minaei, S Chandra Mouli, Mainack Mondal, Bruno Ribeiro & Aniket Kate, *Deceptive Deletions for Protecting Withdrawn Posts on Social Media Platforms*, PROC. 27TH ANN. NETWORK & DISTRIBUTED SYS. SEC. SYMP. (Feb. 2021), https://www.ndss-symposium.org/wp-content/uploads/ndss2021_3A-2_23139_paper.pdf [https://perma.cc/K8RC-UNJQ].

119. *See* Lujo Bauer, Lorrie Faith Cranor, Saranga Komanduri, Michelle L. Mazurek, Michael K. Reiter, Manya Sleeper & Blase Ur, *The Post Anachronism: The Temporal Dimension of Facebook Privacy*, PROC. 12TH ACM WORKSHOP ON PRIV. ELEC. SOC'Y 1, 9 (Nov. 2013), https://dl.acm.org/doi/pdf/10.1145/2517840.2517859 [https://perma.cc/HSU3-FJHD].

Machine learning further complicates the question of data deletion. This is particularly apparent due to the persistent presence of the algorithmic shadow. As explained above, programmers find or create a set of data to be used as training data in contemporary machine learning.[120] Humans collect, create, clean, codify, and otherwise prepare this training data for use in building the machine learning system.[121] Humans then create an algorithm and ask a computer to run that algorithm on the training data set to detect patterns and produce a machine learning model that can then be used to solve problems, generate predictions, or otherwise achieve a programmer's goal.[122]

Data deletion does not eliminate the algorithmic shadow. Deleting data from the training data set (the initial data set fed into the computer to train and produce a machine learning model) has no impact on an already trained model. This means that an imprint from the individual user will still remain, though all "data" has been deleted. The algorithmic shadow persists, which means some measure of privacy loss cannot be undone through the act of data deletion. Thus, there is a disconnect between the practical reality of data deletion and its supposed goals.

V. ALGORITHMIC DISGORGEMENT

Recently, the Federal Trade Commission (FTC) introduced a new potential remedy for consumer protection violations caused by companies' use of improperly obtained data to train machine learning algorithms.[123] This new enforcement mechanism would require companies not only to delete improperly collected data (something that is routine in FTC enforcement) but also to delete any machine learning models trained on that data (something previously unseen in FTC judgments).[124] Requiring model deletion could effectively require companies to "roll back" their models to the time before the improperly obtained data was introduced and thus retrain them without the deleted data.

This rather radical enforcement mechanism entails requiring companies to delete machine learning models or algorithms trained on the misbegotten data—a mechanism journalist Kate Kaye titled "algorithmic destruction."[125] FTC Commissioner Rebecca Kelly Slaughter described it in perhaps less explosive terms as "an innovative disgorgement remedy."[126] Whether we call it algorithmic destruction, algorithmic disgorgement, or perhaps simply machine learning model deletion, this new

120. *Supra* text accompanying notes 44–46.
121. *Id.*
122. *Supra* text accompanying notes 47–48.
123. *See* Kaye, *supra* note 11.
124. *See id.*
125. *Id.*
126. Rebecca Kelly Slaughter, Commissioner, Fed. Trade Comm'n, Keynote Address at the Future of Privacy Forum's Annual Privacy Papers for Policymakers Event: Protecting Consumer Privacy in a Time of Crisis 2 (Feb. 10, 2021) [hereinafter Comm'r Slaughter Keynote Address], https://www.ftc.gov/system/files/documents/public_statements/1587283/fpf_opening_remarks_210_.pdf [https://perma.cc/2LBE-P2TX].

remedy is revolutionary in scope and deserves serious consideration, particularly in light of the existing problems with data deletion and the persistence of algorithmic shadows.

A. The Radical FTC

There is a saying in journalism: one is an accident, two is a coincidence, and three is a trend.[127] Thus far, three major FTC cases have included some variation of algorithmic disgorgement as a remedy.[128] This lasting trend, backed by statements from FTC Commissioners, shows what is likely becoming a new push from the FTC to establish algorithmic disgorgement as a routine privacy remedy.[129]

The FTC first introduced the concept of model deletion as a remedy in its enforcement decree against Cambridge Analytica in 2019 after the company engaged in deceptive practices to exploit its access to users on the social media platform Facebook and collect and use data without user consent.[130] The FTC ordered Cambridge Analytica to "[d]elete or destroy all Covered Information collected from consumers through [its application], and any information or work product, including any algorithms or equations, that originated, in whole or in part, from this Covered Information."[131] Here, the destruction of "work product, including any algorithms or equations[ ] that originated" from the user information in question could be read to include any machine learning models that were created with training data that included ill-gotten data, as well as any resulting models that were refined after being fed ill-gotten user data.[132]

While the 2019 *Cambridge Analytica* case was likely the first time the FTC employed this novel remedy, the idea of algorithmic disgorgement as a remedy really gained traction with the FTC's 2021 enforcement action against Everalbum, Inc. (Ever), a photo storage and organization application company.[133] Ever had created a facial recognition tool for its application, giving users a pop-up message that claimed to allow users to choose whether they wished to turn the face recognition feature on or

127. *See, e.g.*, Renee Montagne, *Two Is a Coincidence, Three Is a Trend*, NPR (Jan. 29, 2013), https://www.npr.org/2013/01/29/170535707/two-is-a-coincidence-threes-a-trend [https://perma.cc/8U2T-XPE2]; Jeffrey Lewis, *It's Not as Easy as 1-2-3*, Foreign Pol'y (Aug. 1, 2012), https://foreignpolicy.com/2012/08/01/its-not-as-easy-as-1-2-3 [https://perma.cc/SA3U-LEDN] (using the phrase in response to behavior from the State Department).

128. *See* Kaye, *supra* note 11.

129. *See id.*

130. *See* Complaint at 1, 12–13, Cambridge Analytica, LLC, FTC File No. 1823107, No. 9383 (F.T.C. July 22, 2019), https://www.ftc.gov/system/files/documents/cases/182_3107_cambridge_analytica_administrative_complaint_7-24-19.pdf [https://perma.cc/2FU6-EN9G].

131. Final Order at 4, Cambridge Analytica, LLC, FTC File No. 1823107, No. 9383 (F.T.C. Nov. 25, 2019), https://www.ftc.gov/system/files/documents/cases/d09389_comm_final_orderpublic.pdf [https://perma.cc/6G84-QEPF].

132. *See id.*

133. *See* Complaint ¶ 3, Everalbum, Inc., FTC File No. 1923172, No. C-4743 (F.T.C. May 6, 2021) [hereinafter Ever Complaint], https://www.ftc.gov/system/files/documents/cases/1923172_-_everalbum_complaint_final.pdf [https://perma.cc/9B83-KPLP].

off.[134] However, the FTC found that Ever did not actually give users a meaningful choice.[135] For users in some jurisdictions, the face recognition feature was enabled by default, and there was no ability for users to disable it.[136] Ever used these users' photos to train its machine learning models for face recognition, arguably without informed consent from users.[137]

Ever also collected other face images from its users and combined them with publicly available datasets to create a variety of datasets used to develop facial recognition technology.[138] Ever did have an option for users to deactivate their accounts, stating in multiple settings that the company would delete the user's information upon deactivation.[139] However, the FTC found that Ever was not deleting photos and videos upon deactivation but was storing them instead.[140]

So far, the case sounds relatively routine for an FTC enforcement action involving privacy. In the absence of federal privacy law, the FTC enforces privacy actions based on its § 5 authority to regulate unfair and misleading practices to consumers.[141] Here, the identified issues were Ever's misrepresentations involving the ability for users to opt in or opt out of face recognition and misrepresentations that Ever would delete user photos and videos upon deactivation.[142]

What made this enforcement action radical was the new remedy introduced by the FTC as part of the settlement order. In addition to requiring Ever to delete the photos and videos of users who had requested their accounts deactivated, the FTC also required the company to "delete or destroy all Face Embeddings derived from Biometric Information Respondent collected from Users who have not . . . consent[ed]" and to "delete or destroy any Affected Work Product."[143] Elsewhere in the Order, the FTC defined "Face Embedding" as "data, such as a numeric vector, derived in whole or in part from an image of an individual's face."[144] It is important to note that such Face Embedding could itself be the product of a machine learning system that was trained on face image data, potentially face image data from non-consenting users. Thus, asking for

134. *See id.* ¶¶ 5–7.
135. *See id.* ¶¶ 10, 23–24.
136. *See id.* ¶ 10.
137. *See id.* ¶ 11.
138. *See id.* ¶ 12.
139. *See id.* ¶¶ 17–18, 20.
140. *See id.* ¶¶ 22, 26.
141. *See* Federal Trade Commission Act, 15 U.S.C. § 45 (granting the FTC power to regulate "unfair or deceptive acts or practices in or affecting commerce"); *see also* Daniel J. Solove & Woodrow Hartzog, *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. 583, 598–99 (2014) (quoting the Act as the "primary source of authority for FTC privacy enforcement").
142. *See* Ever Complaint, *supra* note 133, ¶¶ 23–27.
143. Decision and Order at 4–5, Everalbum, Inc., FTC File No. 1923172, No. C-4743 (F.T.C. May 6, 2021), https://www.ftc.gov/system/files/documents/cases/1923172_-_everalbum_decision_final.pdf [https://perma.cc/JD4K-XYJT].
144. *Id.* at 3.

the deletion of the Face Embeddings could also mean the deletion of some machine learning models.

The Commission defined "Affected Work Product" even more clearly: "any models or algorithms developed in whole or in part using Biometric Information Respondent collected from Users of the 'Ever' mobile application."[145] This is one of the most revolutionary calls the FTC has ever made regarding AI. Essentially, as a remedy for privacy violations suffered by individuals whose data was used without their consent, the FTC required Ever to delete any models or algorithms that had been developed using the misbegotten user data.

This kind of algorithmic destruction or algorithmic disgorgement is a radical remedy. The model deletion remedy goes much further than the traditional remedy of asking companies to delete users' data. Remember that the input of any data into a machine learning system helps shape and refine the machine learning model, thus leaving a persisting algorithmic shadow. While traditional data deletion would not correct the harms of the persistent algorithmic shadow, model deletion has the potential to solve such harms. By requiring companies to delete the models trained on user data, the FTC essentially eliminates the algorithmic shadow. The company would have to retrain its models based on a data set that does not include and was not influenced by the shadow of the individual's data.

More recently, in 2022, the FTC employed algorithmic disgorgement once again in its settlement with WW International, Inc. (WW International), formerly known as Weight Watchers, and a subsidiary called Kurbo, Inc. (Kurbo).[146] WW International and Kurbo had, among other legal violations, collected the personal information of children under thirteen without proper parental information, in defiance of the Children's Online Privacy Protection Act.[147] The final settlement ordered the weight loss company to delete or destroy any "Affected Work Product," defined as "any models or algorithms developed in whole or in part" using the ill-gotten personal information at controversy in this case.[148] Here again, the FTC employed algorithmic disgorgement as a remedy for a privacy violation, acknowledging that simple data deletion is not enough to remedy such violations where companies use misbegotten data to develop machine learning models.

145. *Id.* at 2.

146. *See* Stipulated Order for Permanent Injunction, Civil Penalty Judgment, and Other Relief at 8, United States v. Kurbo, Inc., Case No. 3:22-cv-00946-TSH (N.D. Cal. Mar. 3, 2022), https://www.ftc.gov/system/files/ftc_gov/pdf/wwkurbostipulatedorder.pdf [https://perma.cc/PR4Q-TMXJ] [hereinafter Kurbo Order].

147. *See* Complaint for Permanent Injunction, Civil Penalties, and Other Equitable Relief ¶¶ 1–2, 57, United States v. Kurbo, Inc., Case No. 22-CV-946 (N.D. Cal. Feb. 16, 2022), https://www.ftc.gov/system/files/ftc_gov/pdf/filed_complaint.pdf [https://perma.cc/BE3A-A3HJ].

148. Kurbo Order, *supra* note 146, at 2, 8.

B. Algorithmic Disgorgement as Remedy

Algorithmic disgorgement acts as a remedy by preventing unjust enrichment and eliminating the ill-gotten gains of privacy-violating companies.[149] This concept pulls from a number of legal doctrines, including unjust enrichment and disgorgement from contract law and remedies and the fruit of the poisonous tree analysis in criminal procedure.[150] Scholars have discussed these questions before. For example, Bernard Chao has written on the concept of privacy losses as wrongful gains,[151] a theory which could be extended to justify the use of algorithmic disgorgement as a rather traditional remedy for unjust gain.

In a statement on the *Everalbum* case, former FTC Commissioner Rohit Chopra celebrated the decision, in part because the FTC required the malfeasant company to "forfeit the fruits of its deception."[152] This phrasing directly links the deceptive acts of the company (clearly under the FTC's § 5 authority) with the remedy of algorithmic disgorgement.

Further justifying algorithmic disgorgement, Commissioner Slaughter described the remedy as being similar to monetary disgorgement in a keynote speech at the Future of Privacy Forum in 2021:

> We routinely obtain disgorgement of ill-gotten monetary gains when consumers pay for a product that is marketed deceptively. Everalbum shows how we can apply this principle to privacy cases where companies collect and use consumers' data in unlawful ways: we should require violators to disgorge not only the ill-gotten data[ ] but also the benefits—here, the algorithms—generated from that data.[153]

The message is clear: the FTC views algorithmic disgorgement as a remedy in line with both the FTC's legal enforcement authority as well as existing FTC precedent.

In a law review article, Commissioner Slaughter and co-authors Janice Kopec and Mohamad Batal put it more simply: "The premise is simple: when companies collect data illegally, they should not be able to profit from either the data or any algorithm developed using it."[154] Algorithmic disgorgement, then, could achieve a similar intended purpose as tradi-

---

149. *See* Heather Federman, *Tainted Fruit: Disgorgement of Data from the FTC and Beyond*, IAPP: The Privacy Advisor (Apr. 27, 2021), https://iapp.org/news/a/tainted-fruit-disgorgement-of-data-from-the-ftc-and-beyond [https://perma.cc/ZWS5-2KMP].

150. *See id.*

151. *See* Bernard H. Chao, *Privacy Losses as Wrongful Gains*, 106 Iowa L. Rev. 555, 557 (2021).

152. Statement of Commissioner Rohit Chopra In the Matter of Everalbum and Paravision, Comm'n File No. 1923172 (Jan. 8, 2021), https://www.ftc.gov/system/files/documents/public_statements/1585858/updated_final_chopra_statement_on_everalbum_for_circulation.pdf [https://perma.cc/53J6-PYTP].

153. Comm'r Slaughter Keynote Address, *supra* note 126, at 2.

154. Rebecca Kelly Slaughter, Janice Kopec & Mohamad Batal, *Algorithms and Economic Justice: A Taxonomy of Harms and a Path Forward for the Federal Trade Commission*, 23 Yale J.L. & Tech. (Special Issue) 1, 39 (2021), https://yjolt.org/sites/default/files/23_yale_j.l._tech._special_issue_1.pdf [https://perma.cc/PX84-FB2C].

tional disgorgement—prevention of unjust enrichment for legal wrongdoers.

Because data deletion does not solve the privacy loss that remains with the persistent algorithmic shadow, algorithmic disgorgement may be a more effective remedy. Data deletion as a remedy does not make whole the privacy victim whose information still has a lasting imprint on the resulting machine learning model; the algorithmic shadow of one's information reflects a form of privacy loss that data deletion does not remedy. In contrast, algorithmic disgorgement would erase the algorithmic shadow, and the privacy victim's information would no longer have an imprint on the resulting machine learning model. Any machine learning systems that had been developed with the use of the data subject's information would be deleted, which at least ventures closer to making the victim whole, as no part of them (or their data) is being actively misused anymore.

C. Algorithmic Disgorgement as Right

The impact on individuals and the ability of this remedy to right some algorithmic wrongs has been missing so far from the analysis of this new remedy. While the FTC appears to believe algorithmic disgorgement ought to be a privacy remedy, neither the FTC nor legislatures have taken the next leap: establishing algorithmic disgorgement as a privacy right.

Algorithmic disgorgement creates a financial penalty for companies, but it also provides relief for users who have had their data misused. Algorithmic disgorgement succeeds where data deletion fails in preventing the harms of the persistent algorithmic shadow. While rather blunt, algorithmic disgorgement goes further than data deletion in solving some of the issues this Article raises in understanding the role of algorithmic shadows and the related persistent harms. Algorithmic disgorgement, then, can succeed as a legal remedy, at least on a theoretical basis (though the practical realities of compliance may make it untenable).

However, it is one thing for enforcement agencies to use algorithmic disgorgement as a penalty; it is another thing entirely to introduce the right to request algorithmic disgorgement as a positive right for individuals. So far, no privacy law includes such a right. Introducing algorithmic disgorgement as a privacy right in privacy law would increase the potential compliance burdens on companies but could also increase the deterrent effect, raising the risks to such an extent that companies would be encouraged to be even more careful with their use of data and machine learning. Machine learning in general lacks clear regulatory or legal boundaries,[155] but further legislative or regulatory actions are likely.

155. *See generally*, Mehtab Khan & Alex Hanna, *The Subjects and Stages of AI Dataset Development: A Framework for Dataset Accountability*, 19 Ohio St. Tech. L.J. (forthcoming 2023); Mark A. Lemley & Bryan Casey, *Fair Learning*, 99 Texas L. Rev. 743 (2021).

Ultimately, if we believe in the seriousness of problems like the harms of the persistent algorithmic shadow, as well as other group privacy harms, indirect privacy harms, and proxy privacy harms, then we should consider algorithmic disgorgement as a privacy right to be included in privacy laws. However, the challenge in focusing on privacy rights is generally due to the fact that it is increasingly difficult for individuals to know the extent of the violations that relate to their privacy.[156] In today's complex data ecosystem, individuals are at a huge information disadvantage in comparison to the large corporations and governments that have access to multiple streams of data collection and analysis of individuals' data.[157] Thus, relying on rights that place the burden on individuals for enforcement may make it difficult for any individual to protect their own privacy.

This is particularly apparent when we discuss issues like the improper use of one's data in machine learning systems or the harms of the algorithmic shadow. While these harms are very real, it is also quite possible that an individual may never have knowledge that their data is being used to develop a machine learning system or that a persistent shadow of their data exists even after their data has been deleted.

D. Against Algorithmic Disgorgement

Algorithmic disgorgement as a remedy is relatively novel, and there are still unresolved issues about what it means to suffer a privacy harm, to remedy a harm, to enshrine a right to protect against harms, and so on. For example, while algorithmic disgorgement may prevent a company from unjustly enriching itself at the expense of privacy victims, algorithmic disgorgement may not do much to correct the harms incurred by victims whose privacy rights have been violated. The psychological harms of knowing one's data has been released in a data breach,[158] for example, may not be ameliorated at all by the knowledge that the breaching company did not profit from your data.

However, legal issues aside, the chief problem with algorithmic disgorgement is that it is not a particularly practicable solution. Compliance with algorithmic disgorgement orders may be costly, which could harm smaller companies and potentially chill innovation.[159] It is possible that some developers will look at the risk of being forced to delete their trained models and decide that the risk and costs are too great to proceed. Thus, algorithmic disgorgement could create economic harm for the technology industry, which could, in turn, lead to greater social and policy harms related to decreased national strength in technology. For example,

156. *See* Peter J. van de Waerdt, *Information Asymmetries: Recognizing the Limits of the GDPR on the Data-Driven Market*, 38 Comput. L. & Sec. Rev. 1, 2 (2020).

157. *See id.*

158. *See* Ido Kilovaty, *Psychological Data Breach Harms*, 23 N.C. J.L. & Tech. 1, 4 (2021); Citron & Solove, *supra* note 77, at 841.

159. *See* Tiffany C. Li, *Post-Pandemic Privacy Law*, 70 Am. U. L. Rev. 101, 139 (2021).

losing out on AI innovation could mean the United States has less of a role in shaping global norms surrounding AI and privacy, and thus less of an ability to advocate for democratic ideals with respect to AI, privacy, and the future of technology.[160]

If deletion of machine learning training models becomes an established and routine legal consequence for the deployment of a bad model or a bad AI system in general, this might harm small startups and discourage new market entrants in technology industries.[161] These policies could counterintuitively help further entrench outsized market power from Big Tech companies, as they may be the only ones able to risk the financial consequences of launching new AI-driven projects. Thus, there may need to be limits to protect nonprofits, small startups, and other parties with limited resources for compliance. These protections should be built into any new laws that include model deletion as a consequence for misuse of data or other legal violations.

VI. CONCLUSION

The persistence of the algorithmic shadow is one small reminder that there are large gaps in much of our current privacy law. Even the most recent privacy laws, as well as proposed AI laws, fail to address algorithmic shadow harms. Most of these laws similarly fail to adequately address group privacy harms, proxy privacy harms, or indirect privacy harms. Other gaps persist, including a lack of acknowledgment of critical dimensions of privacy harms and privacy rights and the impact of digital inequity, and the role of privacy as a civil right.[162] While data deletion does provide some measure of right and remedy for individuals, it is ultimately insufficient to meet its goals.[163] Other proposed solutions like algorithmic disgorgement may be a better fit to solve some of the problems that have not yet been addressed by law, including the harms of the algorithmic shadow. Though perhaps none of these solutions are sufficient or even effective, and other algorithmic rights or remedies ought to be considered.

More research is needed on the feasibility of model deletion, as well as research on technical realities and economic costs for compliance. This is certainly an area where interdisciplinary research is critical in order to drive conversations forward in a meaningful way. Organizations, including the FTC, that consider algorithmic disgorgement as a right or remedy must encourage and fund research on this question before lawyers and policymakers treat algorithmic disgorgement as a routine remedy and tool for privacy enforcement.

---

160. *See id.* at 137–39.

161. See Lemley and Casey, *supra* note 155, arguing that limiting the bounds of the fair use defense for machine learning developers or users accused of copyright infringement would chill innovation.

162. *See* Tiffany C. Li, *Privacy as/and Civil Rights*, BERKELEY TECH. L.J. 1265 (2022).

163. *See, e.g.*, *supra* discussion in note 112.

As the FTC is currently the de facto privacy commission for the United States[164] and the agency that first pushed model deletion, it will also be necessary for the FTC to take issues like algorithmic shadow persistence into account. Currently, in the absence of larger federal privacy law, much of U.S. privacy enforcement has fallen to the FTC, which has, in turn, built a body of FTC common law.[165] With the leading role the FTC has taken in privacy enforcement, it is especially important that the agency carefully consider the practical and technical dimensions of model deletion as an enforcement mechanism.

Ultimately, the debate around deletion should not center around data deletion, algorithmic deletion, or model deletion. What must be deleted is the siloed nature of scholarship and policymaking on matters of AI. We must properly examine issues like algorithmic destruction as a legal remedy while there is still time before we fall victim to the destructive effects of unrestrained AI development and applications. AI will only become more important in the future, and it is imperative that we collaborate now to create frameworks for protecting our rights and interests in a techno-social future that is hopefully better for all of us.

164. *See* CHRIS JAY HOOFNAGLE, FEDERAL TRADE COMMISSION PRIVACY LAW AND POLICY 145, 192 (2016). *See generally* Steven Hetcher, *The De Facto Federal Privacy Commission*, 19 J. MARSHALL J. COMPUT. & INFO. L. 109 (2000).

165. *See* Solove & Hartzog, *supra* note 141, at 583.